

# A 14-gene B-cell immune signature in early-stage triple-negative breast cancer (TNBC): a pooled analysis of seven studies



Benedetta Conte,<sup>a,b,m</sup> Fara Brasó-Maristany,<sup>a,b,c,m</sup> Adela Rodríguez Hernández,<sup>a,b</sup> Tomás Pascual,<sup>a,b,d</sup> Guillermo Villacampa,<sup>c,d,e</sup> Francesco Schettini,<sup>a,b,f</sup> Maria J. Vidal Losada,<sup>a,b,d,f,g</sup> Elia Seguí,<sup>a,b,d</sup> Laura Angelats,<sup>a,b</sup> Isabel Garcia-Fructuoso,<sup>a,b</sup> Raquel Gómez-Bravo,<sup>a,b</sup> Natàlia Lorman-Carbó,<sup>a</sup> Laia Paré,<sup>c</sup> Mercedes Marín-Aguilera,<sup>c</sup> Olga Martínez-Sáez,<sup>a,b</sup> Barbara Adamo,<sup>a,b</sup> Esther Sanfeliu,<sup>a,h</sup> Beatrice Frattini,<sup>a</sup> Claudette Falato,<sup>a,d</sup> Núria Chic,<sup>a,b</sup> Ana Vivancos,<sup>c,i</sup> Patricia Villagrasa,<sup>c</sup> Johan Staaf,<sup>j</sup> Joel S. Parker,<sup>k</sup> Charles M. Perou,<sup>k,l</sup> and Aleix Prat<sup>a,b,c,f,g,\*</sup>



<sup>a</sup>Translational Genomics and Targeted Therapies in Solid Tumors Group, August Pi i Sunyer Biomedical Research Institute (IDIBAPS), Barcelona, Spain

<sup>b</sup>Cancer Institute and Blood Diseases, Hospital Clinic de Barcelona, Barcelona, Spain

<sup>c</sup>Reveal Genomics, Barcelona, Spain

<sup>d</sup>SOLTI Cooperative Group, Barcelona, Spain

<sup>e</sup>Oncology Data Science, Vall d'Hebron Institute of Oncology, Barcelona, Spain

<sup>f</sup>Department of Medicine, University of Barcelona, Barcelona, Spain

<sup>g</sup>Institute of Oncology (IOB)-Hospital QuirónSalud, Barcelona, Spain

<sup>h</sup>Pathology Department, Hospital Clinic of Barcelona, Barcelona, Spain

<sup>i</sup>Vall d'Hebron Institute of Oncology (VHIO), Cancer Genomics Group, Barcelona, Spain

<sup>j</sup>Division of Translational Cancer Research, Department of Laboratory Medicine, Lund University, Sweden

<sup>k</sup>Department of Genetics, University of North Carolina, Chapel Hill, NC, USA

<sup>l</sup>Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC, USA

## Summary

**Background** Early-stage triple-negative breast cancer (TNBC) displays clinical and biological diversity. From a biological standpoint, immune infiltration plays a crucial role in TNBC prognosis. Currently, there is a lack of genomic tools aiding in treatment decisions for TNBC. This study aims to assess the effectiveness of a B-cell/immunoglobulin signature (IGG) alone, or in combination with tumor burden, in predicting prognosis and treatment response in patients with TNBC.

**Methods** Genomic and clinical data were retrieved from 7 cohorts: SCAN-B (N = 874), BrighTNess (n = 482), CALGB-40603 (n = 389), METABRIC (n = 267), TCGA (n = 118), GSE58812 (n = 107), GSE21653 (n = 67). IGG and a risk score integrating IGG with tumor/nodal staging (IGG-Clin) were assessed for event-free survival (EFS) and overall survival (OS) in each cohort. Random effects model was used to derive pooled effect sizes. Association of IGG with pathological complete response (pCR) was assessed in CALGB-40603 and BrighTNess. Immune significance of IGG was estimated through CIBERSORTx and EcoTyper.

**Findings** IGG was associated with improved EFS (pooled HR = 0.77, [95% CI = 0.70–0.85], I<sup>2</sup> = 18%) and OS (pooled HR = 0.79, [0.73–0.85], I<sup>2</sup> = 0%) across cohorts, and was predictive of pCR in CALGB-40603 (OR 1.25, [1.10–1.50]) and BrighTNess (OR 1.57 [1.25–1.98]). IGG-Clin was predictive of recurrence (pooled HR = 2.11, [1.75–2.55], I<sup>2</sup> = 0%) and death (pooled HR = 1.99, 95% [0.84–4.73], I<sup>2</sup> = 79%) across cohorts. IGG was associated with adaptive immune response at CIBERSORTx and EcoTyper analysis.

**Interpretation** IGG is linked to improved prognosis and pCR in early-stage TNBC. The integration of IGG alongside tumor and nodal staging holds promise as an approach to identify patients benefitting from intensified or de-intensified treatments.

**Funding** This study received funding from: Associació Beca Marta Santamaria, European Union's Horizon 2020 research and innovation and Marie Skłodowska-Curie Actions programs, Fundación FERO, Fundación CRIS contra el cáncer,

eBioMedicine

2024;102: 105043

Published Online 5 March 2024

<https://doi.org/10.1016/j.ebiom.2024.105043>

1016/j.ebiom.2024.105043

105043

\*Corresponding author. Translational Genomic and Targeted Therapies in Solid Tumors, August Pi i Sunyer Biomedical Research Institute (IDIBAPS), Barcelona, Spain.

E-mail address: [alprat@clinic.cat](mailto:alprat@clinic.cat) (A. Prat).

<sup>m</sup>Both authors contributed equally.

Agència de Gestó d'Ajuts Universitaris i de Recerca, Instituto de Salud Carlos III, Fundación Contigo, Asociación Cáncer de Mama Metastásico IV, Breast Cancer Research Foundation, RESCUER, Fundación científica AECC and FSEOM.

Copyright © 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

**Keywords:** Triple-negative breast cancer (TNBC); Pathological complete response (pCR); Event-free survival (EFS); Overall survival (OS); B-cell/immunoglobulin signature (IGG); Prognostic biomarkers; Predictive biomarkers; Gene expression

### Research in context

#### Evidence before this study

Neoadjuvant chemotherapy with anthracyclines and taxanes has long been the standard treatment for most patients with early-stage TNBC. However, recent advancements have introduced escalated and de-escalated strategies to tailor treatment intensity based on prognosis. For instance, anthracyclines-free regimens combining carboplatin and a taxane have recently emerged as a less toxic treatment option in the neoadjuvant setting. On the other hand, escalated approaches, such as the incorporation of immune checkpoint inhibitors and platinum agents into anthracyclines-containing regimens, have shown to reduce the risk of recurrence in this context. Currently, decisions whether to endorse patients for such approaches are still based on tumor size and nodal status in the pre-treatment context, and on the evaluation of tumor response in the post-neoadjuvant setting. Nevertheless, there is a growing recognition of the substantial role played by tumor biology, particularly its immunological features, in determining patient outcomes. Consequently, there is a pressing need for tools that integrate the prognostic information from these features into a single assay to guide systemic therapy effectively.

#### Added value of this study

This study represents an attempt to construct a prognostic score for early TNBC by combining clinical and genomic

immune variables. We have identified a 14-gene immunoglobulin B-cell signature (IGG) as an independent predictor of pCR and survival outcomes in patients with early-stage TNBC. The IGG signature provided significant prognostic information beyond tumor size and nodal status, and it predicted distinct survival outcomes in patients with low tumor burden. Building upon this discovery, we evaluated the IGG-Clin risk-score, based on the HER2DX genomic assay for HER2+ breast cancer, which comprehensively assesses both the expression of the IGG signature and tumor stage. The IGG-Clin risk-score successfully identified patients at significantly different risks of relapse and death across cohorts.

#### Implications of all the available evidence

Integrating the prognostic information obtained from both IGG expression and tumor stage into a unified assay holds the potential to enhance treatment decision-making in early-stage TNBC. This approach could prove particularly beneficial for patients whose tumors fall into low-risk groups, as it may enable the use of less aggressive treatment strategies while still achieving curative outcomes. Further studies will be crucial to establish the clinical utility of integrating immunogenomic and clinical features into a prognostic assay to personalize treatment for patients with TNBC.

### Introduction

Early-stage triple-negative breast cancer (TNBC) has a poor prognosis despite (neo)adjuvant multi-agent chemotherapy.<sup>1,2</sup> Classical factors associated with risk of recurrence are tumor size and nodal status.<sup>3</sup> These two clinical-pathological factors dictate the possibility to combine multi-agent chemotherapy with pembrolizumab, an immune checkpoint inhibitor (ICI).<sup>4</sup> However, TNBC is clinically and biologically heterogeneous, and immune-related features can influence prognosis beyond type of treatment and tumor burden.<sup>5–16</sup> Therefore, a standardized assay integrating tumor burden with immune-related characteristics could improve prognostic stratification and treatment personalization.

In the last decade, the abundance of tumoral immune infiltrate has emerged as a prominent molecular biomarker in TNBC.<sup>7,10,17</sup> Tumor-infiltrating

lymphocytes (TILs) have been associated with pathological complete response (pCR) in patients receiving neoadjuvant chemotherapy,<sup>9,10</sup> and with improved survival in patients treated with and without chemotherapy.<sup>7,12</sup> Among the diverse TILs subpopulations, the presence of CD8+ T-lymphocytes, assessed either with immunohistochemistry or gene expression, has shown a strong prognostic impact,<sup>18–21</sup> while the role of B-lymphocytes has long been overlooked. Recently, compelling evidence has suggested that B-lymphocytes could indeed play a pivotal role in TNBC outcomes, particularly when they exhibit an antigen-driven phenotype.<sup>22,23</sup> For instance, the oligoclonal expansion of type G immunoglobulins (IgG) in the tumor stroma of patients with early-stage TNBC was strongly associated with long-term survival in the CALGB-40603 neoadjuvant trial.<sup>24</sup>

Similarly, a 14-gene immunoglobulin B-cell signature (IGG) has been identified as a key variable associated with a better prognosis and a higher probability of pCR in HER2-positive (HER2+) breast cancer treated with anti-HER2-based chemotherapy.<sup>25–28</sup> The IGG signature in combination with tumor burden (i.e., tumor size and nodal status), and two other gene signatures tracking proliferation and luminal features, have been integrated into a single assay, called HER2DX, and validated retrospectively across >7 studies and >1800 patients. This genomic test is currently available for clinical use.<sup>25,29,30</sup>

Here, we aimed to clinically validate the association of the IGG signature with prognosis and/or pCR across various cohorts of patients with TNBC. In addition, we aimed to assess the prognostic value of a multi-feature assay based on the HER2DX prognostic algorithm when IGG, tumor size and nodal status are considered.

