CLINICAL TRIAL



# Reducing chemotherapy use in clinically high-risk, genomically low-risk pN0 and pN1 early breast cancer patients: five-year data from the prospective, randomised phase 3 West German Study Group (WSG) PlanB trial

Ulrike Nitz<sup>1,2</sup> · Oleg Gluz<sup>1,2</sup> · Matthias Christgen<sup>3</sup> · Ronald E. Kates<sup>1</sup> · Michael Clemens<sup>4</sup> · Wolfram Malter<sup>5</sup> · Benno Nuding<sup>6</sup> · Bahriye Aktas<sup>7</sup> · Sherko Kuemmel<sup>8</sup> · Toralf Reimer<sup>9</sup> · Andrea Stefek<sup>10</sup> · Fatemeh Lorenz-Salehi<sup>11</sup> · Petra Krabisch<sup>12</sup> · Marianne Just<sup>13</sup> · Doris Augustin<sup>14</sup> · Cornelia Liedtke<sup>1,15</sup> · Calvin Chao<sup>16</sup> · Steven Shak<sup>16</sup> · Rachel Wuerstlein<sup>1,17</sup> · Hans H. Kreipe<sup>3</sup> · Nadia Harbeck<sup>1,17</sup>

Received: 16 June 2017/Accepted: 23 June 2017/Published online: 29 June 2017 © The Author(s) 2018, corrected publication 2019

#### Abstract

*Background* The prospective phase 3 PlanB trial used the Oncotype DX<sup>®</sup> Recurrence Score<sup>®</sup> (RS) to define a genomically low-risk subset of clinically high-risk pN0-1 early breast cancer (EBC) patients for treatment with adjuvant endocrine therapy (ET) alone. Here, we report five-year data evaluating the prognostic value of RS, Ki-67, and other traditional clinicopathological parameters.

Ulrike Nitz and Oleg Gluz have equally contributed to this work.

Oleg Gluz oleg.gluz@wsg-online.com

- <sup>1</sup> West German Study Group, Ludwig-Weber-Str. 15, 41061 Moenchengladbach, Germany
- <sup>2</sup> Ev. Hospital Bethesda, Breast Center Niederrhein, Ludwig-Weber-Str. 15, 41061 Moenchengladbach, Germany
- <sup>3</sup> Institute of Pathology, Medical School Hannover, Carl-Neuberg-Str. 1., 30625 Hannover, Germany
- <sup>4</sup> Department of Oncology, Clinics Mutterhaus der Borromäerinnen, Feldstraße 16, 54290 Trier, Germany
- <sup>5</sup> Department of Gynecology and Obstetrics, Breast Center, University Clinics Cologne, Kerpener Str. 62, 50937 Cologne, Germany
- <sup>6</sup> Department of Gynecology and Obstetrics, Evangelical Hospital, Ferrenbergstraße 24, 51465 Bergisch Gladbach, Germany
- <sup>7</sup> Department of Gynecology and Obstetrics, University Clinics Essen, Hufelandstraße 55, 45147 Essen, Germany
- <sup>8</sup> Clinics Essen-Mitte, Breast Centre, Henricistraße 92, 45136 Essen, Germany

*Methods* A central tumour bank was prospectively established within PlanB. Following an early amendment, hormone receptor (HR)+ , pN0-1 RS  $\leq$  11 patients were recommended to omit chemotherapy. Patients with RS  $\geq$  12, pN2-3, or HR-negative/HER2-negative disease were randomised to anthracycline-containing or anthracycline-free chemotherapy. Primary endpoint: disease-free survival (DFS). PlanB Clinicaltrials.gov identifier: NCT01049425.

*Findings* From 2009 to 2011, PlanB enrolled 3198 patients (central tumour bank, n = 3073) with the median age of

- <sup>9</sup> Department of Gynecology and Obstetrics, Clinics Suedstadt, Südring 81, 18059 Rostock, Germany
- <sup>10</sup> Johanniter Clinics Stendal, Breast Center Altmark, Bahnhofstraße 24 – 26, 39576 Stendal, Germany
- <sup>11</sup> Department of Gynecology, Dr. Horst-Schmidt Clinics, Ludwig-Erhard-Straße 100, 65199 Wiesbaden, Germany
- <sup>12</sup> Klinikum ChemnitzFlemmingstraße 2, 09116 Chemnitz, Germany
- <sup>13</sup> Oncological Practice, Teutoburger Str. 60 33604, Bielefeld, Germany
- <sup>14</sup> Breast Centre Ostbayern, Perlasberger Str. 41, 94469 Deggendorf, Germany
- <sup>15</sup> Department of Gynecology and Obstetrics, University Hospital Schleswig–Holstein Campus Luebeck, Ratzeburger Allee 160, 23562 Luebeck, Germany
- <sup>16</sup> Genomic Health, Inc., Redwood City, CA, USA
- <sup>17</sup> Department of Gynecology and Obstetrics, Breast Center, University of Munich (LMU) and CCCLMU, Munich, Germany

56 years, 41.1% pN+, and 32.5% grade 3 EBC. Chemotherapy was omitted in 348/404 (86.1%) eligible  $RS \le 11$  patients. After 55 months of median follow-up, five-year DFS in ET-treated RS  $\leq$  11 patients was 94% (in both pN0 and pN1) versus 94% (RS 12-25) and 84% (RS > 25) in chemotherapy-treated patients (p < 0.001); five-year overall survival (OS) was 99 versus 97% and 93%, respectively (p < 0.001). Nodal status, central/local grade, tumour size, continuous Ki-67, progesterone receptor (PR), IHC4, and RS were univariate prognostic factors for DFS. In a multivariate analysis including all univariate prognostic markers, only pN2-3, central and local grade 3, tumour size >2 cm, and RS, but not IHC4 or Ki-67 were independent adverse factors. If RS was excluded, IHC4 or both Ki-67 and PR entered the model. The impact of RS was particularly pronounced in patients with intermediate Ki-67 (>10%, <40%) tumours.

