

# Breast cancer subtyping by immunohistochemistry and histological grade outperforms breast cancer intrinsic subtypes in predicting neoadjuvant chemotherapy response

E. H. Lips · L. Mulder · J. J. de Ronde ·  
I. A. M. Mandjes · B. B. Koolen · L. F. A. Wessels ·  
S. Rodenhuis · J. Wesseling

Received: 22 March 2013 / Accepted: 21 June 2013 / Published online: 5 July 2013  
© The Author(s) 2013. This article is published with open access at Springerlink.com

**Abstract** Intrinsic subtypes are widely accepted for the classification of breast cancer. Lacking gene expression data, surrogate classifications based on immunohistochemistry (IHC) have been proposed. A recent St. Gallen consensus meeting recommends to use this “surrogate intrinsic subtypes” for predicting adjuvant chemotherapy resistance, implying that “Surrogate Luminal A” breast cancers should only receive endocrine therapy. In this study we assessed both gene expression based intrinsic subtypes as well as surrogate intrinsic subtypes regarding their power to predict neoadjuvant chemotherapy benefit. Single institution data of 560 breast cancer patients were reviewed. Gene expression data was available for 247 patients. Subtypes were determined on the basis of IHC, Ki67, histological grade, endocrine responsiveness, and gene expression, and were correlated with chemotherapy response and recurrence-free survival. In ER+/HER2– tumors, a high histological grade

was the best predictor for chemotherapy benefit, both in terms of pCR ( $p = 0.004$ ) and recurrence-free survival ( $p = 0.002$ ). The gene expression based and surrogate intrinsic subtype based on Ki67 had no predictive or prognostic value in ER+/HER2– tumors. Histological grade, ER, PR, and HER2 were the best predictive factors for chemotherapy response in breast cancer. We propose to continue the conventional use of these markers.

**Keywords** Breast cancer · Chemotherapy response · Gene expression data · Intrinsic subtypes · Neoadjuvant chemotherapy

## Introduction

Breast cancer can be divided into different subtypes, based on immunohistochemical (IHC) marker expression or on gene expression array data. IHC-markers are routinely used in most diagnostic labs. Hormone receptor status is used for the assignment of endocrine therapy, whereas HER2

**Electronic supplementary material** The online version of this article (doi:10.1007/s10549-013-2620-0) contains supplementary material, which is available to authorized users.

E. H. Lips · L. Mulder · J. J. de Ronde · J. Wesseling  
Department of Molecular Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

E. H. Lips · L. Mulder · J. Wesseling (✉)  
Department of Pathology, The Netherlands Cancer Institute,  
Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands  
e-mail: j.wesseling@nki.nl

J. J. de Ronde · L. F. A. Wessels  
Department of Bioinformatics, The Netherlands Cancer Institute,  
Amsterdam, The Netherlands

I. A. M. Mandjes  
Data Center, The Netherlands Cancer Institute, Amsterdam,  
The Netherlands

B. B. Koolen  
Department of Nuclear Medicine, The Netherlands Cancer Institute, Amsterdam, The Netherlands

B. B. Koolen  
Department of Surgical Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

S. Rodenhuis  
Department of Medical Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

overexpression identifies tumors sensitive to trastuzumab. The intrinsic subtypes based on gene expression analysis, first defined by Sorlie et al. [1] in 2001, have also gained wide spread acceptance as a way of classifying breast cancer tumors. Since gene expression data is often not available, approximations of intrinsic subtypes based on routinely available tests have been proposed. Recently, the expert panel of the St. Gallen meeting [2] proposed to classify tumors for therapeutic purposes based on such “surrogate intrinsic subtypes”. Luminal A disease (characterized by a positive hormone receptor, absence of HER2 amplification and “low” Ki67 [ $<14\%$ ]) is stated to require only endocrine therapy, whereas in Luminal B/HER2–, Luminal B/HER2+, HER2-enriched and triple negative disease, chemotherapy should be considered. In case of HER2 amplification, trastuzumab should be considered as well. For the ER+/HER2– tumors, the new surrogate intrinsic subtype classification replaces the concept of endocrine responsiveness, which was proposed at the St. Gallen 2007 meeting [3] to separate the ER+/HER2– tumors in a chemotherapy-sensitive and a chemotherapy-resistant subgroup.

