SUPPLEMENTARY MATERIALS

Supplementary methods

Patients

The ShortHER phase III trial randomized 1253 patients with HER2-positive BC from 2007 to 2013 to receive adjuvant anthracycline and taxane-containing chemotherapy combined with trastuzumab for 9 weeks or 1 year(1). After randomization, surveillance included clinical examination, complete blood chemistry, CEA, CA 15-3, chest radiogram and liver ultrasound every 6 months for five years and yearly thereafter. Mammography was required every 12 months. Median follow up for the cohort of 437 patients with available gene expression data was 8.6 years (95% CI 8.4-8.8).

The CherLOB phase II trial randomized 121 HER2-positive breast cancer patients from 2006 to 2010 to receive neoadjuvant anthracycline and taxane-based chemotherapy plus either trastuzumab, lapatinib or both. Adjuvant treatment was left at physician's discretion(2). Median follow up for the 84 patients with available gene expression data was 9.1 years (95% CI 8.4-9.7).

The two Institutional cohorts included consecutive patients diagnosed with non-metastatic HER2-positive breast cancer with available gene expression data. Patients were treated from 2005 to 2018 with neoadjuvant chemotherapy and trastuzumab-based regimens as per standard clinical practice, adjuvant treatment included anti-HER2 for up to 1 year and endocrine therapy for at least 5 years if HR-positive. Median follow up was: 7.4 years (95% CI 6.8-8.0, Barcelona) and 7.2 years (95% CI 6.4-7.9, Padova).

The main clinicopathological, treatment, and outcome data were collected from the original study databases and implemented whenever necessary for clinical study cohorts. The timing and type of disease-free survival events (locoregional relapse, second primary tumor, distant relapse, or death without any prior event) were collected if not available in the original study databases. Information regarding the site of distant relapse as the first event was acquired.

Intrinsic subtyping and HER2DX score

In the ShortHER, Hospital Clinic, and Veneto Institute of Oncology cohorts, gene expression analysis was performed on formalin-fixed paraffin-embedded primary tumor samples at the Translational Genomics and Targeted Therapies in Solid Tumours at IDIBAPS. For the ShortHER cohort, the surgical

specimens were analyzed, whereas for patients undergoing neoadjuvant therapy, the baseline core biopsy was processed.

PAM50 analysis: a minimum of ~125 ng of total RNA was used to measure the expression of 50 PAM50 predictor genes and five housekeeping genes (GAPD, PUM1, ACTB, RPLP0, and PSMC4) using the nCounter platform (NanoString Technologies, Seattle, USA).

For the CherLOB samples, research-based PAM50 and HER2DX were applied to publicly available microarray data (GSE66399) (3), as previously described. (4).

Statistical analysis

In the analyses that considered distant relapse at any site (first event) as the event of interest, the following events were treated as competing risks: second primary tumor, locoregional, and death. In the analyses that considered distant relapse at a specific site (first event) as the event of interest, the following events were treated as competing risks: second primary tumor, locoregional relapse, distant relapse at a different site, and death. Patients who had not experienced an event of interest or any of the competing risk events at the last follow-up date were censored at that time. The sites of distant relapse were classified as follows: brain, brain-only (in the absence of other disease sites), lung, bone, bone-only, liver, and others. Patients who experienced concomitant locoregional and distant relapses were considered to have distant relapse as their main event. The baseline date was the date of randomization for the ShortHER and CherLOB trials, and the date of initial diagnosis for the two institutional cohorts.

Supplementary results

Patients' characteristics

A significantly different distribution of HER2DX pCR-scores across intrinsic subtypes was observed: LumA and LumB subtypes were enriched for low HER2DX pCR-score cases (70.8% and 64.8%, respectively), whereas HER2-E and Basal-like subtypes were enriched for high HER2DX pCR-score cases (55.2% and 67.4%, respectively; $p<0.001$; **Supplementary Figure 1**).