Interpretation The excellent five-year outcomes in clinically high-risk, genomically low-risk (RS  $\leq$  11) pN0-1 patients without adjuvant chemotherapy support using RS with standardised pathology for treatment decisions in HR+ HER2-negative EBC. Ki-67 has the potential to support patient selection for genomic testing.

Keywords Genomic signature  $\cdot$  Oncotype DX  $\cdot$  Breast cancer  $\cdot$  Ki-67  $\cdot$  IHC4

# Introduction

Adjuvant treatment decisions in hormone receptor (HR)+/ HER2-negative early breast cancer (EBC) have traditionally relied on clinicopathological characteristics (nodal status, tumour size/grade, age, and co-morbidities). A recent meta-analysis of standard chemotherapy (CT) regimens demonstrated that proportional risk reductions were little affected by age, nodal status, tumour diameter/differentiation, ER status, and tamoxifen use [1].

Increasing evidence suggests that HR+/HER2-negative EBC is a heterogeneous disease [2]. Specifically, patients with molecularly classified [2] luminal-A or immunohistochemically (IHC) classified (Ki-67, progesterone receptor (PR)) luminal-A-like tumours have excellent prognosis [3–6]. However, clinical utility of Ki-67 is limited by interobserver variability [7, 8], despite improvement by quality assurance [9, 10].

Since the advent of genomic signatures, efforts have been undertaken to integrate them into clinical routine for pN0-1 HR+ patients. Until 2015, most clinical evidence for genomic signatures derived from retrospective analyses of prospective trials conducted in the early-CT era [11–15] or used archived samples from untreated patients [16]. Considerable retrospective evidence exists for reproducible prognostic impact of Oncotype DX<sup>®</sup>/Recurrence Score<sup>®</sup> (RS), MammaPrint<sup>®</sup>, EndoPredict<sup>®</sup>, and Prosigna<sup>®</sup>, but their value beyond centrally measured IHC markers or derived scores like IHC4 remains uncertain; [4, 5, 17] the RS is the only test shown to predict CT benefit [12, 18].

While many national/international EBC guidelines incorporate genomic signatures [19], prospective trials exploring CT overtreatment/under-treatment are needed to assess benefits of augmenting conventional pathology with genomic signatures.

Recent prospective evidence has generally confirmed retrospective results: for MammaPrint<sup>®</sup>, MINDACT [20]; for RS, the registry part of TAILORx (in N0 disease) [21]; and the WSG-PlanB three-year analysis in clinically high-risk EBC [22]. Nevertheless, the utility of genomic testing in high-risk patients and the optimal selection for these tests remain unclear.

PlanB is a prospective, randomised multicentre phase 3 CT trial in HER2-negative EBC. RS was prospectively assessed in HR+ patients; following an early amendment, RS  $\leq$  11 participants with  $\leq$ 3 involved lymph nodes (LN) could omit CT; RS  $\leq$  11 patients had excellent three-year survival [22]. The planned translational analyses included independent central pathology review of grade/IHC markers and their utilisation with RS for prognosis in nodal and KI-67 subgroups, based on the "five-year" (median 55-month) PlanB follow-up.

#### Methods

# Study participants

The trial included female patients, 18–75 years, with histologically confirmed, unilateral primary invasive BC, adequate surgical treatment (free margins, sentinel-node biopsy in node-negative, or axillary dissection in nodepositive patients), without evidence of metastasis. Key inclusion criteria: HER2-negativity; pT1-T4c; pN+ [or pN0 with a risk factor ( $\geq$ pT2, grade 2/3, high uPA/PAI-1, <35 years, or HR-negative)]; ECOG performance status <2 or Karnofsky Index  $\geq$  80%; signed informed consent; and (if  $\geq$ 4 positive LN, RS > 11, or HR-negative) willingness to participate in the adjuvant CT PlanB trial.

# Study design

WSG-PlanB was approved by German ethics boards and conducted in accordance with the Declaration of Helsinki. Clinicaltrials.gov identifier: NCT01049425.

PlanB (CONSORT diagram: Fig. 1) began in 2009 as a CT trial comparing anthracycline-containing (four cycles of epirubicin/cyclophosphamide followed by four cycles of



Fig. 1 PlanB CONSORT diagram. CT chemotherapy, ET endocrine therapy, HR hormone receptor, RS recurrence score

docetaxel q3w) and anthracycline-free (six cycles of docetaxel/cyclophosphamide q3w) CT. After including 274 patients, the trial was amended (08/2009) to recommend endocrine therapy (ET) alone for pN0/pN1, locally HR+ patients with RS  $\leq$  11 (based on an initial RS validation study [11]).

Follow-up was performed at three-month intervals for the first three years and every six months thereafter. Data were obtained from electronic case record forms and verified by monitoring visits to the study sites.

RS was assessed on surgically removed primary tumour tissue at the central laboratory of Genomic Health Inc. (Redwood City, CA).

Slide review, IHC, and fluorescence in situ hybridisation were analysed in an independent central laboratory (Institute of Pathology, Hannover Medical School) [22]. Tumours were considered centrally ER+ (antibody SP-1) or PR+ (antibody PgR636, both Zytomed, Berlin, Germany) if immunostaining was present in  $\geq 1\%$  of tumour nuclei. Ki-67 was assessed centrally using rabbit monoclonal Ki-67 antibody 30–9 (Ventana Inc.Tucson, USA) on  $\geq 100$  invasive tumour cells; the semi-quantitative procedure for Ki-67 produces values in 5% increments. Hence, analysed semi-quantitative ranges Ki-67  $\leq 10\%$ ,  $\geq 40\%$  correspond to quantitative KI-67  $\leq 13.25\%^3$ , Ki-67 > 35% [23], respectively; analyses were also performed for semi-quantitative Ki-67 ranges <20% and  $\geq 20\%^6$ . IHC4 was computed as previously described [4, 17].