Recently, a large analysis of the German breast group has focused on these surrogate intrinsic subtypes and chemotherapy responses in an attempt to establish the association between pCR and survival for each of the different subtypes [4]. This analysis suggested that it is important to divide a breast cancer dataset according to subtypes, as the relation between pCR and survival depended strongly on the subtype. pCR was a suitable surrogate end point for patients with luminal B/HER2–, HER2+, and triple-negative disease but not for tumors with luminal B/HER2+ or luminal A tumors. Also, another study applied the surrogate definitions to define different prognostic breast cancer groups [5]. However, no gene expression data were available in both studies, so the association between the original gene expression based classification, the surrogate intrinsic subtypes and chemotherapy response could not be established.

In this study, we compared the power to predict neoadjuvant chemotherapy response of gene expression based intrinsic subtypes, as well as surrogate intrinsic subtype definitions. Therefore, surrogate intrinsic subtypes were determined in 560 patients and response and survival analysis were performed. PAM50 gene expression based intrinsic subtypes were available for 247 tumors.

## Patients and methods

### Patients

Pre-treatment biopsies of 560 primary breast tumors were collected. All patients had received neoadjuvant treatment

at the Netherlands Cancer Institute between 2004 and 2012 as part of ongoing clinical trials, or were treated off protocol according to the standard arms of one of these studies. The studies had been approved by the ethical committee and informed consent was obtained from all patients. All tumors were either at least 3 cm in size, or the presence of—axillary lymph node metastases had been proven by fine needle aspiration (FNA). Biopsies were taken using a 14 G core needle under ultrasound guidance. After collection, specimens were snap-frozen in liquid nitrogen and stored at  $-80\text{ }^{\circ}\text{C}$ . Each patient had two or three biopsies taken to ensure that enough tumor material was available for both diagnostics and research purposes.

Depending on the particular study, a treatment regimen was assigned to each patient, which, for HER2-negative tumors, consisted of one of the following: (1) six courses of dose-dense Doxorubicin/Cyclophosphamide (ddAC); (2) six courses of Capecitabine/Docetaxel (CD) or (3) if the therapy response was considered “unfavorable” by MRI evaluation [6] after three courses, ddAC was changed to CD or vice versa. Study results have not been published yet. However, a first analysis shows that switching of chemotherapy regimen may be beneficial for initial non-responding tumors (manuscript in preparation, de Rigger et al.). pCR rates are not significantly different for patients treated with AC or DC. All HER2+ patients were treated by a regimen of three cycles of 8-weekly courses of paclitaxel, trastuzumab, and carboplatin. Hormone receptor positive patients received both chemotherapy and endocrine therapy.

### Response evaluation

The pathological response in the resection specimen after chemotherapy was used as an endpoint. Patients with a complete absence of invasive tumor cells (irrespective of carcinoma in situ) in the surgical specimen of the breast and of the lymph nodes were considered to have a pathological complete remission (pCR). All other patients (partial and non responses) were included in the “no pCR” category.

### Pathology

All tissue sections were reviewed by a consultant breast cancer pathologist (J.W.) for histological classification and immunohistochemical assessment. Samples were scored as positive for ER and/or PR by immunohistochemistry (IHC), when at least 10 % of the tumor cells nuclei showed staining of the ER or PR, respectively. A sample was scored as being HER2 positive when either a strong membrane staining (3+) could be observed by IHC or if CISH revealed amplification of HER2 in samples with

moderate (2+) membrane staining at IHC. Ki67 staining was performed with the MIB1 antibody (Dako, Glostrup, Denmark), dilution 1:250. In another study by our group (manuscript in preparation Rigter et al.) we assessed the best Ki67 cut-off to differentiate Luminal A and Luminal B tumors in our tumor set (Supplementary Fig. 1). We found that the 95 % confidence interval ranged from 10 to 15 %. 14 % was then taken as a cut-off between low and high proliferation index, as this cut-off was also described by Cheang et al. [7]. For surrogate intrinsic subtyping based on grade, grade 1 and 2 were combined into the low proliferative group and grade 3 represented the high proliferative tumors.

For ER+/HER2– tumors, we determined the endocrine responsiveness, as was described by Colleoni et al. [8]. Tumors were classified as highly endocrine responsive when ER and PR were positive in at least 50 % of the cells and incompletely endocrine responsive when either ER or PR was positive in less than 50 % of the cells. Table 1 summarizes all different subtype definitions used in this study.