## Methods

### Patient datasets

A summary of all the cohorts evaluated is available in [Table S1](#) and [Fig. 1](#). Genomic, clinical and survival data were extracted from 5 non-overlapping, publicly available datasets of patients with TNBC. Data from The Cancer Genome Atlas (TCGA) and METABRIC were retrieved from the cBio Cancer Genomics Portal (<http://cbioportal.org>).<sup>31,32</sup> Data from the last updated version of the SCAN-B dataset was retrieved from Mendeley Data (<https://data.mendeley.com/datasets/yzxtxn4nmd>).<sup>33,34</sup> Two other publicly available datasets were obtained from the NCBI Gene

Expression Omnibus (GEO) repository (<https://www.ncbi.nlm.nih.gov>) under the accession numbers GSE21653, and GSE58812.<sup>13,35</sup> Patients were eligible for inclusion if they exhibited early-stage breast tumors classified as triple-negative by immunohistochemistry and had available gene expression data along with either disease-related survival or overall survival data. In cases where the percentage of tumor nuclei staining positively for estrogen and progesterone receptors was accessible, tumors were deemed triple-negative if both receptor types were below 10%. For datasets that presented estrogen and progesterone receptor expression as binary variables (i.e., “positive” or “negative”), tumors were classified as triple-negative based on the specific definition provided within each dataset ([Table S1](#)). Patients with missing information of (neo)adjuvant treatments or who did not receive any were still regarded as eligible. Both females and male patients were eligible.

To further assess IGG and IGG-Clin in a more homogeneous and contemporary treatment setting and explore potential associations with pCR, two additional datasets of patients treated with neoadjuvant chemotherapy (NACT) were explored. The CALGB-40603 phase II trial randomized 443 patients with stage II-III TNBC to receive either standard NACT (12 cycles of weekly paclitaxel followed by 4 cycles of doxorubicin plus cyclophosphamide), standard NACT plus carboplatin, standard NACT plus bevacizumab, or standard NACT plus carboplatin and bevacizumab.<sup>24,36</sup> Clinical data and RNA-seq data of pre-treatment baseline samples from 389 (88%) patients from CALGB-40603 were downloaded from the dbGaP web site, under accession

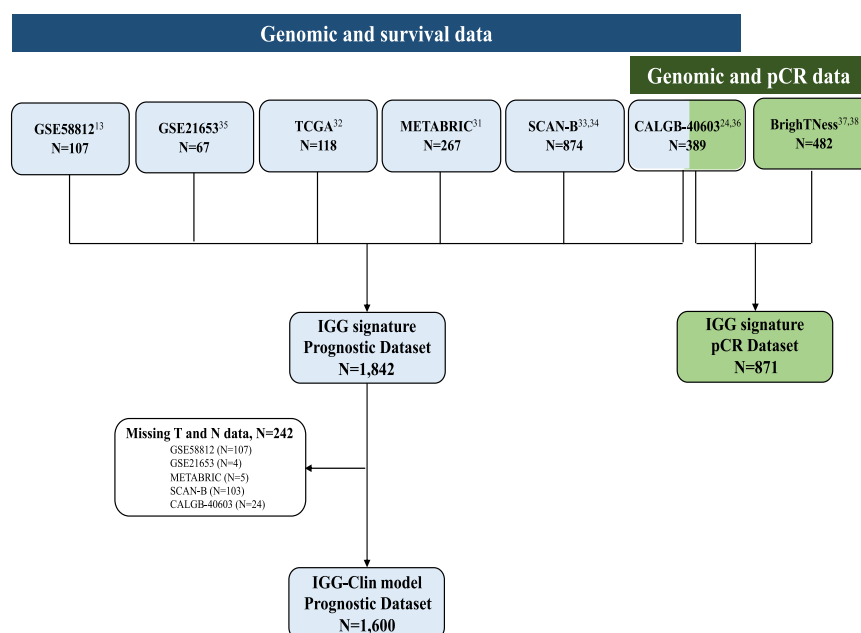


Fig. 1: Overview of the various patient cohorts assessed.

number phs001863.v1.p1. The BrighTNess phase III, placebo-controlled trial randomized 634 patients with stage II-III TNBC to receive either standard NACT (same scheme as CALGB-40603), standard NACT plus carboplatin, or standard NACT plus carboplatin plus veliparib.<sup>37,38</sup> Clinical data and RNA-seq data of pretreatment baseline samples from 482 (76%) patients from BrighTNess were downloaded from the GEO repository web site, under accession number GSE164458. Survival data along with matched genomic data were not publicly available for the BrighTNess dataset. Direct download links to gene expression datasets are provided in the [Supplementary methods](#).

### Gene expression

Gene expression was assessed with different genomic platforms across datasets ([Table S1](#)). Overall, there was moderate correlation between platforms, with the exception of dataset GSE58812 ([Supplementary methods](#)). Z-scores or log<sub>2</sub>-transformed normalized expression from either RNA-seq or Microarray data was used for single genes and to derive the IGG signature, depending on the type of data available for each dataset. Briefly, IGG has been previously identified by an unsupervised clustering of 550 node-negative breast tumors,<sup>39</sup> and externally validated across several clinical cohorts, for a total of >1000 patients.<sup>25–28,40</sup> The IGG signature is composed of 14 genes implicated in maturation of T and B lymphocytes progenitors (*IL2RG*), CD4+ and B lymphocytes activation and survival (*CD27*, *TNFRSF17*, *PIM2*), B lymphocytes differentiation in germinal centers (*POU2AF1*), immunoglobulin production (*CD79a*, *JCHAIN*, *IGKC*, *IGL*, *IGLV3-25*), chemotaxis (*CXCL8*, *NTN3*), and regulation of B, T and NK lymphocytes activity (*LAX1*, *HLA-C*). In datasets where the expression of some IGG genes was missing, the signature score was calculated from the available genes ([Figure S1](#)). For consistency with previous studies, the signature was calculated as mean expression of genes.

### IGG-Clin score

IGG-Clin is based on the prognostic HER2DX algorithm developed for early-stage HER2+ breast cancer.<sup>25–28,41,42</sup> Briefly, HER2DX risk-score is a clinical genomic test, which integrates clinical-pathological information (i.e., tumor size and nodal status) with the expression of 4 gene signatures, including IGG. The 2 clinical-pathological and the 4 genomic variables are multiplied by a risk coefficient to obtain a final HER2DX risk-score (from 0 to 100). In HER2DX, tumor size and nodal status are categorized as T1 vs T2-4 and N0 vs N1 vs N2-3, respectively. In addition, gene signatures are computed as continuous variables. IGG-Clin score is based on the same exact HER2DX risk-score algorithm without the luminal and proliferation signatures yet retains the risk coefficients for tumor size, nodal status and IGG score.

### Biological relevance of IGG signature

To evaluate the immune cell-type composition captured by the IGG signature, CIBERSORTx (<https://cibersortx.stanford.edu/>) and EcoTyper (<https://ecotyper.stanford.edu/carcinoma/>) publicly available tools were interrogated separately in each dataset.<sup>43,44</sup> Differences in immune cell-type composition between IGG groups (i.e., quartiles), were compared using a 2-way ANOVA considering both IGG quartiles and datasets, followed by Tukey's honest significant difference analysis.

To further investigate the relevance of IGG in TNBC, we interrogated a larger panel of 185 breast cancer-related genes encompassing the 14 IGG genes, the 50 genes from the PAM50 subtypes,<sup>45,46</sup> and 121 genes implicated in proliferation, luminal differentiation, HER2-signalling, and immune response. To assess whether the IGG genes would be selected for their prognostic and predictive relevance from this core panel, we examined the association of each gene with event-free survival (EFS) in the CALGB-40603 and SCAN-B datasets, and with pCR in the CALGB-40603 and BrighTNess datasets. Genes showing consistent association with both EFS and pCR across datasets were selected as core significant genes. Next, we aimed at demonstrating that the combined evaluation of core significant genes tracking a specific biological module, like the IGG signature, enhances the accuracy of EFS and pCR prediction compared to individual genes. To do so, we compared the goodness-of-fit of models combining just two core immune genes (corresponding to a minimal signature) to models using individual core immune genes. Finally, the goodness-of-fit of different methods of signatures calculation (i.e., addition, subtraction, multiplication and ratio) were also evaluated.

The univariate and multivariable Cox regression models and logistic regression models tested for the goodness-of-fit were adjusted for treatment arm. The goodness-of-fit of each model was measured by the chi-square ( $\chi^2$ ) statistic obtained from a likelihood ratio test.

### Statistical analysis

All survival analyses were based on time-to-event data, with the start time set as the date of cancer diagnosis. For surrogate survival endpoints, the end time was defined according to the occurrence of the dataset-specific event of interest or at the loss of follow-up. Because surrogate survival endpoint definitions varied across datasets ([Table S1](#)), we uniformly adopted the term EFS in our reporting to represent these various endpoints. For the analysis of overall survival (OS), the end time was determined by death from any cause or the loss of follow-up. The association of IGG and IGG-Clin with disease-related survival outcomes was assessed separately for each dataset ([Table S1](#)) using uni- and multi-variable Cox model analyses. Variables known to be associated with survival in early breast cancer, i.e., age at diagnosis, tumor size (T1 vs T2-4),

and nodal status (N0 vs N+), were included in the multivariable model.<sup>3,47</sup> IGG and IGG-Clin were evaluated as continuous variables and as quartiles. Hazard ratios (HRs) from uni- and multi-variable analyses were pooled using random effects model according to the Hartung-Knapp-Sidik-Jonkman method.<sup>48</sup> To test the contribution of the IGG score, we estimated the log likelihood ratio statistic of IGG quartiles as an addition to a Cox model containing tumor size and nodal status. The definition of disease-related survival outcomes in each dataset can be found in [Table S1](#). Of note, GSE58812 dataset was not included in the pooled analysis since no data on tumor size and nodal status was available.

The association of IGG signature with pCR following neoadjuvant chemotherapy (NACT) was assessed in the CALGB-40603<sup>24,36</sup> and BrighTNess<sup>37,38,49</sup> datasets using uni- and multi-variable logistic regression analyses to calculate the odds ratio (OR) and 95% Confidence Intervals (CIs). Variables known to be associated with pCR, i.e., tumor size (T1 vs T2-4), nodal status (N0 vs N+), and carboplatin receipt, were included in the multivariable model.<sup>50,51</sup> IGG was evaluated as a continuous variable and as quartiles. Interaction tests were used to evaluate whether the benefit of carboplatin in pCR was similar between IGG subgroups. To test the interaction, a new covariate with the interaction between carboplatin and IGG status was included in the logistic model. pCR was defined as absence of invasive disease in breast and axilla (ypT0/Tis ypN0).