Cumulative incidence of distant relapse by HER2DX and correction by intrinsic subtype

The impact of HER2DX pCR-score on the pattern of distant relapse was evaluated. The results are presented in **Supplementary Table 1 and Supplementary Table 2.**

A high vs. low pCR-score was associated with a lower cumulative incidence of any distant metastasis as the first event (5-yr and 10-yr rates of 5.4% and 7.3% for the high pCR-score group vs. 8.2% and 15.7% for the low pCR-score group, Gray's p=0.036, subHR 0.53, 95% CI 0.29-0.70, p=0.042). This result might be partly explained by the prognostic role of pCR at the individual level, although suboptimal. Moreover, the lower incidence of any distant relapse in the high pCR-score group in our analysis was, at least in part, driven by a lower incidence of bone relapse (5-yr and 10-yr rates of 2.1% and 2.6% for the high pCR-score group vs. 4.6% and 11.2% for the low pCR-score group, Gray's p=0.013, subHR 0.28, 955 CI 0.11-0.70, p=0.007). Similar results were observed for bone-only diseases. This observation is consistent with the high pCR-score tracking of tumors with low expression of luminal features. When the competing risk regression models were corrected by intrinsic subtype, both the pCR-score and intrinsic subtype maintained a significant association with any distant relapse, whereas only the pCR-score was significantly associated with bone-only disease (**Supplementary Table 2**).

Supplementary Table 1. Cumulative incidence rates at 3, 5 and 10 years for any distant relapse as first event and site-specific distant relapse as first event according to HER2DX pCR-score.

Abbreviations: pCR, pathologic complete response, p, p-value; yr, years.

		Any distant relapse	Site-specific distant relapse					
			Brain	Brain only	Lung	Bone	Bone only	Liver
HER2DX pCR-score: low	3-yr	6.1%	1.5%	0.5%	1.0%	3.1%	2.0%	2.0%
	5-yr	8.2%	1.5%	0.5%	1.5%	4.6%	3.1%	3.1%
	10-yr	15.7%	1.5%	0.5%	2.3%	11.2%	7.8%	5.7%
HER2DX pCR-score: medium	3-yr	7.9%	2.5%	1.5%	1.5%	2.5%	1.0%	3.5%
	5-yr	11.9%	4.0%	2.0%	3.0%	4.5%	1.0%	4.4%
	10-yr	16.1%	4.0%	2.0%	4.1%	5.0%	1.0%	7.6%
HER2DX pCR-score: high	3-yr	3.3%	1.2%	0.8%	0.8%	1.3%	0%	1.3%
	5-yr	5.4%	1.2%	0.8%	2.1%	2.1%	0.4%	2.1%
	10-yr	7.3%	1.2%	0.8%	2.6%	2.6%	0.9%	2.1%
Gray's p		*0.036*	*0.119*	*0.335*	*0.507*	*0.013*	*<0.001*	*0.122*

Supplementary Table 2. Competing risk regression models for any distant relapse as first event and site-specific distant relapse as first event: univariate models by HER2DX pCR-score and multivariate models by HER2DX pCR-score and intrinsic subtype.

		Univariate competing risk regression models HER2DX pCR-score*		
Endpoint	Variables	subHR	95% CI	p
Any distant	Medium vs Low	1.11	0.65-1.89	0.710
	High vs Low	0.53	0.29-0.98	0.042
Bone	Medium vs Low	0.55	0.25-1.21	0.140
	High vs Low	0.28	0.11-0.70	0.007
Bone only	Medium vs Low	0.15	0.03-0.69	0.015
	High vs Low	0.13	0.03-0.57	0.007
		Multivariate competing risk regression models HER2DX pCR-score and intrinsic subtype#		
Endpoint	Variables	subHR	95% CI	p
Any distant	pCR score Medium vs Low	0.82	0.48-1.41	0.480
	pCR score High vs Low	0.36	0.19-0.68	0.002
	LumA vs others	0.29	0.13-0.64	0.002
Bone only	pCR score Medium vs Low	0.15	0.03-0.70	0.016
	pCR score High vs Low	0.13	0.02-0.87	0.036
	LumA+LumB vs others	0.98	0.28-3.43	0.970

*only significant/evaluable models showed.