#### Gene expression data

mRNA isolation and extraction from the frozen material were performed, as described previously [9]. In short, a 5- $\mu$ m section halfway through the biopsy was stained for hematoxylin and eosin and analyzed by a pathologist for tumor percentage. Only samples that contained at least 60 % tumor cells were subsequently analyzed on a microarray. All samples were labeled and hybridized to Illumina 6v3 arrays (Illumina, La Jolla, CA), according to the manufacturer's protocol. Data were log<sub>2</sub> transformed and between-array normalized using simple scaling. The subtype single sample predictor “PAM50” [10] was used to assign a molecular subtype to the samples based on their expression profiles across the intrinsic gene set. In short, the intrinsic genes were selected from the Illumina 6v3 platform, when a single gene was represented by multiple probes the probe with the highest variance was chosen. Subsequently, for all samples the Spearman correlation of a sample (i.e., the expression of the intrinsic genes of that sample) to the centroid of each corresponding molecular subtype was calculated. Each sample was then assigned to the subtype with the highest correlation.

#### Statistics

The Fisher's exact test was used to assess associations between the dichotomized response measures, pathological and molecular markers. Survival curves were constructed using the Kaplan–Meier technique, and survival was compared using log rank tests. Multivariate analysis for

chemotherapy response (pCR) and cox-regression models for recurrence free survival were performed to correct for age, T- and N-stage. The contribution of the intrinsic subtypes was assessed by the change in log likelihood ratio  $\chi^2$ , between a model containing only clinical variables (age, T-stage, N-stage), and a model including clinical variables and (surrogate) subtype information. To assess concordance between PAM50 and the surrogate intrinsic subtypes cross-tables were presented and kappa values were calculated. All analyses were performed in SPSS 17.0.

## Results

### Patient and tumor characteristics

In this study we show data of all breast cancer patients who received neoadjuvant chemotherapy in a single institution. Since the start of the neoadjuvant program in 2004, 560 patients completed neoadjuvant treatment and had surgery. From a subset of 247 patients, gene expression data was available and PAM50 intrinsic subtypes were determined. Table 1 shows the different definitions of breast cancer subtyping as well as the number of cases per group. We subsequently show the following subtype classifications: IHC-based classification, PAM50 gene expression subtype, surrogate intrinsic subtype based on Ki67, surrogate intrinsic subtype based on grade, and the endocrine responsiveness. Patient and tumor characteristics are presented in Table 2. Note that for the TN and HER2+ samples, the surrogate intrinsic subtypes completely overlap with the IHC-based classification. The concept of endocrine responsiveness was only developed for the ER+/HER2– tumors, so it is only shown for this group.

### Response for the surrogate intrinsic subtypes

We assessed response to neoadjuvant chemotherapy for the different subtype classifications. pCR rates are shown in Table 3 for the IHC-based subtyping, PAM50 subtyping, the surrogate intrinsic subtype definitions, based on both Ki67 staining and histological grade, and endocrine responsiveness. For the TN and HER2+ tumors, the surrogate intrinsic subtypes are identical to the clinical IHC-based subtypes, so the response rates for surrogate subtypes are only shown for the ER+/HER2– tumors. In this subgroup the various classifications are meant to select a chemotherapy-sensitive and a chemotherapy-resistant group. Therefore, we focus in the rest of this study mainly on the ER+/HER2– group. The surrogate intrinsic subtype based on grade could divide this subgroup in a Luminal B/HER2– subgroup with limited

**Table 1** Subtype definitions and numbers per subtype in this study

	Definition	Comments	Count	%
Immunohistochemical				
TN	ER–, PR–, HER2–		152	27.1
HER2+/ER–	HER2+, ER–, PR–		56	10.0
HER2+/ER+	HER2+, ER+, PR±		65	11.6
ER+/HER2–	ER+ HER2–, PR±		287	51.3
PAM50 <sup>a</sup>				
Basal	Based on		48	19.4
HER2-enriched	expression of 50 genes		43	17.4
Luminal A			81	32.8
Luminal B			49	19.8
Normal			26	10.5
Surrogate intrinsic (based on Ki67)				
Basal	ER–, PR–, HER2–	<sup>c</sup>	152	27.1
HER2-enriched	HER2+, ER–, PR–	<sup>c</sup>	56	10.0
Luminal A	ER+, HER2–, Ki67 < 14 %		179	32.0
Luminal B/HER2–	ER+, HER2–, Ki67 ≥ 14 %		73	13.0
Luminal B/HER2+	HER2+, ER+	<sup>c</sup>	67	12.0
ND			33	5.9
Surrogate intrinsic (based on grade)				
Basal	ER–, PR–, HER2–	<sup>c</sup>	152	27.1
HER2-enriched	HER2+, ER–, PR–	<sup>c</sup>	56	10.0
Luminal A	ER+, HER2–, grade 1/2		213	38.0
Luminal B/HER2–	ER+, HER2–, grade 3		48	8.6
Luminal B/HER2+	HER2+, ER+	<sup>c</sup>	67	12.0
ND			24	4.3
Endocrine responsiveness <sup>b</sup>				
Incompletely endocrine sensitive	ER % or PR % < 50 %		142	49.5
Highly endocrine sensitive	ER % and PR % ≥ 50 %		144	50.2
ND			1	0.3