For all statistical analyses, the significance level was set to a 2-sided alpha of 0.05.

### Role of the funding source

The funding sources had no role in the study design, data interpretation, manuscript preparation and final submission.

### Ethics

The study was performed in accordance with Good Clinical Practice guidelines and the Declaration of Helsinki. No informed consents were needed for the present study. Data were retrieved from public repositories.

## Results

### Description of prognostic datasets

Survival follow-up was available for 1842 patients from SCAN-B, CALGB-40603, METABRIC, TCGA, GSE58812, and GSE21653 ([Fig. 1](#)). Datasets were substantially heterogeneous in terms of genomic data, clinical data, and follow-up ([Table 1](#) and [Tables S1 and S2](#)). Microarray gene expression data was available for METABRIC, GSE58812 and GSE21653, while RNA-seq data was available for SCAN-B, CALGB-40603 and TCGA. Different surrogate survival endpoints were reported in each dataset: disease-

free survival (DFS) in TCGA and GSE21653, distant-DFS in GSE58812, relapse-free interval (RFI) in SCAN-B, event-free survival (EFS) in CALGB-40603, and breast cancer specific survival (BCSS) in METABRIC. OS was missing in GSE21653 and CALGB-40603 datasets. Thereafter, we used the term EFS and OS across all datasets.

With respect to patients' characteristics, 100% of them were females, 67% (n = 1228) received (neo) adjuvant chemotherapy, while all patients in GSE21653 (n = 67) and GSE58812 (n = 107), and 33% (n = 39) of patients in TCGA, had missing information regarding the use of systemic treatments. Of 1842 patients with survival follow-up, 1600 (86.9%) had IGG-Clin score available. In GSE58812 dataset (n = 107, 9.8% of the combined cohort), tumor size and nodal status were missing ([Table 1](#), and [Fig. 1](#)).

### Association of IGG with EFS and OS

IGG as continuous variable was significantly associated with better EFS in SCAN-B (HR = 0.72, 95% CI = 0.62–0.83,  $p < 0.0001$ ), CALGB-40603 (HR = 0.82, 95% CI = 0.68–0.99,  $p = 0.040$ ), METABRIC (HR = 0.81, 95% CI = 0.67–0.99,  $p = 0.040$ ), and GSE21653 (HR = 0.54, 95% CI = 0.35–0.82,  $p = 0.004$ ). No significant association of IGG with EFS was observed in TCGA (HR = 0.86, 95% CI = 0.52–1.45,  $p = 0.577$ ) and GSE58812 (HR = 1.00, 95% CI = 0.69–1.45,  $p = 0.778$ ). With respect to OS, a significant association was observed in SCAN-B (HR = 0.78, 95% CI = 0.71–0.85,  $p < 0.0001$ ) and METABRIC (HR = 0.81, 95% CI = 0.67–0.96,  $p = 0.041$ ), while no association was observed in TCGA (HR = 0.63, 95% CI = 0.32–1.26,  $p = 0.319$ ) and GSE58812 (HR = 0.99, 95% CI = 0.68–1.46,  $p = 0.723$ ).

When results were pooled with a random effects model, IGG was found significantly associated with better EFS (HR = 0.77, 95% CI = 0.70–0.85,  $p = 0.003$ , heterogeneity statistic  $I^2 = 18\%$ ) and better OS (HR = 0.79, 95% CI = 0.73–0.85,  $p = 0.002$ ,  $I^2 = 0\%$ ). The association of IGG with survival outcome was found to be independent of age (continuous variable), tumor size (i.e., T1 vs T2-4), nodal status (i.e., N0 vs N+), and receipt of (neo)adjuvant chemotherapy (i.e., no chemotherapy vs chemotherapy vs unknown) both for EFS (HR = 0.77, 95% CI = 0.70–0.84,  $p < 0.0001$ , heterogeneity statistic  $I^2 = 11\%$ ) and OS (HR = 0.80, 95% CI = 0.73–0.88,  $p < 0.0001$ , heterogeneity statistic  $I^2 = 0\%$ ) ([Figures S2 and S3](#)).

### IGG in node-negative and stage 1 disease

To identify the optimal patient population with TNBC for therapy de-escalation, those with node-negative or stage 1 disease were considered. To concentrate on these specific groups, each dataset was divided into four equal parts (quartiles) based on IGG expression levels. Subsequently, all datasets were merged to analyse

	TCGA	METABRIC	GSE21653	GSE58812	SCAN-B	CALGB 40,603	Total	Eligible for IGG-Clin
<b>N</b>	118	267	87	107	874	389	1842	1600
<b>Tumour size, N (%)</b>								
T1	30 (25.4)	104 (39.0)	16 (18.4)	-	454 (52.0)	42 (10.8)	646 (35.1)	634 (39.6)
T2-4	88 (74.6)	160 (60.0)	68 (78.2)	-	341 (39.0)	340 (87.4)	997 (54.1)	966 (60.4)
NA	0 (0)	3 (1.0)	3 (3.4)	107 (100)	79 (9.0)	7 (1.8)	199 (10.8)	0 (0)
<b>Nodal status, N (%)</b>								
NO	75 (63.6)	128 (47.9)	51 (58.6)	-	517 (59.2)	164 (42.1)	935 (50.8)	692 (43.3)
N+	43 (36.4)	139 (52.1)	34 (39.1)	-	289 (33.0)	201 (51.7)	706 (38.3)	908 (56.7)
NA	0 (0)	0 (0)	2 (2.3)	107 (100)	68 (7.8)	24 (6.2)	201 (10.9)	0 (0)
<b>Treatment, N (%)</b>								
None	15 (12.7)	113 (42.3)	-	-	207 (23.7)	0 (0)	335 (18.2)	323 (20.2)
(Neo)adjuvant CT	64 (54.2)	154 (57.7)	-	-	621 (71.0)	389 (100)	1228 (66.8)	1139 (71.2)
NA	39 (33.1)	0 (0)	87 (100)	107 (100)	46 (5.3)	0 (0)	279 (15.1)	138 (8.6)
<b>Median follow-up, years (IQR)</b>	2.8 (1.4-5.0)	14.4 (8.0-17)	4.8 (3.3-8.5)	7.0 (4.3-9.2)	5.0 (4.5-8.0)	5.7 (4.8-6.4)	6	6
<b>Disease specific events</b>								
Deaths	19	109	27	31	131	115	432	373
Deaths	16	142	NA	29	244	NA	431	376

NA, not available; CT, chemotherapy.

**Table 1: Characteristics of patients in the prognostic datasets.**

survival associations across these quartiles. After a median follow-up period of 6.2 years, the quartiles of IGG expression demonstrated significant prognostic value in node-negative disease (n = 935) and stage 1 disease (n = 432), both in terms of EFS and OS (Fig. 2A–D). In node-negative disease, patients in the highest quartile showed a 69% and 63% relative reduction in the risk of event or death compared to those in the lowest quartile (EFS HR = 0.31, 95% CI = 0.19–0.52], p < 0.0001; OS HR = 0.37, 95% CI = 0.24–0.58, p = 0.001), which corresponded to a 5-year EFS and OS events rate of 10% and 8% in the highest quartile group, and of 32–28% in the lowest quartile group, respectively. In stage 1 disease, patients in the top two quartiles (i.e., IGG expression above median) showed a 54% and 39% relative reduction in the risk of event or death compared to those in the two lowest quartiles (i.e., below median) (EFS HR = 0.46, 95% CI = 0.25–0.84, p = 0.012; OS HR = 0.61, 95% CI = 0.38–0.97, p = 0.040). Such difference translated in to a 5-year EFS and OS event rate of 8% and 6% in the top two quartiles group and of 19% and 12% in the two lowest quartiles group, respectively (Fig. 2C and D).

**Prognostic value of IGG beyond tumor stage**

To evaluate the potential of a risk-score integrating IGG with tumor stage, we used the combined dataset (n = 1600) to build a multivariable Cox model containing (in sequential order) tumor stage, nodal status and IGG quartiles. Next, we estimated the relative contribution of each variable to the final log likelihood ratio statistic of the model, as a measure of the amount of prognostic information conveyed by each one these variables. The analysis demonstrated a substantial

prognostic contribution of IGG expression beyond tumor stage, amounting to 37% for EFS and 21% for OS (Fig. 3).

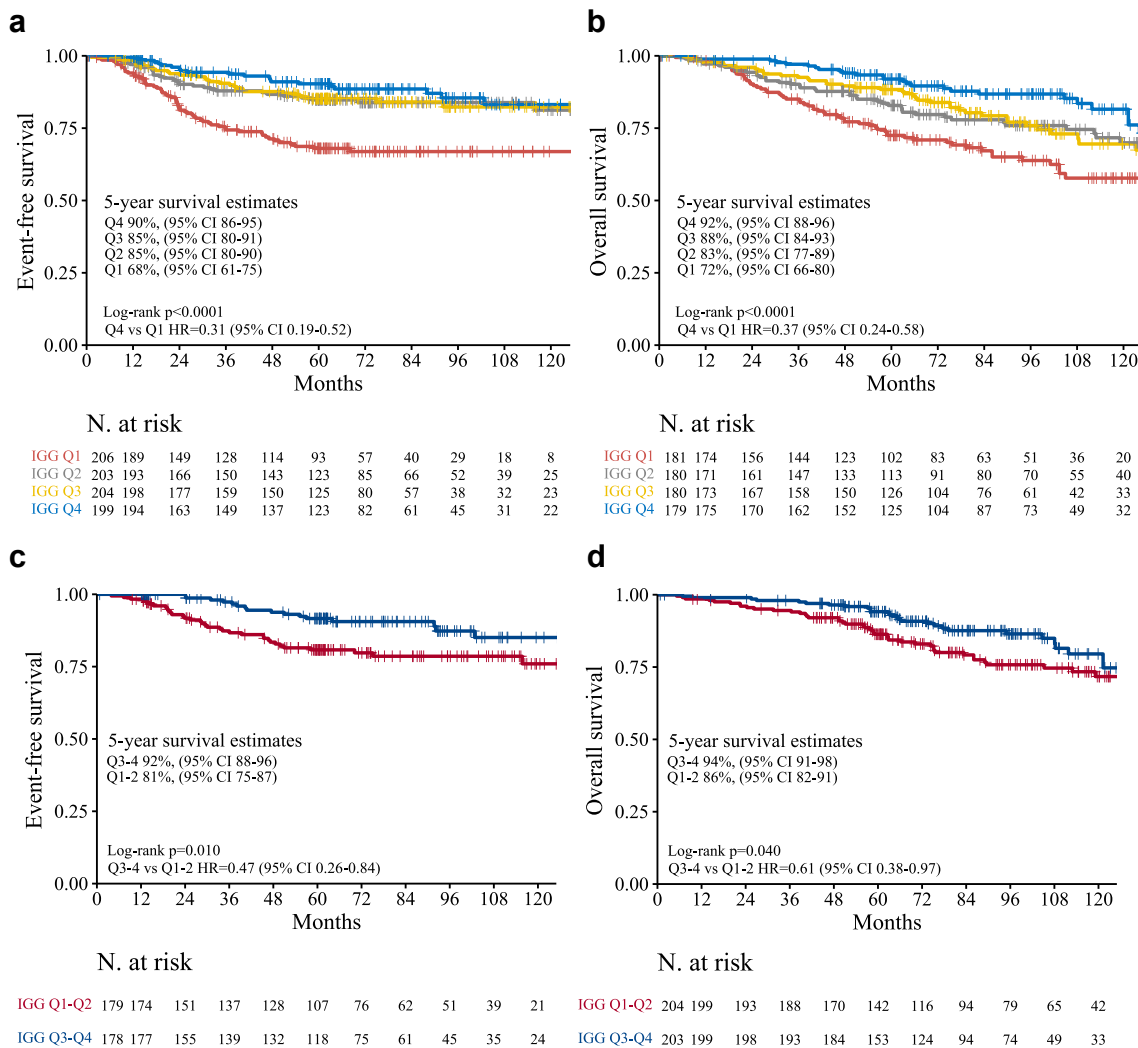
**Association of IGG-Clin with EFS and OS**