#models for those events for which both HER2DX pCR score and intrinsic subtype showed a significant association in univariate analysis.

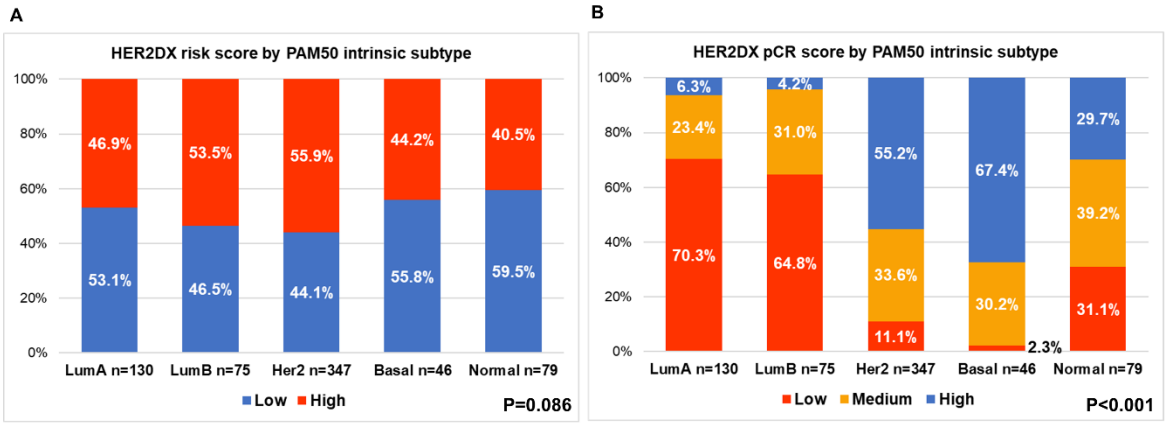
Abbreviations: subHR, sub Hazard Ratio; CI, confidence interval; pCR, pathologic complete response; LumA, Luminal A; LumB, Luminal B.

Supplementary Table 3. Frequency of site-specific metastases according to classic clinicopathological and molecular features among the 77 patients who experienced a distant relapse as a first event.

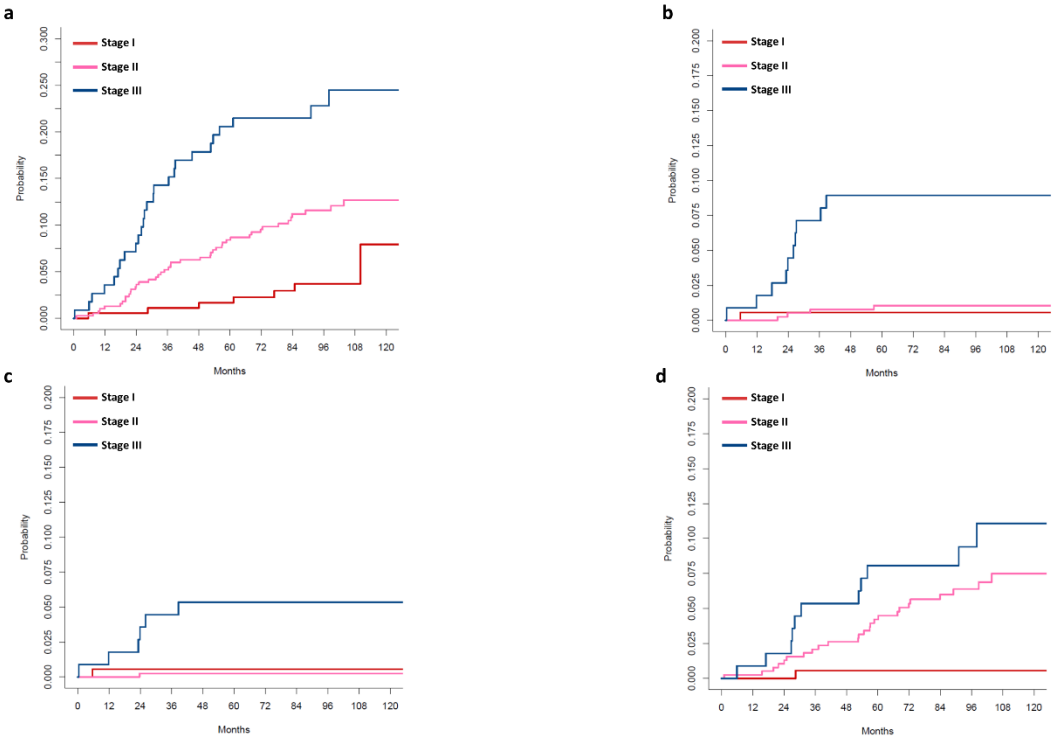
Abbreviations: Neg, negative; Pos, positive; n, number; p, p-value; LumA, Luminal A; LumB, Luminal B, HER2-e, HER2-enriched; Basal, Basal-like; Normal, Normal -like.