<sup>a</sup> The PAM50 classification was determined for 247 tumors for which we had gene expression data

<sup>b</sup> The endocrine responsiveness is only shown for the ER+/HER2– tumors

<sup>c</sup> These subtypes are the same as the immunohistochemical subtypes

response to chemotherapy (14.6 % pCR), and the Luminal A group with almost no pCRs (2.3 % pCR) ( $p = 0.002$ ). The PAM50 gene expression based subtype, the surrogate intrinsic subtype based on Ki67 and endocrine responsiveness did not differentiate ER+/HER2– tumors in groups with significantly different chemotherapy responses.

For the ER+/HER2– tumors, we computed multivariate models to assess if the subtypes have predictive power in assessing chemotherapy response. We tested four different models, one employing PAM50, one employing Ki67, one employing grade, and one employing endocrine responsiveness, as these variables are used to define a Luminal A and Luminal B tumor group. Age, T-stage, and N-stage were included as covariates. In the model including grade as a variable, tumors with grade 3 tumors had a higher probability of achieving a pCR than grade 1 or 2 tumors (OR = 6.38,  $p = 0.004$ ). PAM50, Ki67 and endocrine responsiveness were not associated with pCR rates in a multivariate model (Table 4). By computing the change in log likelihood  $\chi^2$  between a model with only clinical covariates (age, T-stage, and N-stage) and a model including both clinical covariates and the subtype, we can compare the fit of the models. Higher values indicate more added predictive information. The model, including grade, adds the most information in predicting pCR rates.

### Survival analysis

To assess if the surrogate intrinsic subtypes can discriminate tumors with different survival, Kaplan–Meier curves for recurrence free survival (Fig. 1) were computed, for IHC, PAM50, surrogate subtypes (one based on Ki67 and one based on grade), and for endocrine responsiveness. For the surrogate definitions and for endocrine responsiveness, we only show the curves for the ER+/HER2–, as the triple negative and HER2+ subtypes are similar in the surrogate subtype definitions to the IHC-subtypes. Indeed, the different subtypes differentiated the tumors in classes with distinct relapse free survival (Fig. 1a–e). As expected, the triple negative or basal groups had the worst survival. Luminal A tumors had the best survival. From the surrogate definitions, the definition based on grade had the lowest  $p$  value, indicating that the grade was most strongly associated with survival. We should note that our data are preliminary, as median follow-up time was only 29 months (range 2–82), with 169 patients with a follow-up of less than 2 years. This is very limited, especially for the Luminal subtypes, which have mostly late recurrences between 5 and 10 years after treatment. We performed Cox regression analysis to determine if the surrogate definitions are still predicting survival when we correct for age, T-stage, and N-stage. Table 5 shows that the definition based on grade resulted in the best predictor, with the lowest  $p$  value and the highest hazard ratio. Patients with a Luminal B/HER2– surrogate intrinsic subtype (based on grade) have a five times higher chance on a recurrence than patients with a Luminal A subtype tumor within the first years after surgery (HR = 5.3,  $p = 0.002$ ). Also, PAM50 based subtypes and endocrine responsiveness

**Table 2** Patient and tumor characteristics

	<i>n</i>	%
Age		
<50 years	328	58.6
≥50 years	232	41.4
T-stage		
T1/2	364	65.0
T3/4	194	34.6
ND	2	0.4
N-stage		
Negative	152	27.1
Positive	404	72.1
ND	4	0.7
ER		
Negative	208	37.1
Positive	352	62.9
PR		
Negative	310	55.4
Positive	239	42.7
ND	11	2.0
HER2		
Negative	437	78.0
Positive	123	22.0
Grade		
1/2	327	58.4
3	195	34.8
ND	38	6.8
Ki67		
<14 %	191	34.1
≥14 %	148	26.4
ND	221	39.5
Chemotherapy		
6 × AC	310	55.4
6 × DC	22	3.9
3 × AC, 3 × DC	119	21.3
Trastuzumab based	109	19.5
Local Recurrence		
No recurrence	447	79.8
Recurrence	54	9.6
ND	59	10.5
Response		
No pCR	440	78.6
pCR	120	21.4

ND not determined, AC doxorubicin/cyclophosphamide, DC docetaxel/capecitabine, pCR pathological complete remission

are significantly associated with recurrence free survival (HR = 9.65,  $p = 0.02$  for PAM50 and HR = 4.81,  $p = 0.01$  for endocrine responsiveness, respectively). When we consider the change in log likelihood ratio  $\chi^2$ , both grade

and endocrine responsiveness add the most predictive information, followed by PAM50 based subtypes.