We assessed the IGG-Clin score, a comprehensive risk-score that combines IGG, tumor size, and nodal status into a unified measure. To create IGG-Clin, we used the risk coefficients of the previously reported HER2DX prognostic algorithm,<sup>25,41</sup> which integrates tumor size, nodal status and IGG to obtain a continuous risk score which is positively associated with the risk of recurrence. Of note, we omitted the luminal and proliferation signatures which are also part of the HER2DX risk-score.

IGG-Clin risk score as a continuous variable was found significantly associated with both EFS and OS across all datasets (Table S3). Upon pooling the results from each dataset using a random effects model a significant effect size was observed for EFS (HR = 2.11, 95% CIs = 1.75–2.55, p = 0.0004, I<sup>2</sup> = 0%), while the effect size for OS was not significant due to high heterogeneity (HR = 1.99, 95% CIs = 0.84–4.73, p = 0.076, I<sup>2</sup> = 79%) (Figure S4A-B).

To further investigate the prognostic potential of IGG-Clin risk score, patients from each dataset were categorized into quartiles based on their IGG-Clin risk scores. The resulting data were combined, resulting in a pooled dataset of 1600 patients. Most patients (n = 966; 60%) had primary tumors larger than 2 cm, and 908 (57%) were found to be node-positive, as outlined in Table 1.

Analysing the IGG-Clin quartiles revealed significant prognostic implications for both EFS and OS.



**Fig. 2: Kaplan-Meier estimates in node-negative or stage 1 TNBC according to the expression of the IGG signature (quartiles).** (a) EFS in patients with node-negative disease; (b) OS in patients with node-negative disease; (c) EFS in patients with stage I disease; (d) OS in patients with stage I disease.

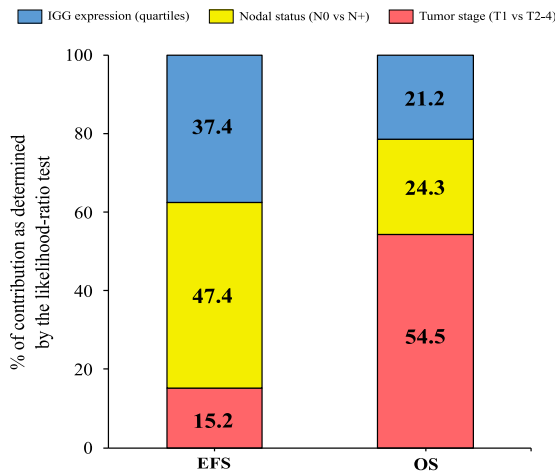
Compared to patients in the highest quartile, those in the lowest quartile experienced a notable 77% relative reduction in the risk of relapse or death (HR = 0.23, 95% CI = 0.16–0.32,  $p < 0.001$ ), leading to a substantial 31% absolute improvement in EFS at the 5-year mark (Fig. 4A). Similarly, the analysis for OS showed promising outcomes, with patients in the lowest IGG-Clin quartile exhibiting a 52% reduction in the risk of death (HR = 0.52, 95% CI = 0.40–0.67,  $p < 0.001$ ), translating to a remarkable 37% absolute increase in 5-year OS (Fig. 4B).

#### Association of IGG with pCR

To assess the connection between IGG and pCR after neoadjuvant multi-agent chemotherapy, we examined the CALGB-40603 and BrighTNess datasets (see

Table S4). The pCR rates were similar in both datasets: 48% (186/389) in CALGB-40603 and 49% (236/482) in BrighTNess.

Our analysis revealed a significant association of IGG as a continuous variable with pCR in both the CALGB-40603 cohort (OR 1.25, [95% CIs 1.10–1.50],  $p = 0.006$ ) and the BrighTNess cohort (OR 1.57 [1.25–1.98],  $p < 0.001$ ). This association remained significant across different subgroups, including tumor size, nodal status, and carboplatin administration (Table 2). Furthermore, patients in the highest quartile of IGG expression exhibited significantly higher pCR rates compared to those in the lowest quartile in both CALGB-40603 (61% vs. 38%, OR = 2.56, 95% CIs = 1.45–4.60,  $p = 0.001$ ) and BrighTNess (58% vs. 34%, OR = 2.64, 95% CIs = 1.57–4.48,  $p < 0.001$ ) cohorts (Fig. 5A and B and Table S5).



**Fig. 3: Impact of the IGG signature beyond tumor stage.** Visual depiction showcasing the distinct contributions of variables to improve the overall fit of the prognostic model for both event-free survival (EFS) and overall survival (OS) outcomes. The sequence of variable inclusion in each model followed the order of tumor stage, nodal stage, and IGG signature. The assessment employed the likelihood-ratio chi-squared statistic as a measure of how well the model fits the data.

pCR improvement with carboplatin was noted across all quartiles of IGG (Fig. 5C and D), however, no significant interaction was observed between IGG as a continuous variable and carboplatin administration, both in CALGB-40603 (p-value for interaction = 0.800) and in BrightTness (p-value for interaction = 0.200).

**Association of IGG with survival beyond pCR status and residual cancer burden**

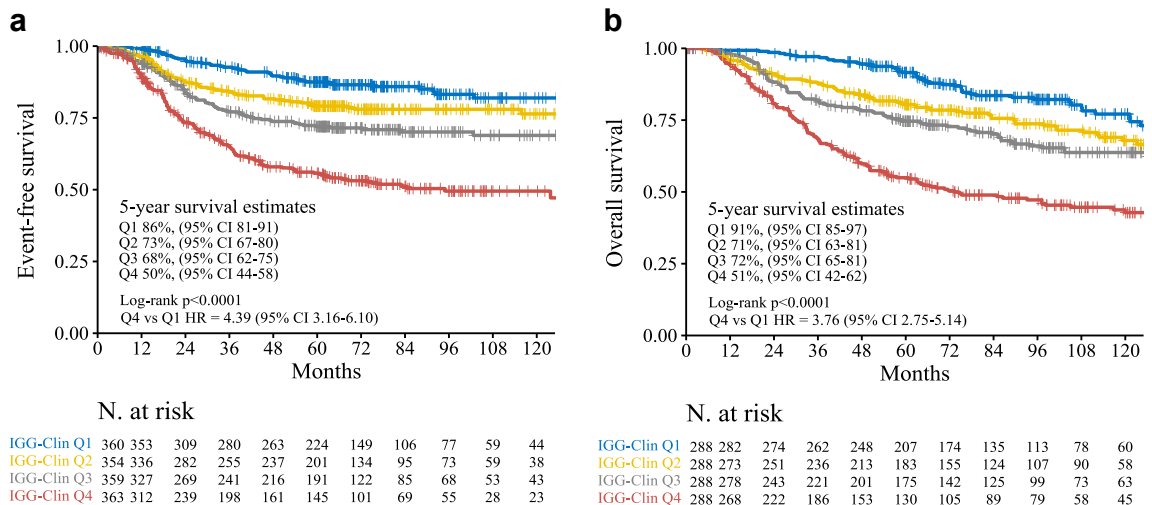
We investigated the relationship between IGG expression and survival outcomes in patients who underwent

neoadjuvant multi-agent chemotherapy and had known pCR status. This evaluation was conducted using the CALGB-40603 dataset.<sup>24</sup> The main results from CALGB-40603 trial (n = 443) did not show any significant association of tumor size and carboplatin with EFS, while pCR and nodal status were independent predictor of survival.<sup>24</sup> These results were confirmed in our genomic dataset (n = 389). In the univariate analysis, nodal status (HR = 1.96, 95% CI = 1.29–2.97, p = 0.002), pCR (HR = 0.30, 95% CI: 0.20–0.46, p < 0.0001), and residual cancer burden (HR 1.37, 95% CI: 1.25–1.49, p < 0.0001) were significantly associated with EFS, while tumor size (HR = 0.81, 95% CI: 0.46–1.43, p = 0.469) and the use of carboplatin (HR = 0.78, 95% CI: 0.32–3.11, p = 0.498) were not. The IGG signature was also significantly associated with EFS (HR = 0.86, 95% CI = 0.75–0.99, p = 0.040).

Upon incorporating IGG, nodal status, and pCR status into a multi-variable model, all three variables continued to exhibit significant associations with EFS, including IGG (HR = 0.84, 95% CI: 0.72–0.98, p = 0.026). Likewise, IGG remained an independent predictor of improved EFS in a multi-variable model including nodal status and residual cancer burden (HR 0.83, 95% CI = 0.71–0.97, p = 0.017).

**Association of IGG-Clin with survival beyond pCR status and residual cancer burden**

In the CALGB-40603 dataset, IGG-Clin, as a continuous variable, exhibited a significant association with EFS (HR = 3.22, 95% CI = 1.78–5.84, p < 0.001). When conducting bivariate analyses, considering pCR status and residual cancer burden (RCB), IGG-Clin remained significantly associated with survival (HR adjusted for pCR = 2.00, 95% CI = 1.47–2.66, p < 0.001; HR adjusted for RCB = 1.96, 95%



**Fig. 4: Kaplan-Meier estimates in TNBC according to the expression of the IGG-Clin risk score (quartiles).** (a) EFS in the pooled dataset; (b) OS in the pooled dataset.