		Brain		Brain only		Lung		Bone		Bone only		Liver	
		%	p	%	p	%	p	%	p	%	p	%	p
Hormone receptors	**Neg, n=26**	23.1%		7.7%		26.9%		30.8%		7.7%		30.8%	
	Pos, n=51	15.7%	0.427	9.8%	0.761	25.5%	0.892	52.9%	0.065	27.5%	0.043	35.3%	0.691
Histologic Grade	**1-2, n=19**	26.3%		5.3%		26.3%		52.6%		26.3%		36.8%	
	3, n=48	18.8%	0.492	12.5%	0.383	25.0%	0.911	41.7%	0.416	20.8%	0.628	31.3%	0.660
Stage	**I, n=7**	14.3%		14.3%		14.3%		14.3%		0.0%		42.9%	
	II, n=44	9.1%		2.3%		31.8%		52.3%		22.7%		38.6%	
	III, n=26	34.6%	0.027	19.2%	0.051	19.2%	0.388	42.3%	0.159	23.1%	0.364	23.1%	0.358
Intrinsic subtype	**LumA, n=8**	0.0%		0.0%		25.0%		100.0%		62.5%		25.0%	
	LumB, n=10	0.0%		0.0%		20.0%		60.0%		30.0%		40.0%	
	Her2, n=45	26.7%		15.6%		24.4%		31.1%		11.1%		35.6%	
	Basal, n=7	0.0%		0.0%		71.4%		42.9%		0.0%		28.6%	
	Normal, n=7	28.6%	0.083	0.0%	0.242	0.0%	0.037	57.1%	0.006	42.9%	0.004	28.6%	0.954
HER2DX risk-score	**Low, n=15**	13.3%		6.7%		20.0%		33.3%		6.7%		40.0%	
	High, n=56	19.6%	0.575	8.9%	0.780	26.8%	0.592	48.2%	0.304	25.0%	0.122	33.9%	0.662
HER2DX pCR-score	**Low, n=25**	12.0%		4.0%		16.0%		64.0%		44.0%		32.0%	
	Medium, n=29	27.6%		13.8%		27.6%		34.5%		6.9%		41.4%	
	High, n=17	11.8%	0.244	5.9%	0.396	35.3%	0.346	35.3%	0.061	11.8%	0.002	29.4%	0.665