## Discussion

In this study, we collected data from 560 patients in a single cancer center who took part in an ongoing neoadjuvant chemotherapy program. For 247 patients gene expression data were available. We investigated if different intrinsic subtype definitions could help to predict neoadjuvant chemotherapy benefit in breast cancer. Subsequently we compared IHC-based subtypes, gene expression based PAM50 subtypes, surrogate intrinsic subtypes both based on Ki67 and on grade and endocrine responsiveness. The surrogate intrinsic subtyping was identical to the IHC subtypes for the basal and the HER2-positive tumors. Clinically, the surrogate intrinsic subtypes have been suggested to be mainly important to divide the clinical ER+/HER2– tumors in classes with different prognoses and chemotherapy responses. In our series, histological grade had the best predictive power. We therefore prefer to continue the use of the conventional IHC-subtyping based on hormone receptors, HER2 and histological grade rather than the “surrogate intrinsic subtype” classification, at least in the setting of neoadjuvant chemotherapy.

Our results are in line with a recent study by Von Minckwitz et al. [4], as we both show that subtypes based on grade can differentiate tumors with different chemotherapy sensitivity. However, von Minckwitz et al. did not compare the different subtype definitions. By comparing the association between the different subtype classifications and chemotherapy response, we see that grade is better than Ki67 or endocrine responsiveness in identifying groups of tumors with different chemotherapy response in this dataset. A recent review by the Breast International Group and North American Breast Cancer Group [11] recommends to use surrogate intrinsic subtypes based on Ki67, but acknowledge that the definition of low and high proliferation might be subject to change. Our study contradicts this recommendation, as we see that grade has more predictive power than Ki67. We should note that the number of responding ER+/HER2– tumors is small in our dataset, and a definite conclusion should therefore be based on larger datasets.

Breast cancer intrinsic subtypes have been subjected to many studies and a variety of classifications exist [1, 10, 12–15]. Different classifications are said to lead to the same molecular classification of breast cancer in the five subtypes, however, a recent study by Weigelt et al. [16] contradicts this assumption. In that study the agreement between three different types of gene expression based intrinsic subtyping methods was assessed. Only the basal



**Table 3** Response according to different subtype classifications

	%	Ratio	<i>p</i> value <sup>a</sup>
Immunohistochemical ( <i>n</i> = 560)			
TN	36.2	55/152	
HER2+/ER–	55.4	31/56	
HER2+/ER+	30.8	20/65	
ER+/HER2–	4.9	14/287	
PAM50 ( <i>n</i> = 247)			
Basal	41.7	20/48	
HER2-enriched	48.8	21/43	
Luminal A	4.9	4/81	
Luminal B	8.2	4/49	0.47
Normal	19.2	5/26	
Surrogate intrinsic (Ki67) ( <i>n</i> = 252)			
Luminal A	3.9	7/179	
Luminal B/HER2–	5.5	4/73	0.74
Surrogate intrinsic (grade) ( <i>n</i> = 261)			
Luminal A	2.3	5/213	
Luminal B/HER2–	14.6	7/48	<b>0.002</b>
Endocrine responsiveness ( <i>n</i> = 286)			
Highly endocrine sensitive	3.5	5/142	
Incompletely endocrine sensitive	6.3	9/144	0.41

<sup>a</sup> *p* values are shown for the comparison between Luminal A versus Luminal B tumors or highly versus incompletely endocrine sensitive tumors

subtype showed a good agreement between classifiers, whereas for the other subtypes the agreement was fair to moderate. Consequently, the authors conclude that it is too early to use the intrinsic subtypes into clinical practice. We determined the concordance between the PAM50 based classification, immunohistochemical subtypes, surrogate intrinsic subtypes, and endocrine sensitivity (Supplementary Table 1). In line with Weigelt's findings, the concordance for the basal likes was good, but modest for the other subtypes.