# Endpoints

The endpoints included prospective evaluation of RS prognostic impact at follow-up target of five years: Clinical outcomes (disease-free survival [DFS], overall survival [OS]) in RS  $\leq$  11 patients treated with ET alone, and prospective evaluation of the prognostic value of other parameters (Ki-67, IHC4 and histological grade [Elston-Ellis] by local/central assessment). Study was performed according to the reporting recommendations for tumour marker prognostic studies (REMARK) guidelines [24].

# Statistical analysis

For DFS analysis, an event was defined as any invasive cancer event or death (with/without recurrence). Estimates of five-year DFS or OS with approximate 95% confidence intervals [given in brackets] were obtained by the Kaplan–

Meier method. Comparisons of DFS or OS among subgroups used pairwise log-rank tests (reported as significant for p < 0.05).

Subgroup analyses were performed in RS  $\leq$  11, RS 12–25, RS > 25, and in Ki-67 subgroups (see below). Univariate and multivariate (forward elimination) Cox proportional hazard models for DFS were estimated; RS, Ki-67, ER, PR, and IHC4 were coded as continuous variables using fractional ranks. For a realistic measure of effect sizes, hazard ratios of fractionally ranked variables are reported for 75th versus 25th percentile. Factors with significant impact on DFS or OS are referred to as "prognostic". Nodal status was coded as pN1-3 *vs* nodenegative, pN2-3 versus pN0-1, and pN3 versus pN0-2; tumour stage was coded as pT2-4 versus pT1; local and central grades were coded as grade 3 versus grade 1–2. Statistical analyses were performed using SPSS v.23 (IBM Corp., Armonk, NY).

#### Role of the funding source

The industry supporters in this trial had no role in study design, data collection, analysis/interpretation, writing, or decision to submit the manuscript. The authors (OG, UN, NH, RK) had full data access and hold final responsibility for manuscript submission.

# RESULTS

#### Study participants

From 4/2009 to 12/2011, 3198 patients were recruited from 93 centres; 2449 were randomised for CT. As previously reported [22], the central tumour bank population included 2642 locally HR+ cases (97.4% centrally confirmed). The analyses presented here focus on locally HR + patients (all pN) unless otherwise stated. Their median age was 56 years; median tumour size was 19 mm; 61.9% had central grade 2; 58.8% were N0; 35.2% were pN1; 6.0% were pN2-3; 2553 (96.6%) had available RS: 17.4% RS  $\leq$  11, 58.4% RS 12–25, and 20.8% RS > 25 (Table 1).

Within the 2274 locally HR+ pN0-1 post-amendment patients, 404 (17.8%) had RS  $\leq$  11; CT was omitted in 348 (86.1%) of these patients: 238 (68.4%) pN0 and 110 (31.6%) pN1.

435 tumours were HR-negative/HER2– by central pathology (local HR-negative status centrally confirmed in 93.5%).

Compliance with treatment recommendations within the trial was 95.2% for pN0 and 75.2% for pN1.

Median follow-up was 55 months (range 3–72 months).

 
 Table 1
 Patient characteristics in ER and/or PR positive population (by local assessment)

Characteristic	N = 2642
Age (years)	
Median (range)	56 (25-77)
Tumour size (mm)	
Median (range)	19 (1.2–130.)
Nodal status [N (%)]	
pN0	1554 (58.8)
pN1	930 (35.2)
pN2	122 (4.6)
pN3	36 (1.4)
Therapy [N (%)]	
Endocrine	348 (13.0)
Chemotherapy randomization	1970 (74.6)
Out of study	324 (12.4)
RS result [N (%)]	
≤11	459 (17.4)
12–25	1544 (58.4)
> 25	550 (20.8)
Unknown	89 (3.4)
Central hormone receptor status, [N (%)]	
ER positive	2118 (89.7)
PR positive	1751 (74.1)
ER and/or PR positive	2202 (93.2)
ER and PR negative	57 (2.4)
Unknown	160 (6.8)
Central grade [N (%)]	
Grade 1	134 (5.1)
Grade 2	1636 (61.9)
Grade 3	825 (31.2)
Unknown	47 (1.8)
Median Ki-67 (%)	15

ER oestrogen receptor, PR progesterone receptor, RS recurrence score

# DFS in RS groups, impact of nodal status, and Ki-67 within these RS groups

In locally HR+ patients with follow-up (including all pN), DFS was higher in RS  $\leq 11$  and RS 12–25 compared to RS > 25 (both p < 0.001); five-year DFS was 93.6% [90.8–96.4%] in low-RS (about three-fourths receiving no CT), 94.3% [92.8–95.8%] in intermediate RS (all CT-treated), and 84.2% [80.6–87.8%] in high-RS patients (all CT-treated) (Fig. 2a). Five-year DFS in those RS  $\leq 11$  patients treated with ET alone was 94.2% [91.2–97.3%] and was similar in pN0 (n = 238) (94.2%; [90.4–98.0%]) and pN1 (n = 110) subgroups (94.4%; [89.5–99.3%]).



Fig. 2 DFS for patients with RS  $\leq$  11, 12–25, and > 25 overall (a), node-negative patients (b), for patients with pN1 disease (c), and for patients with pN2-3 disease (d)

Higher DFS in RS  $\leq$  11 and RS12-25 compared to RS  $\geq$  25 patients held separately in N0, N1, and N2 + (Fig. 2b-d). In pN2-3 patients with RS > 25, five-year DFS was 61.7% [46.6–76.8%].

Within RS subgroups, a significant impact of Ki-67 on DFS was seen only for RS > 25. In terms of fractionally ranked Ki-67, the hazard ratio (75th to 25th percentile) was 2.62 [1.39–4.91] (p = 0.003), i.e. poorer DFS. The corresponding CI in low and intermediate RS were [0.39–3.01] and [0.87–2.82], respectively. (Similar results were seen for grouped Ki-67).