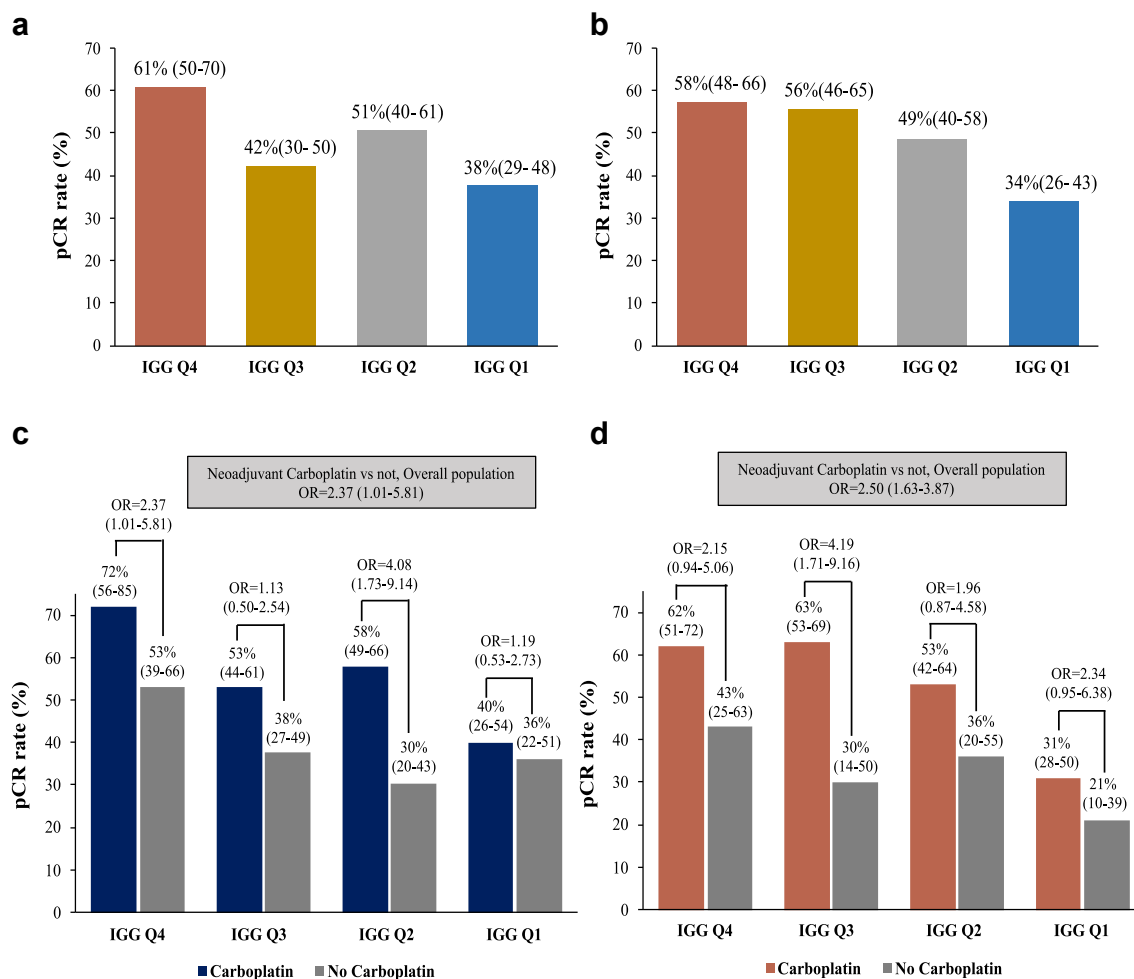
Dataset	CALGB-40603		BrightNess	
	HR (95% CI)	p value	HR (95% CI)	p value
IGG/immune	1.21 (1.10-1.53)	0.023	1.68 (1.32-2.15)	<0.0001
Tumour stage (T1 vs T2-4)	0.82 (0.42-1.50)	0.490	NA	NA
Nodal stage (N0 vs N+)	1.23 (0.80-1.91)	0.401	0.50 (0.32-0.74)	0.0003
Carboplatin (yes vs no)	1.90 (1.20-2.98)	0.003	2.70 (1.71-4.33)	<0.0001
Interaction IGG/immune*carboplatin	-	0.800	-	0.20

NA, not available; HR, hazard ratio.

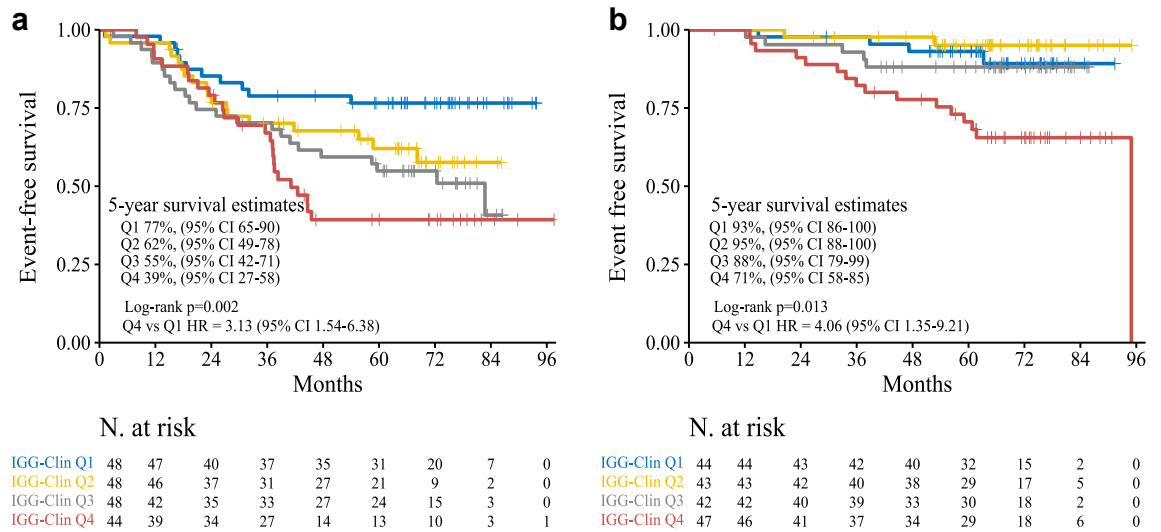
**Table 2: Multivariable logistic regression analyses for pCR in CALGB-40603 and BrightNess datasets.**

CI = 1.40–2.74,  $p < 0.001$ ). Within patients with residual disease and pCR, the IGG-Clin score effectively stratified individuals into different risk groups (Fig. 6A and B). In patients with a pCR, the IGG-Clin

higher risk group (i.e., top-quartile) displayed a 5-year EFS estimate of 71% (95% CI = 58–85), while the lower risk group (bottom-quartile) displayed a 5-year EFS of 93% (95% CI = 86–100) (Fig. 6B).



**Fig. 5: Association of the IGG signature with pCR in CALGB-40603 and BrightNess datasets.** (a, b) pCR rates according to the expression of the IGG signature according to quartiles in CALGB-40603 (a) and BrightNess (b); (c, d) effect of carboplatin on pCR rates according to the expression of the IGG signature in (c) CALGB-40603 and (d) BrightNess.



**Fig. 6:** Kaplan-Meier estimates in TNBC according to the IGG-Clin score and the type of pathological response in the CALGB-40603 trial. EFS according to IGG-Clin score (quartiles) in patients with (a) residual disease or (b) pCR following neoadjuvant chemotherapy.

**Relationship between IGG expression and immune cell subpopulations**

To explore the immunological features reflected by the IGG signature, each dataset was analysed through CIBERSORTx and Ecotyper deconvolution algorithms, in order to estimate the proportion or abundance of different immune cell types (CIBERSORTx) and immune activation states (EcoTyper) in the tumor micro-environment.<sup>43,44</sup> The results showed that tumors in the top IGG quartile were significantly enriched for plasma cells (p = 0.004), CD8+ lymphocytes (p = 0.002), follicular helper T cells (0.040), and M1 macrophages (0.035), and significantly depleted of M0 and M2 macrophages (0.004 and 0.003, respectively) (Fig. 7, Figure S5, and Table S6). Accordingly, high IGG expression was significantly associated with immune ecotypes CE9 and CE10, which correspond to tumors with the highest levels of immunoreactivity of the tumor microenvironment and potential benefit from immunotherapy (2-way ANOVA p = 0.006 and p = 0.002, respectively) (Figure S6 and Table S7).

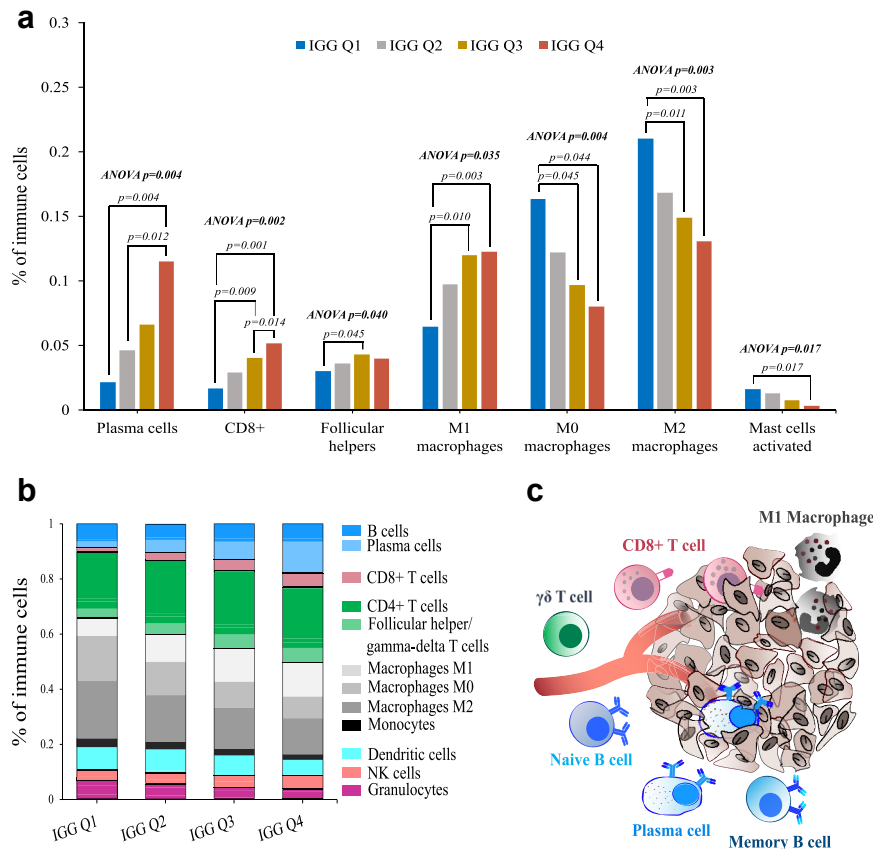
**Further relevance of the IGG signature**