This study has several limitations. First, there were missing values for Ki67 staining and for grade. We used both to differentiate between Luminal A and Luminal B/HER2– tumors [2]. As seen in Supplemental Fig. 2, there is a correlation between Ki67 staining and grade, but it is far from perfect. Both Ki67 staining and histological grade determination have their pros and cons. Cheang et al. determined a cut-off percentage of positive nuclei to classify a sample as either Luminal A or Luminal B, and found that the false positive and false negative rates were both 25 % [7]. Also, the cut-off of Cheang et al. is based on patients treated in the adjuvant setting, which is different from the neoadjuvant setting. It has been recommended that each laboratory establishes its own optimal value. In our dataset, an optimal cut-off is between 10 and 15 %, we therefore kept to use the 14 % cut-off as was assessed by

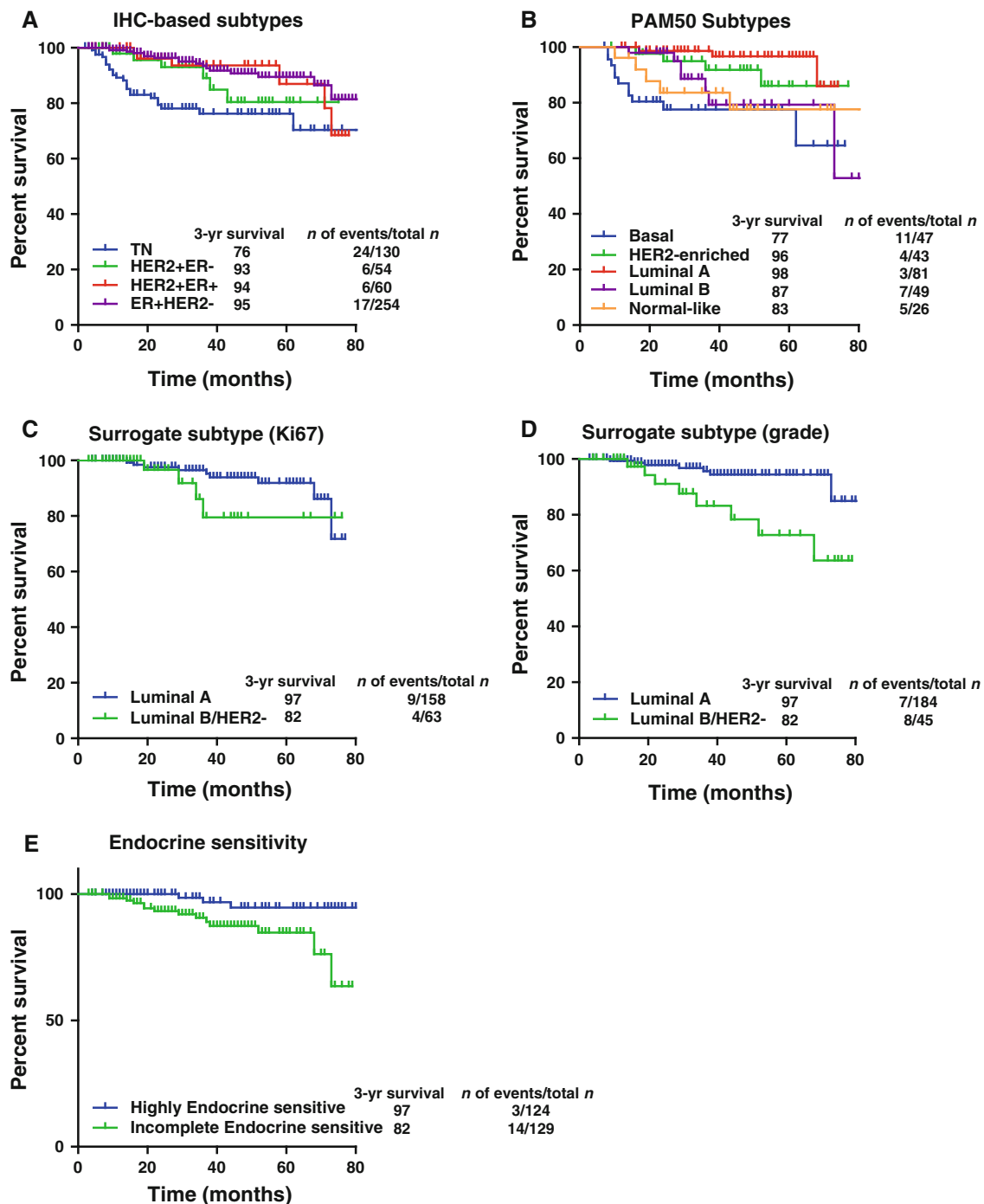
**Table 4** Multivariate analysis of pCR in ER+/HER2– tumors