# DFS in Ki-67 subgroups and impact of RS within these subgroups

DFS was assessed (Appendix Fig. 5) in ("luminal") patients (all pN) with central HR+ and HER2-negative status and follow-up for the Ki-67 subgroups "low"

(0–10%, n = 810), "intermediate" (>10%, <40%, n = 988), and "high" (≥40%, n = 74); and for comparison in triple-negative (TN) (n = 405) patients (central HR-negative and HER2-negative); CT was administered in 79.9%, 86.7%, 97.3%, and 100%, respectively. Five-year DFS rates were 94.7% [92.9–96.6%], 91.0% [89.0–93.0%], 73.4% [61.8–84.9%], and 80.3% [76.2%–84.4%], respectively (p < 0.003 for "intermediate" versus "low" Ki-67, p = 0.5 for "high" versus "TN", p < 0.001 for the other four comparisons).

Figure 4 illustrates the impact of RS groups on DFS within the low-Ki-67 and intermediate Ki-67 subgroups defined above (there were only 62 patients with follow-up and measured RS in the entire high-Ki-67 subgroup, almost all with high RS, Table 2). Within the intermediate Ki-67 subgroup, RS > 25 was associated with poorer DFS than either in RS  $\leq$  11 or RS 12–25 (both *p* < 0.001, log-rank); five-year DFS was 94.4% [89.9–98.9%], 93.8%

Table 2 Joint distribution of recurrence score (RS) and semiquantitative Ki-67 in the analysed population

	(semi-quantitative) Ki-67 group								
	0-10%	15%	20%	25%	30%	35%	40%	>40%	Total
<b>RS</b> ≤ 11									
Ν	223	87	68	23	1	1	2	0	405
% of Ki-67 group	22.7%	20.1%	19.1%	10.6%	1.0%	2.5%	4.7%	0.0%	0.0%
RS12-25									
Ν	680	283	219	108	42	16	5	2	1355
% of Ki-67 group	69.3%	65.5%	61.5%	50.0%	40.8%	40.0%	11.6%	6.3%	0.0%
RS > 25									
Ν	78	62	69	85	60	23	36	30	443
% of Ki-67 group	8.0%	14.4%	19.4%	39.4%	58.3%	57.5%	83.7%	93.8%	0.0%

[91.5-96.1%], and 83.6% [78.7-88.6%] in these RS  $\leq 11$ , RS 12–25, and RS > 25 patients, respectively; the corresponding hazard ratios for RS > 25 were 3.61 [1.61–8.13] versus RS < 11 and 2.82 [1.72–4.61] versus RS 12–25. Within the low-Ki-67 subgroup, an impact of RS was not seen, but only 33 events occurred in the whole low-Ki-67 subgroup (94.7% five-year DFS, see above), and only 8.4% had high RS.

Ν % of Total Ν

% of RS group

981

44.5%

432

19.6%

356

16.2%

216

9.8%

103

4.7%

In pN0-1 patients, five-year DFS rates for Ki-67 low (0-10%), intermediate (>10%, <40%), and high (>40%) were 95.0% [93.1-96.8%], 92.0% [90.0-94.0%], and 75.4% [63.4-87.3%], respectively. Within intermediate Ki-67 (but restricting to pN0-1), the relative impact of RS groups on DFS was as above: RS > 25 was associated with poorer DFS than either RS < 11 or RS12-25 (both p < 0.001, logrank). Five-year DFS was 94.5% [89.8-99.3%], 93.8% [91.5-96.2%], and 87.1% [82.3-91.9%] in RS  $\leq 11$ , RS 12-25, and RS > 25 patients, respectively (pN0-1, intermediate Ki-67).

Separately for two subgroups defined by Ki-67  $\ge$  20 and Ki-67 < 20 (St. Gallen-inspired cutoff), high-RS (>25) patients had poorer DFS than low- or intermediate RS patients (all pN): The hazard ratios for RS > 25 versus RS 0–25 within Ki-67  $\geq$  20% and Ki-67 < 20% subgroups were 2.69 [1.65-4.40] and 2.14 [1.07-4.26], respectively. Again,  $RS \le 11$  and RS12-25 subgroups had similar fiveyear DFS: 92.2 versus 92.3% for Ki-67  $\geq$  20% and 94.9 versus 95.1% for Ki-67 < 20%, respectively.

# Univariate and multivariate analyses of DFS

In univariate analyses of locally HR+ patients with available RS (Table 2), nodal status (pN2-3 versus pN0-1, pN3 versus pN0-2), central and local grade 3 (vs. grade 1 or 2), tumour size >2 cm, as well as continuous, fractionally

ranked RS, Ki-67, PR, and IHC, were all significant factors for DFS; higher levels of all these factors were unfavourable except for PR.

40

1.8%

43

2.0%

32

1.5%

2203

100.0%

Table 3 also shows the results of a multivariate analysis including all the markers identified by univariate analysis (and fractionally ranked ER). Besides nodal status (pN2-3 vs. pN0-1, pN3 vs. pN0-2), grade 3 (independently by central and local assessment), tumour size >20 mm, and fractionally ranked RS-but not IHC4 or Ki-67-were independent factors for poorer DFS. The same results were true if analysis of the model was limited to chemotherapyonly-treated patients. When RS was excluded from the multivariate model, IHC4 or both Ki-67 and PR became independent predictors for DFS. When local grade was excluded from the multivariate model, Ki-67 became independent prognostic factor for DFS.

#### OS analysis by RS and nodal status

Consistent with DFS, among all locally HR + patients with available RS, better OS was observed in RS < 11 or RS12–25 patients than in RS > 25 (p < 0.001 for both comparisons). Five-year OS (Fig. 3a) was 99.1% [98.5-100%] in RS  $\leq 11, 97.2\%$  [96.0-98.5%] in RS 12-25, and 93.3% [90.8%-95.8%] in high-RS (>25) patients, despite only about one-fourth of these  $RS \leq 11$  patients receiving CT, compared to all RS > 11 patients; moreover, five-year OS in those RS < 11 patients receiving ET alone was also 99.1%. The differences correspond to guite substantial hazard ratios of 6.46 [2.27–18.42] for RS > 25 versus RS  $\leq$  11 and 3.26 [1.87–5.70] for RS > 25 versus RS 12–25.