To further assess the relevance of the IGG signature in TNBC within the SCAN-B, CALGB-40603, and BrightNess trials, we expanded our analysis and evaluated a panel of 185 breast cancer-related genes. This panel included the 14 IGG genes, the 50 genes of the PAM50 subtypes, and additional 121 genes linked to proliferation, luminal differentiation, HER2 signalling, and different immune processes (Table S8). From these 185 genes, only 10 showed consistent association with EFS in SCAN-B and CALGB-40603, and pCR in CALGB-40603 and BrightNess (Figure S7). Notably, all

10 genes (thereafter called core immune genes or CIGs) are involved in immune response, with 5 (i.e., *CD79A*, *LAX1*, *PIM2*, *POU2AF1*, and *TNFRSF17*) being part of the IGG signature and tracking B-cell immunity. The observed overlap between the IGG signature and the identified CIGs is statistically significant and not expected by chance (p < 0.0001, exact hypergeometric probability test). As a proof of concept, we aimed to illustrate the improved prognostication achieved by amalgamating relevant immune genes into a unified signature, in contrast to the assessment of individual genes. We investigated whether a minimal signature composed of two CIGs could surpass the predictive and prognostic accuracy of individual CIGs. These immune signatures were constructed by applying addition, subtraction, multiplication, and division to the gene expression values. The prognostic and predictive information conveyed by each signature and gene was quantified using the chi-square ( $\chi^2$ ) statistic obtained from a likelihood ratio test. Signatures generated by addition, mirroring the method used for the calculation of the IGG signature, exhibited superior prognostic and predictive information compared to single genes (Figure S8) and signatures derived from subtraction, multiplication, and division (Table S9).

**Discussion**

In this study, we aimed to construct a prognostic score for TNBC utilising a combination of clinical and immune-genomic variables. We have successfully established that the IGG signature serves as an independent predictor of pCR and survival outcomes in early-stage TNBC. Building upon this discovery, we



**Fig. 7: Immune infiltrate composition in TNBC based on IGG expression in the SCAN-B, BrightNess and CALGB-40603 datasets.** (a) Significant changes in immune cell types as determined by CIBERSORTx across IGG quartiles. Statistical differences were evaluated using a 2-way ANOVA considering both IGG quartiles and datasets, followed by Tukey's honest significant difference analysis; (b) Summary of the distribution of immune cell types as determined by CIBERSORTx across IGG quartiles. Percentages were calculated by averaging the expression of each cell type in each IGG quartile, with weighting based on the number of patients in each dataset; (c) Graphical representation of immune cell types associated with IGG in TNBC.

developed the IGG-Clin risk-score, which is a comprehensive assessment incorporating both the expression of the IGG signature and tumor stage. Notably, our study revealed that the IGG-Clin risk-score effectively stratifies patients into distinct risk groups with substantial variations in their likelihood of relapse.

After decades in which standard anthracycline/taxane-based chemotherapy has been the only option for TNBC, new approaches to either de-escalate or escalate treatment have now become available. These involve the exclusion of anthracyclines from platinum-taxane neoadjuvant regimens,<sup>52-56</sup> the addition of immune checkpoint inhibitors (ICIs) and platinum agents to standard neoadjuvant chemotherapy,<sup>57</sup> adjuvant capecitabine to treat residual disease in the post-neoadjuvant setting,<sup>58</sup> and adjuvant olaparib in high-risk patients with germline *BRCA1/2* mutations.<sup>59</sup> Currently, there is a lack of standardized prognostic tools that enable personalized treatment selection,

leading to the reliance on tumor stage assessment in the pre-treatment context and evaluation of tumor response in the post-neoadjuvant setting for patient selection. However, it is crucial to recognize that tumor biology plays a substantial role in determining patients' outcomes. In stage I disease, historical series report recurrence rates between 5 and 30%, underlining prognostic heterogeneity in patients with low clinical risk, and the need of novel biomarkers to identify those who may benefit from escalated approaches.<sup>60,61</sup>

According to our study, IGG expression discriminated between tumors with excellent (5-years relapse risk <10%) and dismal prognosis (5-years relapse risk ~20%) within stage I TNBC, and between those at high (>30%), intermediate (15%) and lower (~10%) risk of relapse within the T1-2 N0 group. Therefore, this signature could represent a valuable biomarker to stratify patients with low tumor burden. In stage II-III disease, recurrence rates are estimated to

be  $\geq 30\text{--}40\%$ , a sufficient risk to recommend treatment escalation with neoadjuvant carboplatin and ICIs.<sup>4,50</sup> However, evidence indicates that some patients could be cured with less intense treatments. In the control arms of neoadjuvant immunotherapy trials, 77–81% of participants remained free of recurrence at 3 years with chemotherapy alone.<sup>57,62–64</sup> Similarly, neoadjuvant trials assessing the addition of platinum agents show that up to 70% of patients with stage II–III TNBC remain free of recurrence with standard anthracycline-based chemotherapy.<sup>49,51</sup> Finally, phase II trials show that 12–18 cycles of taxane-carboplatin are sufficient to achieve pCR in 49–55% of stage I–III TNBCs, suggesting that de-escalation might be an option for some patients.<sup>52–56</sup>

In this scenario, combining the prognostic information conveyed by IGG expression and tumor burden could improve patient selection. The IGG-Clin algorithm re-classified patients in 4 groups with large differences in their risk of relapse and death. With a 50% probability of recurrence at 5 years, the highest risk group would be the most likely to benefit from combined escalated approaches, such as the addition of both carboplatin and ICIs to NACT. Patients in lower risk groups might benefit instead from intermediate approaches, whose toxicities do not outbalance potential benefits. A biomarker analysis of the NeoPACT trial showed that patients in the IGG-Clin low-risk group have excellent outcomes with the anthracycline-free regimen of carboplatin/docetaxel plus pembrolizumab (3-year EFS of 93%), while for IGG-Clin high-risk patients the prognosis is less favorable with such de-escalated approach (3-EFS of 80%).<sup>65</sup> Of note, the same analysis also showed that tumors with high IGG expression (above median) achieve pCR rates of 70% with the NeoPACT regimen, while in low IGG tumors, the rate dropped to 43%.<sup>65</sup>

In the post-neoadjuvant setting, patients with residual disease are currently endorsed for adjuvant systemic treatments, while no additional therapy is recommended for patients with pCR.<sup>58,59,66</sup> In this framework, novel biomarkers like RCB are being proposed to improve upon such dichotomous distinction.<sup>67,68</sup> In CALGB-40603, IGG expression and IGG-Clin risk score were associated with survival outcomes independently of pCR status and RCB, suggesting that IGG-based biomarkers might add prognostic information beyond the evaluation of pathological response. IGG-Clin identified patients with pCR at substantial risk of relapse ( $\sim 30\%$  at 5 years), who might benefit from adjuvant treatments despite the achievement of a complete response upon NACT.

From a biological standpoint, our findings highlight the significant role of B-cell immunity in the prognosis of TNBC. While B-lymphocytes represent a minor component of TILs, their presence could indicate the activation of an antigen-driven immune response against tumor cells.<sup>69</sup> In a study conducted by Shepherd

and colleagues in the CALGB-40603 trial,<sup>24</sup> oligoclonal expansion of IgG-biased B-lymphocytes within the tumor infiltrate was highly prognostic. Similar results have been observed in another retrospective series.<sup>22</sup>

By employing immune deconvolution analysis with CIBERSORTx and EcoTyper, we made a significant observation regarding TNBCs with heightened IGG expression. We found an enrichment of various cell types involved in the initiation, regulation, and maintenance of adaptive immune responses. Among these cell types were plasma cells, B memory lymphocytes, gamma-delta T lymphocytes, CD8+ lymphocytes, and M1 macrophages. Remarkably, similar results have been reported recently in the NeoPACT trial,<sup>65</sup> reinforcing the validity of our findings.

Our study has a few noteworthy limitations. Overall, datasets exhibited significant heterogeneity in terms of clinical data and patient characteristics. A limitation in terms of internal validity is the use of different surrogate survival endpoints across datasets. While this variability could have potentially impacted the association of IGG and IGG-Clin with the risk of recurrence, it is worth noting that the OS analysis consistently demonstrated a strong association of these biomarkers with patients' survival, reaffirming their prognostic significance. Methodological limitations such as the small number of studies included in the random-effects models, unmeasured confounding and a potential selection bias in HRs estimation should also be acknowledged.<sup>70</sup> Additionally, genomic profiling was conducted using different platforms across the datasets, making it challenging to create a unified dataset with continuous gene expression information. Consequently, several factors could have contributed to the lack of a significant prognostic effect of the IGG signature in the TCGA and GSE58812 datasets, including the limited sample size, shorter follow-up periods, and variations in gene expression analysis methodologies. From a clinical standpoint, the main limitation is the unavailability of TILs data in patients with survival or pCR information, which precluded to assess whether the addition of IGG would enhance the prognostication of TNBC in conjunction with TILs. Patients with node-negative TNBCs high in TILs ( $\geq 75\%$ ) have shown excellent survival outcomes without (neo)adjuvant systemic therapy, making TILs a valuable biomarker for potential treatment de-escalation.<sup>12</sup> Whether immune genomic signatures could outperform TILs in selecting patients for treatment de-escalation and escalation is currently a matter of research. Recently, a correlative analysis of the CALGB-40601 and the PAMELA trial showed that in HER2-positive early breast cancer the IGG signature outperformed TILs for both pCR and prognosis.<sup>40</sup> Similar analyses are warranted in trials including patients with TNBC to compare the prognostic ability of these different means of measuring immune activation. Finally, in the CALGB-40603 and BrightNess trials, we

did not observe any association of IGG with carboplatin response. However, it is important to note that these trials involved the administration of additional investigational treatments (bevacizumab and veliparib, respectively), which could have acted as confounding factors impacting the results.

To conclude, this study presents groundbreaking evidence indicating the potential for enhanced prognostication in TNBC through the utilisation of multi-feature classifiers that combine immune biomarkers and clinical data. Furthermore, our findings contribute to the expanding body of research emphasizing the crucial involvement of B-cell immunity in the context of TNBC.