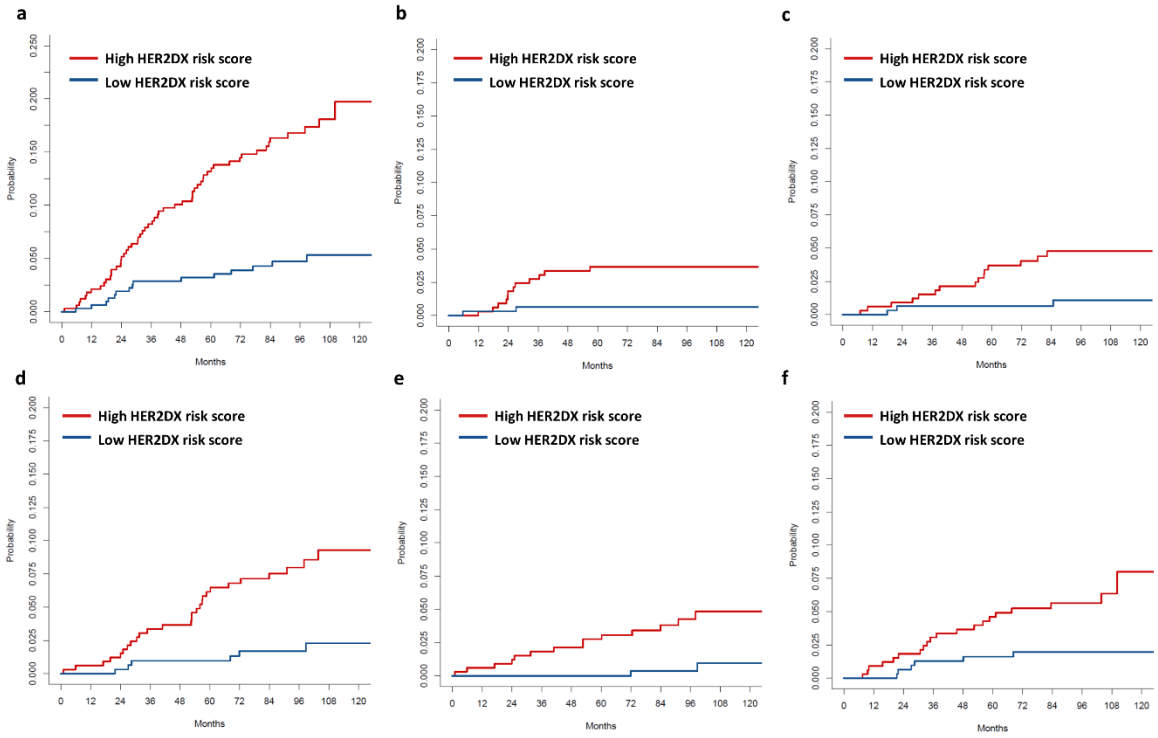
Supplementary Figure 1. Distribution of HER2DX risk-score (A) and pCR score (B) according to PAM50 intrinsic molecular subtypes.



Supplementary Figure 2. Cumulative incidence curves according to stage at diagnosis for the following distant relapse as first event: any distant (a), brain metastasis (b), brain-only metastasis (c), and bone metastasis (d).



Supplementary Figure 3. Cumulative incidence curves according to HER2DX risk-score for the following distant relapse as first events: any relapse (a), brain metastasis (b), lung metastasis (c), bone metastasis (d), bone-only metastasis €, and liver metastasis (f).



References

1. Conte P, Frassoldati A, Bisagni G, et al. Nine weeks versus 1 year adjuvant trastuzumab in combination with chemotherapy: final results of the phase III randomized Short-HER study‡. *Annals of oncology : official journal of the European Society for Medical Oncology*. 2018;29(12):2328-2333. doi:10.1093/annonc/mdy414

2. Guarneri V, Generali DG, Frassoldati A, et al. Double-Blind, Placebo-Controlled, Multicenter, Randomized, Phase IIB Neoadjuvant Study of Letrozole-Lapatinib in Postmenopausal Hormone Receptor–Positive, Human Epidermal Growth Factor Receptor 2–Negative, Operable Breast Cancer. *Journal of Clinical Oncology*. 2014;32(10):1050-1057. doi:10.1200/JCO.2013.51.4737

3. Dieci MV, Prat A, Tagliafico E, et al. Integrated evaluation of PAM50 subtypes and immune modulation of pCR in HER2-positive breast cancer patients treated with chemotherapy and HER2-targeted agents in the CherLOB trial. *Annals of oncology : official journal of the European Society for Medical Oncology*. 2016;27(10):1867-1873. doi:10.1093/annonc/mdw262

4. Prat A, Guarneri V, Pascual T, et al. Development and validation of the new HER2DX assay for predicting pathological response and survival outcome in early-stage HER2-positive breast cancer. *EBioMedicine*. 2022;75:103801. doi:10.1016/j.ebiom.2021.103801