In node-negative patients, five-year OS was a remarkable 99.2% [98.0-100%] in RS < 11 compared to 98.3% [97.0-99.5%] in RS 12-25 and 96.7% [94.4-99.0%] in RS > 25. OS in RS 12-25 was significantly higher than in

579

Table 3 Univ	variate and	multivariate	analysis	(DFS)	for all	locally	HR	+	tumours	with	available	RS
--------------	-------------	--------------	----------	-------	---------	---------	----	---	---------	------	-----------	----

Variable	Coding	Univariate HR (95% CI)	p value	Multivariate HR (95% CI)	p value
RS	Fractionally ranked	2.33 (1.73-3.14)*	<0.001	1.73 (1.21-2.47)*	0.001
Nodal status					
	pN1-3 vs pN0	1.55 (1.11-2.15)	0.009	NS	
	pN2-3 vs pN0-1	3.23 (2.11-4.94)	<0.001	2.24 (1.27-3.96)	0.005
	pN3 vs pN0-2	6.35 (3.34–12.07)	<0.001	2.88 (1.27-6.52)	0.011
Tumour stage	pT2-4 vs pT1	1.77 (1.27–2.47)	0.001	1.49 (1.04–2.15)	0.04
Grade					
Local assessment	Grade 3 vs grade 1/2	2.36 (1.69-3.29)	<0.001	1.68 (1.13-2.51)	0.02
Central assessment	Grade 3 vs grade 1/2	2.47 (1.77-3.43)	<0.001	1.77 (1.18-2.67)	0.01
Ki-67 (%), semi- quantitative	Fractionally ranked	2.66 (1.88-3.75)*	<0.001	NS	
ER (%)	Fractionally ranked	0.75 (0.53-1.05)*	0.10	NS	
PR (%)	Fractionally ranked	0.53 (0.39-0.72)*	<0.001	NS	
IHC4	Fractionally ranked	2.04 (1.47-2.83)*	<0.001	NS	

\* 75th to 25th percentile

CI confidence interval, ER oestrogen receptor, HR hazard ratio, NS nonsignificant, PR progesterone receptor, RS recurrence score Significant p values are specified in bold

RS > 25 patients within all nodal subgroups. OS in low-RS patients was also significantly higher than in high-RS patients within nodal subgroups 1–3 and 4–9 involved nodes (Fig. 3b–d).

# Discussion

Our data provide the first prospective evidence in clinically high-risk EBC demonstrating strong impact of the RS on DFS and OS, not only in the collective as a whole, but also within key subgroups for clinical decision making. The independent impact of RS (as a continuous variable) on DFS was seen in a multivariate analysis including clinicopathological factors and IHC measurements. The observed survival impacts are particularly remarkable considering the PlanB trial design (CT omitted in a large fraction of HR+, pN0-1 patients with RS  $\leq$  11); similarity of DFS in RS  $\leq$  11 and RS12-25 subgroups might be attributable to mitigating effects of CT in all RS > 11 (but only some RS  $\leq$  11) patients.

Based on their excellent five-year DFS (>94%) and OS (>99%), the omission of CT in 348 clinically high-risk (up to three involved LN), genomically low-risk (RS  $\leq$  11) patients seems justified. These favourable survival rates extend those of the explorative PlanB analysis after three-year follow-up [22] and are consistent with the reported five-year invasive DFS of 93.8% and OS of 98% in the RS < 11 subgroup treated with ET alone from TAILORx (pN0 pT1-2 EBC) [21]. PlanB and TAILORx, the first prospective trials using RS for adjuvant decision making, also confirm previous retrospective data suggesting very

low relapse rates in RS < 18 pN0–N1 EBC patients without CT [11–13, 18].

MammaPrint<sup>®</sup> is the only other genomic signature supported by a prospective trial (MINDACT) [20]. MINDACT met its primary objective, five-year distant DFS > 92% without adjuvant CT in patients with high clinical risk (AdjuvantOnline! 9.0) and low genomic risk, suggesting that 46% of the clinically high-risk population (including 46% pN1 and/or 28.6% grade 3 tumours) may not require CT [20].

These prospective trials show that CT can be safely spared in a clinically meaningful fraction of HR +/HER2negative patients with 0–3 affected LN and low genomic risk. High tumour burden remained a strong unfavourable prognostic factor in PlanB; five-year DFS in pN2+ patients (a small group within PlanB) ranged from 100% (RS  $\leq$  11) to 61% (RS > 25).

Our analyses add prospective evidence to the large body of consistent, yet retrospective/observational data for Oncotype DX including the SEER (>45,000 patients) and Clalit registry (n = 2028) analyses. These studies demonstrated excellent five-year breast cancer-specific survival (BCSS) of 99–100% in RS < 18 pN0 patients (2–7% CT use) and 95–99% in RS < 18 pN1 disease (7-41% CT use, increasing from one to three positive LN) [25–27]. All studies observed poor outcome in RS > 25 patients, indicating a need for further targeted therapies in this population.

PlanB is the first prospective study comparing the prognostic value of histological grade and IHC markers (ER, PR, Ki-67, and IHC4) determined by independent central pathology with that of a genomic signature in EBC.



Fig. 3 OS for patients with RS  $\leq$  11, 12–25, and > 25 overall (a), for node-negative patients (b), for patients with pN1 disease (c), and for patients with pN2-3 disease (d)

Our univariate/multivariate analyses confirm that these markers analysed by an experienced laboratory have prognostic value, if RS is excluded. However, RS eliminated all IHC markers (and IHC4) in multivariate analysis, consistent with most [4, 5], although not all [17] retrospective studies.

Consistent with Denkert et al. [23]., HR +/HER2-negative patients with Ki-67 levels  $\geq$ 40% had poor survival, similar to that of triple-negative patients. However, an unfavourable impact of higher Ki-67 on DFS was seen only in the RS > 25 subgroup of locally HR+ patients, whereas no impact of Ki-67 was seen in corresponding RS  $\leq$  11 and RS 12–25 subgroups.

