

Comparative performance of an AI-based digital pathology tool and genomic signatures in early ER+/HER2- breast cancer

V. Aubert¹, V. Gaury¹, I. Garberis², Z. Vaquette¹, E. Hocquet¹, D. Almaraz-Klippel³, F. Daidj², D. Drubay⁴, D. Jacobs¹, N. Arfaoui¹, L. Guillou¹, D. Lin¹, J. Guillon¹, C. Barcenas³, F. André^{2,5}, S. Krishnamurthy^{6*}, M. Lacroix-Triki^{7*}

1: Owkin, Paris, France; 2: INSERM U981, Gustave Roussy, Paris-Saclay University, Villejuif, France; 3: Department of Breast Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, TX, USA; 4: Gustave Roussy, Office of Biostatistics and Epidemiology, Université Paris-Saclay, Villejuif, France Inserm, Université Paris-Saclay, CESP U1018, Oncostat, labeled Ligue Contre le Cancer, Villejuif, France; 5: Department of Cancer Medicine, Gustave Roussy, Paris-Saclay University, Villejuif, France; 6: Department of Anatomical Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA; 7: Department of Pathology, Gustave Roussy, Paris-Saclay University, Villejuif, France — * These authors contributed equally to this work.

Background

→ Prognostic tools, such as genomic signatures, are important for **guiding adjuvant therapy in early breast cancer (EBC)**, particularly for patients with estrogen receptor-positive, HER2-negative (**ER+/HER2-**) tumors, where optimizing the treatment strategy, like making a de-escalation decision, remains a clinical challenge [1].

→ Recent advances in artificial intelligence (AI) applied to whole-slide images (WSI) have opened new possibilities for capturing prognostic information directly from **routinely available H&E slides** of invasive breast tumors.

→ This may offer a scalable, cost-effective, and accessible approach to **personalized risk assessment**.

Objective

→ Evaluate the prognostic performance of **RlarsRisk BC**, an AI digital pathology-based tool developed to estimate the risk of distant recurrence, in comparison with **EndoPredict (EP)** [3] and **OncoType DX (ODX)** [4], using five-years distant recurrence-free survival (dRFS) as the primary endpoint.

Data

Cohort	MDA (MD Anderson, USA)	GR (Gustave Roussy, FR)
n_{TOTAL} (n_{events})	154 (42)	381 (7)
Age		
<50 y.o.	61 39.6%	108 28.35%
>50 y.o.	93 60.4%	273 71.65%
Histological grade		
Grade 1	35 22.7%	23 6.04%
Grade 2	89 57.8%	277 72.7%
Grade 3	30 19.5%	81 21.3%
Nodal status		
N+	154 100%	283 74.3%
N-	0 0%	98 25.7%
Tumor size		
pT1	99 64.3%	248 65.1%
pT2/3/4	55 35.7%	133 34.9%
mean SD(mm)	19.59 ± 12.29	19.83 ± 9.41
Adjuvant treatment		
ET+ only	154 100%	180 47.2%
ET + chemo	N/A	201 52.8%
Routine test result*		
Low Risk	96 62.3%	128 32.6%
High Risk	58 37.7%	257 67.4%

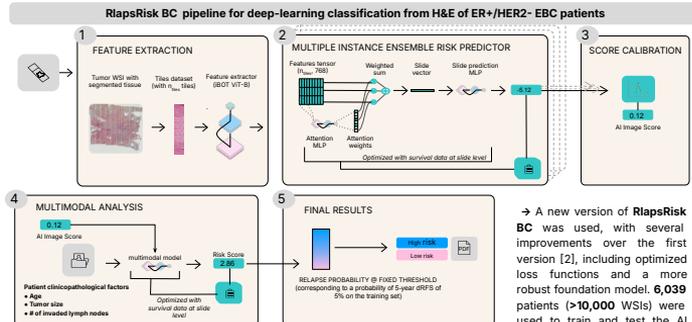
→ Two independent retrospective cohorts of ER+/HER2- EBC patients who underwent genomic testing as part of their **routine clinical care**.

→ For both cohorts, the median clinical follow-up is **more than 5 years** (MDA: 67.81 months; GR: 61.51 months).

→ Diagnostic H&E-stained slides from **surgical resection specimens** were digitized at x20.

*ODX on MDA, EP on GR
**Endocrine Therapy

Method



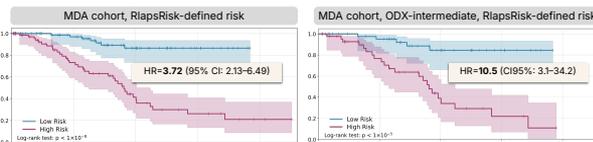
→ For **ODX**, patients were considered high risk based on standard thresholds: a recurrence score (RS) ≥25 for postmenopausal patients and ≥16 for premenopausal patients [5], for the intermediate category: a RS from 11 to 25 [6].

→ For **EP**, we used the validated EpiCn cutoff of 3.3 to define the high risk subgroup [7].

→ A new version of **RlarsRisk BC** was used, with several improvements over the first version [2], including optimized loss functions and a more robust foundation model. **6,039** patients (>10,000 WSIs) were used to train and test the AI model, which was **locked and integrated into a product before results generation** on the study cohorts.

Results

Kaplan Meier Analyses

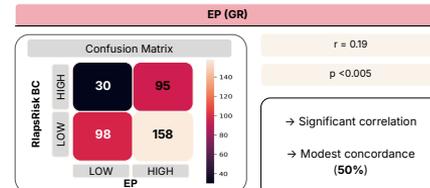
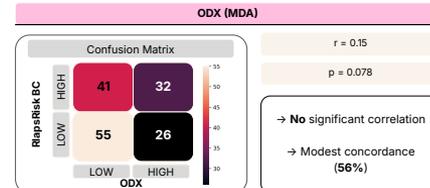


→ 50% of ODX-intermediate classified low risk by RlarsRisk BC

→ Potential for enhanced decision support

Results

Comparison metrics



Prognostic performances metrics

Metric	RlarsRisk BC (MDA)	ODX (MDA)	RlarsRisk BC (GR)	EP (GR)
AUC	0.78	0.53	0.73	0.57
% classified as low-risk	52.6%	62.3%	67%	34%
% distant recurrence events in low-risk group at 5y	4.7%	19.8%	0.4%	1.7%
% distant recurrence events in high-risk group at 5y	34.7%	16%	3.7%	1.4%

In both settings:

→ Stronger discriminative power

→ Better alignment with observed risk

RlarsRisk BC high-risk classification captured the majority of observed events in both cohorts: 6 of 7 in GR and 35 of 42 in MDA.

Conclusion

These findings suggest that RlarsRisk BC may support treatment de-escalation by safely expanding the proportion of patients classified as low risk who experience excellent outcomes within their risk group, while also improving identification of patients with a genuinely higher likelihood of recurrence.

This ability to **improve risk stratification** makes it a promising tissue-efficient alternative or complementary option to current genomic assays such as ODX and EP in ER+/HER2- early breast cancer.