	OR	95 % CI	<i>p</i> value	LR- $\chi^2$
Model based on PAM50				
Age ( $\geq 50$ vs. $< 50$ years)	0.62	0.06–6.32	0.68	
T-stage (T3/4 vs. $< T1/2$ )	0.83	0.08–8.93	0.88	
N-stage (N+ vs. N–)	0.67	0.07–6.75	0.73	
PAM50 (Luminal B vs. Luminal A)	1.74	0.22–13.64	0.60	0.28
Model based on Ki67 %				
Age ( $\geq 50$ vs. $< 50$ years)	0.12	0.01–0.95	<b>0.04</b>	
T-stage (T3/4 vs. $< T1/2$ )	0.65	0.28–1.51	0.32	
N-stage (N+ vs. N–)	1.07	0.52–2.2	0.86	
Ki67 % ( $\geq 14$ % vs. $< 14$ %)	1.12	0.31–4.05	0.87	–0.32
Model based on grade				
Age ( $\geq 50$ vs. $< 50$ years)	0.12	0.01–0.95	<b>0.04</b>	
T-stage (T3/4 vs. $< T1/2$ )	0.75	0.3–1.85	0.53	
N-stage (N+ vs. N–)	0.94	0.53–1.66	0.82	
Grade (3 vs. 1/2)	6.36	1.82–22.3	<b>0.004</b>	<b>9.08</b>
Model based on endocrine responsiveness				
Age ( $\geq 50$ vs. $< 50$ years)	0.19	0.04–0.9	<b>0.04</b>	
T-stage (T3/4 vs. $< T1/2$ )	0.70	0.33–1.48	0.35	
N-stage (N+ vs. N–)	0.92	0.62–1.36	0.67	
Endocrine responsive (incomplete vs. highly)	2.07	0.66–6.44	0.21	–0.32

Values in bold indicate significant *p* values (*p* < 0.05) and associated best fit models

LR- $\chi^2$ , change in log likelihood  $\chi^2$

Viale and co-authors [17–19]. Apart from this, the superiority of grade over Ki67 is perhaps not surprising. In addition to proliferative activity, grade also incorporates two other characteristics of the tumor: the degree of tubule formation and the degree of nuclear polymorphism, both which could correlate with chemosensitivity. A second limitation is that survival data were derived from a limited follow-up period, as the study included many patients from recent years. The median follow-up was 29 months (range 2–82). This is relatively brief for the luminal subtypes, which are subjected to late recurrences which may occur up to 15 years or later after the diagnosis of their primary tumor [20]. Strong points of this study are that we have a large, single institution dataset. For half of patients we had both gene expression and IHC data. HER2 status was checked by CISH in tumors with intermediate IHC scores. Further, all pathology data was reviewed by the same expert breast pathologist. Also, chemotherapy regimens were the same for most patients: the large majority of the HER2– tumors were treated by six courses of anthracycline based or anthracyclin-taxane-based chemotherapy and the HER2+ tumors received trastuzumab in combination with carboplatin and paclitaxel.



**Fig. 1** RFS curves for subsequently the traditional IHC subtypes (log rank  $p$  value = 0.0002) (a), the PAM50 subtypes (log rank  $p$  value = 0.003) (b), the surrogate intrinsic subtypes-Ki67 (log rank

$p$  = 0.19) (c), the surrogate intrinsic subtypes-grade (log rank  $p$  = 0.003) (d), and endocrine sensitivity (log rank  $p$  = 0.0067) (e)

In conclusion, our findings do not make a strong case for the “surrogate intrinsic subtype” terminology. For the ER+/HER2- tumors, the routine determination of IHC positivity for the estrogen receptor, HER2 in situ hybridization and grade yield the best separation of prognostic—and predictive—in terms of pCR rates—

subgroups. For the other surrogate subtypes, ER and HER2 define the surrogate subtype (by definition). As the concordance between surrogate intrinsic subtype and PAM50 subtype is far from perfect, the terminology appears needlessly confusing without adding clinically useful information.

**Table 5** Multivariate cox proportional hazard analysis of the risk of recurrence (recurrence free survival) in ER+/HER2– tumors