Despite Ki-67 inter-laboratory/inter-observer variability [8], cutoff uncertainty (13.25 versus 20%) [3, 5, 28], and conflicting results regarding its predictive value concerning adjuvant CT benefit [28, 29], the St. Gallen Consensus currently includes Ki-67 for identifying luminal-A-like patients, who should not receive CT. Furthermore, several studies

indicate that a substantial proportion of patients would be reclassified from luminal-A to luminal-B if genomic signatures were added to an IHC-based allocation [4, 30]. In PlanB, 10% of luminal-A-like tumours (i.e. Ki-67 < 20%) had RS > 25 [22], a high-risk group for whom CT is recommended.

CT indication based on IHC-defined luminal subtypes or grade is associated with substantial inter-observer variability, especially outside a central pathology setting [22]. It should thus be carefully re-evaluated given our prospective data demonstrating that the highly reproducible RS result outperforms IHC markers as a prognostic factor in EBC. However, the multivariate analysis suggests that genomic assays are best used in the context of established factors such as grade, nodal status, and tumour size. Ki-67, as determined by an experienced laboratory, seems to be useful for selecting patients for the more expensive genomic testing when financial resources are limited. That is because concordance between RS and Ki-67 risk



Fig. 4 DFS in HR+/Ki-67 0-10% group (a) and in HR+/Ki-67 > 10% and <40% (b)

assessment is relatively high in the Ki-67  $\leq$  10 and  $\geq$ 40% groups and the most prognostic value of RS was in the subgroup of intermediate Ki-67 (>10 to <40%). Hence, using genomic signatures may be particularly useful for intermediate/high-risk pN0-1 disease according to clinical and/or IHC markers. In clinically low-risk disease, using MammaPrint<sup>®</sup> is controversial due to the missing CT predictive effect within this group [20]. Using more than one signature per patient is currently not recommended.

Our results have some important limitations. First, this fiveyear follow-up, translational research analysis of the PlanB phase 3 trial is exploratory; the primary endpoint, comparing CT arms, will be reported subsequently. Second, so far, clinical consequences for CT omission can only be drawn for the relatively small group of RS  $\leq$  11, pN0-1 patients. Yet, approximately 60% of patients have RS 12–25, where CT benefit is uncertain [31]. Two large prospective trials (TAI-LORx, RxPONDER) randomise patients to chemo-endocrine or ET alone within this RS range. The ongoing WSG–ADAPT trial uses dynamic proliferation response to preoperative ET as a selection tool for CT allocation in pN0-1 patients with intermediate RS (data availability: 2020).

In conclusion, the present findings regarding the PlanB five-year follow-up may be helpful in guiding/refining CT decisions in HR+/HER2-negative EBC, based on genomic signature and clinicopathological factors, particularly in pN0-1 patients otherwise considered as intermediate to high risk. The PlanB results support sparing adjuvant CT in pN0-1 patients with RS  $\leq$  11 (though possibly higher "clinical") risk. However, whereas five-year data seem sufficient for assessing early risk reduction mitigated by CT [1], longer follow-up is needed to explore questions such as the duration of adjuvant ET. The additional predictive impact of RS (or other genomic signatures) for late

recurrence beyond that of clinicopathological factors is a complex and still controversial issue. Finally, additional targeted therapies (e.g. CDK 4/6 inhibitors) need to be evaluated in pN0-1 luminal-B patients with high genomic risk who have significant residual risk despite adjuvant CT.

Acknowledgements The study was supported by Genomic Health Inc, Sanofi Aventis, and Amgen. Medical editing support was provided by Avital Bareket-Samish, PhD, and funded by Genomic Health Inc. We thank all the study centres, patients, and pathologists for providing study materials. We also thank the teams of the WSG, Palleos, and central pathology lab at the MHH for their work.

Funding Genomic Health Inc, Sanofi Aventis, and Amgen.

#### Compliance with ethical standards

**Conflict of interest** HK received honoraria from Roche, Genomic Health, Novartis, and Astra-Zeneca. The West German Study group provided grant for central pathology work (to HK). RW received honoraria from Roche, Celgene, Novartis, Genomic Health, Amgen, MSD, Pfizer, and Pierre Fabre. NH received honoraria from Genomic Health, Nanostring and, Agendia. OG received honoraria from Genomic Health, Nanostring, and Roche. UN (as a representative of the WSG) received grants from Genomic Health, Sanofi Aventis, and Amgen for the conduct of the trial and received honoraria from Genomic Health, Agendia, Nanostring, Amgen, and Sanofi Aventis. CL and SK received honoraria from Genomic Health. All the other authors (TR, MJ, MC, MC) have declared that they have no conflict of interest. SS is an employee and a shareholder of Genomic Health (SS patent rights assigned to Genomic Health). CC is an employee of Genomic Health. All the other authors had nothing to disclose.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, duplication, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, a link is provided to the Creative Commons license and any changes made are indicated.

# Appendix

See Fig. 5.



**Fig. 5** DFS by central Ki-67 expression levels in centrally HR+ patients. DFS in the triple-negative (centrally ER/PR/HER2-negative) subgroup is included for comparison

# References

- Early Breast Cancer Trialists' Collaborative G (2005) Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 365(9472):1687–1717. doi:10.1016/S0140-6736(05)66544-0
- Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D (2000) Molecular portraits of human breast tumours. Nature 406(6797):747–752
- Cheang MC, Chia SK, Voduc D, Gao D, Leung S, Snider J, Watson M, Davies S, Bernard PS, Parker JS, Perou CM, Ellis MJ, Nielsen TO (2009) Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. J Natl Cancer Inst 101(10):736–750. doi:10.1093/jnci/djp082
- Prat A, Cheang MC, Martin M, Parker JS, Carrasco E, Caballero R, Tyldesley S, Gelmon K, Bernard PS, Nielsen TO, Perou CM (2013) Prognostic significance of progesterone receptor-positive tumor cells within immunohistochemically defined luminal A breast cancer. J Clin Oncol 31(2):203–209. doi:10.1200/JCO. 2012.43.4134
- Gluz O, Liedtke C, Huober J, Peyro-Saint-Paul H, Kates RE, Kreipe HH, Hartmann A, Pelz E, Erber R, Mohrmann S, Mobus V, Augustin D, Hoffmann G, Thomssen C, Janicke F, Kiechle M, Wallwiener D, Kuhn W, Nitz U, Harbeck N, Investigators EC-D (2016) Comparison of prognostic and predictive impact of genomic or central grade and immunohistochemical subtypes or IHC4 in HR+/HER2- early breast cancer: WSG-AGO EC-Doc Trial. Ann Oncol 27(6):1035–1040. doi:10.1093/annonc/mdw070