#### Contributors

Alex Prat, Benedetta Conte, and Fara Brasó-Maristany contributed to the conceptualization of the study. Benedetta Conte wrote the original draft of the manuscript, with supervision provided by Alex Prat and Fara Brasó-Maristany. Data curation was handled by Alex Prat, Benedetta Conte, Fara Brasó-Maristany, Laia Paré, Tomás Pascual, Guillermo Villacampa, Adela Rodríguez Hernández, Francesco Schettini, Maria J. Vidal Losada, Elia Seguí, Laura Angelats, Isabel García-Fructuoso, Raquel Gomez-Bravo, Natàlia Lorman-Carbó, Mercedes Marín-Aguilera, Olga Martínez-Sáez, Barbara Adamo, Esther Sanfeliu, Beatrice Fratini, Claudette Falato, Núria Chic, Ana Vivancos, Patricia Villagrasa González, Johan Staaf, Joel S. Parker, and Charles M. Perou. The methodology was designed by Alex Prat, Benedetta Conte, and Fara Brasó-Maristany, with the formal analysis being performed by them. Guillermo Villacampa reviewed the formal analysis. All authors contributed to the review and editing of the manuscript. Alex Prat, Fara Brasó-Maristany, and Laia Paré verified the underlying data (validation). All authors have read and agreed to the published version of the manuscript.

#### Data sharing statement

All genomic and clinical data used in this paper can be retrieved from public repositories. The Cancer Genome Atlas (TCGA) and METABRIC data were retrieved from the cBio Cancer Genomics Portal (<http://cbioportal.org>). Data from the last updated version of the SCAN-B dataset was retrieved from Mendeley Data (<https://data.mendeley.com/datasets/yzxtxn4nmd>). Data from GSE21653, GSE58812 and BrighTNess (accession number GSE164458) were retrieved from the NCBI Gene Expression Omnibus (GEO) repository (<https://www.ncbi.nlm.nih.gov>). Data from CALGB-40603 were obtained from the dbGap repository.

#### Declaration of interests

B. Conte reports speaker fees from Veracyte and payment for educational events from Medsite and Novartis. A. Prat reports consulting fees from Roche, Pfizer, Novartis, Amgen, BMS, Puma, Oncolytics Biotech, MSD, Guardant Health, Peptomyc and Lilly, lecture fees from Roche, Pfizer, Novartis, Amgen, BMS, Nanostring Technologies and Daiichi Sankyo; patents filed PCT/EP2016/080056, PCT/EP2022/086493, PCT/EP2023/060810, EP23382703 and EP23383369; stockholder and consultant of Reveal Genomics, SL; and institutional financial interests from Boehringer, Novartis, Roche, Nanostring, Sysmex Europa GmbH, Medica Scientia Innovation Research, SL, Celgene, Astellas and Pfizer. F. Schettini has declared consulting fees from Pfizer, honoraria for lectures from Novartis, Gilead and Daiichi Sankyo, and travel expenses from Novartis and Daiichi Sankyo. O. Martínez-Sáez has declared institutional grants from Roche; personal consulting fees from Roche and Reveal Genomics; honoraria for presentations by Daiichi Sankyo, Pierre Fabre, and Reveal Genomics; and travel expenses by Gilead, Pierre Fabre, Novartis, and MSD. A. Vivancos has declared institutional grants from Bristol Meyers Squibb,

Incyte, and Roche; personal consulting fees from Bayer, Bristol Meyers Squibb (BMS); Guardant, Incyte, and Roche; and personal stock options from Reveal Genomics. F. Brasó-Maristany has patents filed: PCT/EP2022/086493, PCT/EP2023/060810, EP23382703 and EP23383369. J. Parker as declared individual and institutional royalties from Veracyte–Prosigna, consulting fees from Bristol Meyers Squibb, Reveal Genomics, and GeneCentric, and holds a patent for breast cancer subtyping. Additionally, he has an equity interest in Reveal Genomics. M. Vidal has declared honoraria for presentations from Novartis, Roche, Pfizer, and Daiichi, and travel expenses from Roche and Pfizer. Additionally, she has participated on a Data Safety Monitoring Board or Advisory Board for Novartis and Roche. C. Perou has declared consulting fees and personal stock options from Reveal Genomics. T. Pascual has declared consulting fees from Novartis; honoraria from Novartis, Astra-zeneca, Veracyte, and Argenetics. I. García-Fructuoso has declared honoraria for presentations from Novartis, Daiichi Sankyo, Esteve, GSK; and travel expenses from Novartis, Gilead, Daiichi Sankyo, Lilly, and BMS. L. Paré has declared contract from Reveal Genomics, a HER2DX patent filed (PCT/EP2022/086493), and the TNBCDX patent filed (EP23382703.9). M. Marín-Aguilera has declared contract from Reveal Genomics. P. Villagrasa has declared contract and personal stock options from Reveal Genomics, the HER2DX patent filed (PCT/EP2022/086493), and the DNADX patent filed (EP22382387.3). G. Villacampa has received a speaker's fee from Merck Sharp & Dohme, Pfizer, GlaxoSmithKline and Pierre Fabrer, and received consultant fees from Reveal Genomics. C. Falato is currently employed in AstraZeneca.

The remaining authors declare no competing interests.

#### Acknowledgements

B. Conte received funding the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska–Curie grant agreement No 955951. A. Prat received funding from Fundació CRIS contra el càncer PR\_EX\_2021-14, Agència de Gestió d'Ajuts Universitaris i de Recerca 2021 SGR 01156, Fundació Fero BECA ONCOXXI21, Instituto de Salud Carlos III PI22/01017, Asociación Cáncer de Mama Metastásico IV Premios M. Chiara Giorgetti, Breast Cancer Research Foundation BCRF-22-198, Associació Beca Marta Santamaria, and RESCUER, funded by European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 847912. F. Brasó-Maristany received funding from Fundació científica AECC Ayudas Investigador AECC 2021 (INVES21943BRAS) and Fundació Contigo (HIMALAIA). N. Chic received a Beca FSEOM de Estancia de 2 Años para Formación en Investigación Translacional en Centros de Referencia en el Extranjero. L. Paré, M. Marín-Aguilera, and P. Villagrasa González are currently employed in Reveal Genomics.

#### Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.ebiom.2024.105043>.

#### References

- Dent R, Trudeau M, Pritchard KI, et al. Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res*. 2007;13:4429–4434.
- Hudis CA, Gianni L. Triple-negative breast cancer: an unmet medical need. *Oncol*. 2011;16(Suppl 1):1–11.
- Hernandez-Aya LF, Chavez-MacGregor M, Lei X, et al. Nodal status and clinical outcomes in a large cohort of patients with triple-negative breast cancer. *J Clin Orthod*. 2011;29:2628–2634.
- Korde LA, Somerfield MR, Hershman DL. For the neoadjuvant chemotherapy, endocrine therapy, and targeted therapy for breast cancer guideline expert panel. Use of immune checkpoint inhibitor pembrolizumab in the treatment of high-risk, early-stage triple-negative breast cancer: ASCO guideline rapid recommendation update. *J Clin Orthod*. 2022;40:1696–1698.
- Lehmann BD, Bauer JA, Chen X, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest*. 2011;121:2750–2767.