	HR	95 % CI	<i>p</i> -value	LR- $\chi^2$
Model based on PAM50				
Age ( $\geq 50$ vs. $< 50$ years)	1.81	0.38–8.62	0.46	
T-stage (T3/4 vs. $< T1/2$ )	1.46	0.31–6.84	0.63	
N-stage (N+ vs. N–)	1.13	0.09–14.12	0.92	
PAM50 (Luminal B vs. Luminal A)	9.65	1.36–68.4	<b>0.02</b>	<b>8.23</b>
Model based on Ki67 %				
Age ( $\geq 50$ vs. $< 50$ years)	2.94	0.86–10.06	0.09	
T-stage (T3/4 vs. $< T1/2$ )	1.23	0.48–3.19	0.66	
N-stage (N+ vs. N–)	1.40	0.36–5.47	0.63	
Ki67 % ( $\geq 14$ % vs. $< 14$ %)	3.32	0.82–13.45	0.09	0.11
Model based on grade				
Age ( $\geq 50$ vs. $< 50$ years)	2.29	0.78–6.7	0.13	
T-stage (T3/4 vs. $< T1/2$ )	2.09	0.88–4.94	0.09	
N-stage (N+ vs. N–)	1.10	0.34–3.53	0.87	
Grade (3 vs. 1/2)	5.30	1.82–15.39	<b>0.002</b>	<b>9.16</b>
Model based on endocrine responsiveness				
Age ( $\geq 50$ vs. $< 50$ years)	2.34	0.85–6.46	0.10	
T-stage (T3/4 vs. $< T1/2$ )	1.86	0.89–3.87	0.10	
N-stage (N+ vs. N–)	1.04	0.41–2.64	0.94	
Endocrine responsive (incomplete vs. highly)	4.81	1.36–17.07	<b>0.01</b>	<b>9.06</b>

Values in bold indicate significant *p* values ( $p < 0.05$ ) and associated best fit models

**Acknowledgments** This study was carried out within the framework of CTMM, the Center for Translational Molecular Medicine ([www.ctmm.nl](http://www.ctmm.nl)), Project Breast CARE (030-104).

**Conflict of interest** The authors declare no conflicts of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Sorlie T, Perou CM, Tibshirani R et al (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 98:10869–10874
- Goldhirsch A, Wood WC, Coates AS et al (2011) Strategies for subtypes—dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol* 22:1736–1747
- Goldhirsch A, Wood WC, Gelber RD et al (2007) Progress and promise: highlights of the international expert consensus on the primary therapy of early breast cancer 2007. *Ann Oncol* 18:1133–1144
- von Minckwitz G, Untch M, Blohmer JU et al (2012) Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. *J Clin Oncol* 30:1796–1804
- Brouckaert O, Laenen A, Vanderhaegen J et al (2012) Applying the 2011 St Gallen panel of prognostic markers on a large single hospital cohort of consecutively treated primary operable breast cancers. *Ann Oncol* 23(10):2578–2584
- Loo CE, Straver ME, Rodenhuis S et al (2011) Magnetic resonance imaging response monitoring of breast cancer during neoadjuvant chemotherapy: relevance of breast cancer subtype. *J Clin Oncol* 29:660–666
- Cheang MC, Chia SK, Voduc D et al (2009) Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J Natl Cancer Inst* 101:736–750
- Colleoni M, Bagnardi V, Rotmensz N et al (2009) Increasing steroid hormone receptors expression defines breast cancer subtypes non responsive to preoperative chemotherapy. *Breast Cancer Res Treat* 116:359–369
- Hannemann J, Oosterkamp HM, Bosch CA et al (2005) Changes in gene expression associated with response to neoadjuvant chemotherapy in breast cancer. *J Clin Oncol* 23:3331–3342
- Parker JS, Mullins M, Cheang MC et al (2009) Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 27:1160–1167
- Fumagalli D, Bedard PL, Nahleh Z et al (2012) A common language in neoadjuvant breast cancer clinical trials: proposals for standard definitions and endpoints. *Lancet Oncol* 13:e240–e248
- Haibe-Kains B, Desmedt C, Loi S et al (2012) A three-gene model to robustly identify breast cancer molecular subtypes. *J Natl Cancer Inst* 104:311–325
- Hu Z, Fan C, Oh DS et al (2006) The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* 7:96
- Perou CM, Sorlie T, Eisen MB et al (2000) Molecular portraits of human breast tumours. *Nature* 406:747–752
- Sorlie T, Tibshirani R, Parker J et al (2003) Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA* 100:8418–8423
- Weigelt B, Mackay A, A'hem R et al (2010) Breast cancer molecular profiling with single sample predictors: a retrospective analysis. *Lancet Oncol* 11:339–349



17. Viale G, Regan MM, Mastropasqua MG et al (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* 100:207–212
18. Goodson WH III, Moore DH, Ljung BM et al (2000) The prognostic value of proliferation indices: a study with in vivo bromodeoxyuridine and Ki-67. *Breast Cancer Res Treat* 59:113–123
19. Liu S, Edgerton SM, Moore DH et al (2001) Measures of cell turnover (proliferation and apoptosis) and their association with survival in breast cancer. *Clin Cancer Res* 7:1716–1723
20. Early Breast Cancer Trialists' Collaborative Group (EBCTCG) (2005) Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 365:1687–1717