- Coates AS, Winer EP, Goldhirsch A, Gelber RD, Gnant M, Piccart-Gebhart M, Thurlimann B, Senn HJ, Panel M (2015) Tailoring therapies-improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. Ann Oncol 26(8):1533–1546. doi:10.1093/annonc/mdv221
- Varga Z, Diebold J, Dommann-Scherrer C, Frick H, Kaup D, Noske A, Obermann E, Ohlschlegel C, Padberg B, Rakozy C, Sancho Oliver S, Schobinger-Clement S, Schreiber-Facklam H, Singer G, Tapia C, Wagner U, Mastropasqua MG, Viale G, Lehr HA (2012) How reliable is Ki-67 immunohistochemistry in grade 2 breast carcinomas? A QA study of the Swiss Working Group of Breast- and Gynecopathologists. PloS one 7(5):e37379. doi:10. 1371/journal.pone.0037379
- Dowsett M, Nielsen TO, A'Hern R, Bartlett J, Coombes RC, Cuzick J, Ellis M, Henry NL, Hugh JC, Lively T, McShane L, Paik S, Penault-Llorca F, Prudkin L, Regan M, Salter J, Sotiriou C, Smith IE, Viale G, Zujewski JA, Hayes DF, International Ki-67 in Breast Cancer Working G (2011) Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. J Natl Cancer Inst 103(22):1656–1664. doi:10.1093/jnci/djr393
- Leung SCY, Nielsen TO, Zabaglo L, Arun I, Badve SS, Bane AL, Bartlett JMS et al (2016) Analytical validation of a standardized scoring protocol for Ki67: phase 3 of an international multicenter collaboration. NPJ Breast Cancer 2:16014
- Polley MY, Leung SC, McShane LM, Gao D, Hugh JC, Mastropasqua MG, Viale G, Zabaglo LA, Penault-Llorca F, Bartlett JM, Gown AM, Symmans WF, Piper T, Mehl E, Enos RA, Hayes DF, Dowsett M, Nielsen TO, International Ki67 in Breast Cancer Working Group of the Breast International G, North American Breast Cancer G (2013) An international ki67 reproducibility study. J Natl Cancer Inst 105(24):1897–1906. doi:10.1093/jnci/djt306
- Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N (2004) A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 351(27):2817–2826
- Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, Cronin M, Baehner FL, Watson D, Bryant J, Costantino JP, Geyer CE Jr, Wickerham DL, Wolmark N (2006) Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol 24(23):3726–3734. doi:10.1200/JCO.2005.04.7985
- 13. Dowsett M, Cuzick J, Wale C, Forbes J, Mallon EA, Salter J, Quinn E, Dunbier A, Baum M, Buzdar A, Howell A, Bugarini R, Baehner FL, Shak S (2010) Prediction of risk of distant recurrence using the 21-gene recurrence score in node-negative and node-positive postmenopausal patients with breast cancer treated with anastrozole or tamoxifen: a TransATAC study. J Clin Oncol 28(11):1829–1834. doi:10.1200/JCO.2009.24.4798
- 14. Gnant M, Filipits M, Greil R, Stoeger H, Rudas M, Bago-Horvath Z, Mlineritsch B, Kwasny W, Knauer M, Singer C, Jakesz R, Dubsky P, Fitzal F, Bartsch R, Steger G, Balic M, Ressler S, Cowens JW, Storhoff J, Ferree S, Schaper C, Liu S, Fesl C, Nielsen TO, Austrian B, Colorectal Cancer Study G (2014) Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the PAM50 Risk of Recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone. Ann Oncol 25(2):339–345. doi:10.1093/annonc/mdt494
- Filipits M, Rudas M, Jakesz R, Dubsky P, Fitzal F, Singer CF, Dietze O, Greil R, Jelen A, Sevelda P, Freibauer C, Muller V, Janicke F, Schmidt M, Kolbl H, Rody A, Kaufmann M, Schroth W, Brauch H, Schwab M, Fritz P, Weber KE, Feder IS, Hennig

G, Kronenwett R, Gehrmann M, Gnant M, Investigators EP (2011) A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. Clin Cancer Res 17(18):6012–6020. doi:10.1158/1078-0432.CCR-11-0926