- 6 Prat A, Parker JS, Karginova O, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res.* 2010;12:R68.
- 7 Loi S, Drubay D, Adams S, et al. Tumor-infiltrating lymphocytes and prognosis: a pooled individual patient analysis of early-stage triple-negative breast cancers. *J Clin Orthod.* 2019;37:559–569.
- 8 Loi S, Sirtaine N, Piette F, et al. Prognostic and predictive value of tumor-infiltrating lymphocytes in a phase III randomized adjuvant breast cancer trial in node-positive breast cancer comparing the addition of docetaxel to doxorubicin with doxorubicin-based chemotherapy: BIG 02-98. *J Clin Orthod.* 2013;31:860–867.
- 9 Denkert C, Loibl S, Noske A, et al. Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer. *J Clin Oncol.* 2010;28:105–113.
- 10 Denkert C, von Minckwitz G, Darb-Esfahani S, et al. Tumor-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. *Lancet Oncol.* 2018;19:40–50.
- 11 Dieci MV, Mathieu MC, Guarneri V, et al. Prognostic and predictive value of tumor-infiltrating lymphocytes in two phase III randomized adjuvant breast cancer trials. *Ann Oncol.* 2015;26:1698–1704.
- 12 de Jong VMT, Wang Y, Ter Hoeve ND, et al. Prognostic value of stromal tumor-infiltrating lymphocytes in young, node-negative, triple-negative breast cancer patients who did not receive (neo) Adjuvant systemic therapy. *J Clin Oncol.* 2022;40:2361–2374.
- 13 Jézéquel P, Loussouarn D, Guérin-Charbonnel C, et al. Gene-expression molecular subtyping of triple-negative breast cancer tumours: importance of immune response. *Breast Cancer Res.* 2015;17:43.
- 14 Abramson VG, Lehmann BD, Ballinger TJ, Pietenpol JA. Subtyping of triple-negative breast cancer: implications for therapy. *Cancer.* 2015;121:8–16.
- 15 Adams S, Gray RJ, Demaria S, et al. Prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancers from two phase III randomized adjuvant breast cancer trials: ECOG 2197 and ECOG 1199. *J Clin Oncol.* 2014;32:2959–2966.
- 16 Denkert C, von Minckwitz G, Brase JC, et al. Tumor-infiltrating lymphocytes and response to neoadjuvant chemotherapy with or without carboplatin in human epidermal growth factor receptor 2-positive and triple-negative primary breast cancers. *J Clin Oncol.* 2015;33:983–991.
- 17 Salgado R, Denkert C, Demaria S, et al. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. *Ann Oncol.* 2015;26:259–271.
- 18 Liu S, Lachapelle J, Leung S, Gao D, Foulkes WD, Nielsen TO. CD8+ lymphocyte infiltration is an independent favorable prognostic indicator in basal-like breast cancer. *Breast Cancer Res.* 2012;14:R48.
- 19 Mahmoud SMA, Paish EC, Powe DG, et al. An evaluation of the clinical significance of FOXP3+ infiltrating cells in human breast cancer. *Breast Cancer Res Treat.* 2011;127:99–108.
- 20 Miyashita M, Sasano H, Tamaki K, et al. Prognostic significance of tumor-infiltrating CD8+ and FOXP3+ lymphocytes in residual tumors and alterations in these parameters after neoadjuvant chemotherapy in triple-negative breast cancer: a retrospective multicenter study. *Breast Cancer Res.* 2015;17:124.
- 21 Raskov H, Orhan A, Christensen JP, Gögenur I. Cytotoxic CD8+ T cells in cancer and cancer immunotherapy. *Br J Cancer.* 2021;124:359–367.
- 22 Harris RJ, Cheung A, Ng JCF, et al. Tumor-infiltrating B lymphocyte profiling identifies IgG-biased, clonally expanded prognostic phenotypes in triple-negative breast cancer. *Cancer Res.* 2021;81:4290–4304.
- 23 Hu Q, Hong Y, Qi P, et al. Atlas of breast cancer infiltrated B-lymphocytes revealed by paired single-cell RNA-sequencing and antigen receptor profiling. *Nat Commun.* 2021;12:2186.
- 24 Shepherd JH, Ballman K, Polley M-YC, et al. CALGB 40603 (alliance): long-term outcomes and genomic correlates of response and survival after neoadjuvant chemotherapy with or without carboplatin and bevacizumab in triple-negative breast cancer. *J Clin Oncol.* 2022;40:1323–1334.
- 25 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. <https://doi.org/10.1016/j.ebiom.2021.103801>.
- 26 Villacampa G, Tung NM, Pernas S, et al. Association of HER2DX with pathological complete response and survival outcomes in HER2-positive breast cancer. *Ann Oncol.* 2023;34:783–795.
- 27 Waks AG, Ogayo ER, Paré L, et al. Assessment of the HER2DX assay in patients with ERBB2-positive breast cancer treated with neoadjuvant paclitaxel, trastuzumab, and pertuzumab. *JAMA Oncol.* 2023;9:835–840.
- 28 Bueno-Muiño C, Echavarría I, López-Tarruella S, et al. Assessment of a genomic assay in patients with ERBB2-positive breast cancer following neoadjuvant trastuzumab-based chemotherapy with or without pertuzumab. *JAMA Oncol.* 2023;9:841–846.
- 29 Tolaney SM, Tarantino P, Graham N, et al. Adjuvant paclitaxel and trastuzumab for node-negative, HER2-positive breast cancer: final 10-year analysis of the open-label, single-arm, phase 2 APT trial. *Lancet Oncol.* 2023;24:273–285.
- 30 Garcia-Saenz JA, Blancas I, Echavarría I, et al. SEOM-GEICAM-SOLTI clinical guidelines in advanced breast cancer (2022). *Clin Transl Oncol.* 2023;25:2665–2678.
- 31 Pereira B, Chin S-F, Rueda OM, et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat Commun.* 2016;7:11479.
- 32 Rueda OM, Sammut S-J, Seoane JA, et al. Dynamics of breast-cancer relapse reveal late-recurring ER-positive genomic subgroups. *Nature.* 2019;567:399–404.
- 33 Staaf J, Häkkinen J, Hegardt C, et al. RNA sequencing-based single sample predictors of molecular subtype and risk of recurrence for clinical assessment of early-stage breast cancer. *NPJ Breast Cancer.* 2022;8:1–17.
- 34 Brueffer C, Vallon-Christersson J, Grabau D, et al. Clinical value of RNA sequencing-based classifiers for prediction of the five conventional breast cancer biomarkers: a report from the population-based multicenter Sweden cancerome analysis network—breast initiative. *JCO Precis Oncol.* 2018;2:1–18.
- 35 Sabatier R, Finetti P, Cervera N, et al. A gene expression signature identifies two prognostic subgroups of basal breast cancer. *Breast Cancer Res Treat.* 2011;126:407–420.
- 36 Sikov WM, Berry DA, Perou CM, et al. Impact of the addition of carboplatin and/or bevacizumab to neoadjuvant once-per-week paclitaxel followed by dose-dense doxorubicin and cyclophosphamide on pathologic complete response rates in stage II to III triple-negative breast cancer: CALGB 40603 (Alliance). *J Clin Oncol.* 2015;33:13–21.
- 37 Loibl S, O’Shaughnessy J, Untch M, et al. Addition of the PARP inhibitor veliparib plus carboplatin or carboplatin alone to standard neoadjuvant chemotherapy in triple-negative breast cancer (BrightTNESS): a randomised, phase 3 trial. *Lancet Oncol.* 2018;19:497–509.
- 38 Metzger-Filho O, Collier K, Asad S, et al. Matched cohort study of germline BRCA mutation carriers with triple negative breast cancer in brightness. *NPJ Breast Cancer.* 2021;7:142.
- 39 Fan C, Prat A, Parker JS, et al. Building prognostic models for breast cancer patients using clinical variables and hundreds of gene expression signatures. *BMC Med Genom.* 2011;4:3.
- 40 Fernandez-Martinez A, Pascual T, Singh B, et al. Prognostic and predictive value of immune-related gene expression signatures vs tumor-infiltrating lymphocytes in early-stage ERBB2/HER2-positive breast cancer: a correlative analysis of the CALGB 40601 and PAMELA trials. *JAMA Oncol.* 2023;9:490.
- 41 Prat A, Guarneri V, Paré L, et al. A multivariable prognostic score to guide systemic therapy in early-stage HER2-positive breast cancer: a retrospective study with an external evaluation. *Lancet Oncol.* 2020;21:1455–1464.
- 42 Guarneri V, Bras-Maristany F, Dieci MV, et al. HER2DX genomic test in HER2-positive/hormone receptor-positive breast cancer treated with neoadjuvant trastuzumab and pertuzumab: A correlative analysis from the PerELISA trial. *EBioMedicine.* 2022;85:104320.
- 43 Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* 2015;12:453–457.
- 44 Luca BA, Steen CB, Matusiak M, et al. Atlas of clinically distinct cell states and ecosystems across human solid tumors. *Cell.* 2021;184:5482–5496.e28.
- 45 Perou CM, Sørlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature.* 2000;406:747–752.
- 46 Parker JS, Mullins M, Cheang MCU, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol.* 2009;27:1160–1167.

- 47 Cardoso F, Kyriakides S, Ohno S, et al. Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2019;30:1194–1220.
- 48 Int'Hout J, Ioannidis JP, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med Res Methodol*. 2014;14:25.
- 49 Geyer CE, Sikov WM, Huober J, et al. Long-term efficacy and safety of addition of carboplatin with or without veliparib to standard neoadjuvant chemotherapy in triple-negative breast cancer: 4-year follow-up data from BrighTNess, a randomized phase III trial. *Ann Oncol*. 2022;33:384–394.
- 50 Poggio F, Bruzzone M, Ceppi M, et al. Platinum-based neoadjuvant chemotherapy in triple-negative breast cancer: a systematic review and meta-analysis. *Ann Oncol*. 2018;29:1497–1508.
- 51 Poggio F, Tagliamento M, Ceppi M, et al. Adding a platinum agent to neoadjuvant chemotherapy for triple-negative breast cancer: the end of the debate. *Ann Oncol*. 2022;33:347–349.
- 52 Sharma P, Kimler BF, O'Dea A, et al. Randomized phase II trial of anthracycline-free and anthracycline-containing neoadjuvant carboplatin chemotherapy regimens in stage I–III triple-negative breast cancer (NeoSTOP). *Clin Cancer Res*. 2021;27:975–982.
- 53 Sharma P, López-Tarruella S, García-Saenz JA, et al. Efficacy of neoadjuvant carboplatin plus docetaxel in triple-negative breast cancer: combined analysis of two cohorts. *Clin Cancer Res*. 2017;23:649–657.
- 54 Sharma P, López-Tarruella S, García-Saenz JA, et al. Pathological response and survival in triple-negative breast cancer following neoadjuvant carboplatin plus docetaxel. *Clin Cancer Res*. 2018;24:5820–5829.
- 55 Ando M, Yamauchi H, Aogi K, et al. Randomized phase II study of weekly paclitaxel with and without carboplatin followed by cyclophosphamide/epirubicin/5-fluorouracil as neoadjuvant chemotherapy for stage II/III breast cancer without HER2 overexpression. *Breast Cancer Res Treat*. 2014;145:401–409.
- 56 Zhang P, Yin Y, Mo H, et al. Better pathologic complete response and relapse-free survival after carboplatin plus paclitaxel compared with epirubicin plus paclitaxel as neoadjuvant chemotherapy for locally advanced triple-negative breast cancer: a randomized phase 2 trial. *Oncotarget*. 2016;7:60647–60656.
- 57 Schmid P, Cortes J, Pusztai L, et al. Pembrolizumab for early triple-negative breast cancer. *N Engl J Med*. 2020;382:810–821.
- 58 Masuda N, Lee S-J, Ohtani S, et al. Adjuvant capecitabine for breast cancer after preoperative chemotherapy. *N Engl J Med*. 2017;376:2147–2159.
- 59 Tutt ANJ, Garber JE, Kaufman B, et al. Adjuvant olaparib for patients with BRCA1- or BRCA2-mutated breast cancer. *N Engl J Med*. 2021;384:2394–2405.
- 60 Zhai Z, Zheng Y, Yao J, et al. Evaluation of adjuvant treatments for T1 N0 M0 triple-negative breast cancer. *JAMA Netw Open*. 2020;3:e2021881.
- 61 Shum K, Hussein A, Hamm C. Are we overtreating stage I triple-negative breast cancer in Ontario? A population-based retrospective epidemiological analysis using the ICES database. *Med Oncol*. 2022;39:228.
- 62 Loibl S, Schneeweiss A, Huober J, et al. Neoadjuvant durvalumab improves survival in early triple-negative breast cancer independent of pathological complete response. *Ann Oncol*. 2022;33:1149–1158.
- 63 Mittendorf EA, Zhang H, Barrios CH, et al. Neoadjuvant atezolizumab in combination with sequential nab-paclitaxel and anthracycline-based chemotherapy versus placebo and chemotherapy in patients with early-stage triple-negative breast cancer (IMpassion031): a randomised, double-blind, phase 3 trial. *Lancet*. 2020;396:1090–1100.
- 64 Gianni L, Huang C-S, Egle D, et al. Abstract GS3-04: pathologic complete response (pCR) to neoadjuvant treatment with or without atezolizumab in triple negative, early high-risk and locally advanced breast cancer. NeoTRIPaPDL1 Michelangelo randomized study. *Cancer Res*. 2020;80. GS3-04.
- 65 Sharma P, Stecklein SR, Yoder R, et al. Abstract PD11-07: PD11-07 Association of TNBC-DX scores with outcomes in triple-negative breast cancer (TNBC) treated with neoadjuvant pembrolizumab and chemotherapy: a correlative analysis from NeoPACT and NeoSTOP trials. *Cancer Res*. 2023;83:PD11–PD17.
- 66 Cortazar P, Zhang L, Untch M, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet*. 2014;384:164–172.
- 67 Symmans WF, Peintinger F, Hatzis C, et al. Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. *J Clin Oncol*. 2007;25:4414–4422.
- 68 Symmans WF, Wei C, Gould R, et al. Long-term prognostic risk after neoadjuvant chemotherapy associated with residual cancer burden and breast cancer subtype. *J Clin Oncol*. 2017;35:1049–1060.
- 69 Lauss M, Donia M, Svane IM, Jönsson G. B cells and tertiary lymphoid structures: friends or foes in cancer immunotherapy? *Clin Cancer Res*. 2022;28:1751–1758.
- 70 Hernán MA. The hazards of hazard ratios. *Epidemiology*. 2010;21:13–15.