- 16. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, Parrish M, Atsma D, Witteveen A, Glas A, Delahaye L, van der Velde T, Bartelink H, Rodenhuis S, Rutgers ET, Friend SH, Bernards R (2002) A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 347(25):1999–2009
- 17. Cuzick J, Dowsett M, Pineda S, Wale C, Salter J, Quinn E, Zabaglo L, Mallon E, Green AR, Ellis IO, Howell A, Buzdar AU, Forbes JF (2011) Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the Genomic Health recurrence score in early breast cancer. J Clin Oncol 29(32):4273–4278. doi:10.1200/JCO.2010. 31.2835
- 18. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, Ravdin P, Bugarini R, Baehner FL, Davidson NE, Sledge GW, Winer EP, Hudis C, Ingle JN, Perez EA, Pritchard KI, Shepherd L, Gralow JR, Yoshizawa C, Allred DC, Osborne CK, Hayes DF (2010) Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with nodepositive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol 11(1):55–65. doi:10.1016/S1470-2045(09)70314-6
- Markopoulos C, van de Velde C, Zarca D, Ozmen V, Masetti R (2016) Clinical evidence supporting genomic tests in early breast cancer: Do all genomic tests provide the same information? Eur J Surg Oncol. doi:10.1016/j.ejso.2016.08.012
- 20. Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delaloge S, Pierga JY, Brain E, Causeret S, DeLorenzi M, Glas AM, Golfinopoulos V, Goulioti T, Knox S, Matos E, Meulemans B, Neijenhuis PA, Nitz U, Passalacqua R, Ravdin P, Rubio IT, Saghatchian M, Smilde TJ, Sotiriou C, Stork L, Straehle C, Thomas G, Thompson AM, van der Hoeven JM, Vuylsteke P, Bernards R, Tryfonidis K, Rutgers E, Piccart M, Investigators M (2016) 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. N Engl J Med 375(8):717–729. doi:10.1056/NEJMoa1602253
- 21. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, Geyer CE Jr, Dees EC, Perez EA, Olson JA Jr, Zujewski J, Lively T, Badve SS, Saphner TJ, Wagner LI, Whelan TJ, Ellis MJ, Paik S, Wood WC, Ravdin P, Keane MM, Gomez Moreno HL, Reddy PS, Goggins TF, Mayer IA, Brufsky AM, Toppmeyer DL, Kaklamani VG, Atkins JN, Berenberg JL, Sledge GW (2015) Prospective Validation of a 21-Gene Expression Assay in Breast Cancer. N Engl J Med 373(21):2005–2014. doi:10.1056/NEJMoa1510764
- 22. Gluz O, Nitz UA, Christgen M, Kates RE, Shak S, Clemens M, Kraemer S, Aktas B, Kuemmel S, Reimer T, Kusche M, Heyl V, Lorenz-Salehi F, Just M, Hofmann D, Degenhardt T, Liedtke C, Svedman C, Wuerstlein R, Kreipe HH, Harbeck N (2016) West German Study Group phase III planB Trial: first prospective outcome data for the 21-gene recurrence score assay and concordance of prognostic markers by central and local pathology assessment. J Clin Oncol. doi:10.1200/JCO.2015.63.5383
- 23. Denkert C, Loibl S, Muller BM, Eidtmann H, Schmitt WD, Eiermann W, Gerber B, Tesch H, Hilfrich J, Huober J, Fehm T,

Barinoff J, Jackisch C, Prinzler J, Rudiger T, Erbstosser E, Blohmer JU, Budczies J, Mehta KM, von Minckwitz G (2013) Ki67 levels as predictive and prognostic parameters in pretherapeutic breast cancer core biopsies: a translational investigation in the neoadjuvant GeparTrio trial. Ann Oncol 24(11):2786–2793. doi:10.1093/annonc/mdt350

- McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM, for the Statistics Subcommittee of the NCIEWGoCD (2005) Reporting recommendations for tumor marker prognostic studies (REMARK). J Natl Cancer Inst 97(16):1180–1184. doi:10.1093/jnci/dji237
- Petkov V, Miller DP, Howlader N (2016) Breast-cancer-specific mortality in patients treated based on the 21-gene assay: a SEER population-based study. NPJ Breast Cancer 2:16017
- 26. Stemmer SM, Steiner M, Rizel S, Soussan-Gutman L, Geffen DB, Nisenbaum B, Ben-Baruch N, Isaacs K, Fried G, Rosengarten O, Uziely B, Svedman C, Rothney M, Klang S, Ryvo L, Kaufman B, Evron E, Zidan J, Shak S, N. L (2015) Real-life analysis evaluating 1594 N0/Nmic breast cancer patients for whom treatment decisions incorporated the 21-gene recurrence score result: 5-year KM estimate for breast cancer specific survival with recurrence score results ≤30 is >98%. Paper presented at the San Antonio Breast Cancer Symposium (SABCS), San Antonio, TX, Dec 8-12, 2015
- 27. Roberts MC, Miller DP, Shak S, Petkov VI (2017) Breast cancerspecific survival in patients with lymph node-positive hormone receptor-positive invasive breast cancer and Oncotype DX Recurrence Score results in the SEER database. Breast Cancer Res Treat. doi:10.1007/s10549-017-4162-3
- Nielsen TO, Jensen MB, Burugu S, Gao D, Tykjaer Jorgensen CL, Balslev E, Ejlertsen B (2016) High risk premenopausal Luminal A breast cancer patients derive no benefit from adjuvant cyclophosphamide-based chemotherapy: results from the DBCG77B clinical trial. Clin Cancer Res. doi:10.1158/1078-0432.CCR-16-1278
- 29. Viale G, Regan MM, Mastropasqua MG, Maffini F, Maiorano E, Colleoni M, Price KN, Golouh R, Perin T, Brown RW, Kovacs A, Pillay K, Ohlschlegel C, Gusterson BA, Castiglione-Gertsch M, Gelber RD, Goldhirsch A, Coates AS, International Breast Cancer Study G (2008) Predictive value of tumor Ki-67 expression in two randomized trials of adjuvant chemoendocrine therapy for node-negative breast cancer. J Natl Cancer Inst 100(3):207–212. doi:10.1093/jnci/djm289
- 30. Wuerstlein R, Sotlar K, Gluz O, Otremba B, von Schumann R, Witzel I, Schindlbeck C, Janni W, Schem C, Bauerfeind I, Hasmueller S, Tesch H, Paulenz A, Ghali N, Orujov E, Kates RE, Cowens W, Hornberger J, Pelz E, Harbeck N (2016) The West German Study Group Breast Cancer Intrinsic Subtype study: a prospective multicenter decision impact study utilizing the Prosigna assay for adjuvant treatment decision-making in estrogenreceptor-positive, HER2-negative early-stage breast cancer. Current medical research and opinion. doi:10.1185/03007995. 2016.1166102
- 31. Penault-Llorca FM, Filleron T, Asselain B, Baehner FL, Fumoleau P, Lacroix-Triki M, Butler SM, Jamshidian F, Cherbavaz DB, Shak S, Roca L, Sagan C, Lemonnier J, Martin A-L, Roche HR (2014) Prediction of recurrence with the Oncotype DX recurrence score in node-positive, HR-positive, breast cancer patients treated with adjuvant chemotherapy: results from PACS01 trial. J Clin Oncol 32:11052