CLINICAL TRIAL

Chemotherapy response and recurrence-free survival in neoadjuvant breast cancer depends on biomarker profiles: results from the I-SPY 1 TRIAL (CALGB 150007/150012; ACRIN 6657)

Laura J. Esserman · Donald A. Berry · Maggie C. U. Cheang · Christina Yau ·
Charles M. Perou · Lisa Carey · Angela DeMichele · Joe W. Gray · Kathleen Conway-Dorsey ·
Marc E. Lenburg · Meredith B. Buxton · Sarah E. Davis · Laura J. van't Veer ·
Clifford Hudis · Koei Chin · Denise Wolf · Helen Krontiras · Leslie Montgomery ·
Debu Tripathy · Constance Lehman · Minetta C. Liu · Olufunmilayo I. Olopade ·
Hope S. Rugo · John T. Carpenter · Chad Livasy · Lynn Dressler · David Chhieng ·
Baljit Singh · Carolyn Mies · Joseph Rabban · Yunni-Yi Chen · Dilip Giri ·
Alfred Au · Nola Hylton · The I-SPY 1 TRIAL Investigators

Abstract Neoadjuvant chemotherapy for breast cancer allows individual tumor response to be assessed depending on molecular subtype, and to judge the impact of response to therapy on recurrence-free survival (RFS). The multicenter I-SPY 1 TRIAL evaluated patients with ≥ 3 cm tumors by using early imaging and molecular signatures, with outcomes of pathologic complete response (pCR) and RFS. The current analysis was performed using data from patients who had molecular profiles and did not receive trastuzumab. The various molecular classifiers tested were highly correlated. Categorization of breast cancer by molecular signatures enhanced the ability of pCR to predict improvement in RFS compared to the population as a whole. In multivariate analysis, the molecular signatures that added to the ability of HR and HER2 receptors, clinical stage, and pCR in predicting RFS included 70-gene signature, wound healing signature, p53 mutation

L. J. Esserman (✉) · C. Yau · M. B. Buxton ·
S. E. Davis · L. J. van't Veer · K. Chin · D. Wolf ·
H. S. Rugo · J. Rabban · Y.-Y. Chen · A. Au · N. Hylton ·
The I-SPY 1 TRIAL Investigators
University of California San Francisco, 1600 Divisadero Street,
Box 1710, San Francisco, CA 94115, USA
e-mail: Laura.esserman@ucsfmedctr.org

D. A. Berry · The I-SPY 1 TRIAL Investigators
MD Anderson Cancer Center, Houston, USA

M. C. U. Cheang · C. M. Perou · L. Carey · K. Conway-Dorsey
· C. Livasy · L. Dressler · The I-SPY 1 TRIAL Investigators
University of North Carolina, Chapel Hill, USA

A. DeMichele · C. Mies · The I-SPY 1 TRIAL Investigators
University of Pennsylvania, Philadelphia, USA

J. W. Gray · The I-SPY 1 TRIAL Investigators
Oregon Health Sciences University, Portland, USA

J. W. Gray · The I-SPY 1 TRIAL Investigators
Lawrence Berkeley National Labs, Berkeley, USA

M. E. Lenburg · The I-SPY 1 TRIAL Investigators
Boston University, Boston, USA

C. Hudis · D. Giri · The I-SPY 1 TRIAL Investigators
Memorial Sloan-Kettering Cancer Center, New York, USA

H. Krontiras · J. T. Carpenter · The I-SPY 1 TRIAL
Investigators
University of Alabama, Tuscaloosa, USA

L. Montgomery · The I-SPY 1 TRIAL Investigators
Albert Einstein College of Medicine, New York, USA

D. Tripathy · The I-SPY 1 TRIAL Investigators
University of Southern California, Pasadena, USA

signature, and PAM50 risk of recurrence. The low risk signatures were associated with significantly better prognosis, and also identified additional patients with a good prognosis within the no pCR group, primarily in the hormone receptor positive, HER-2 negative subgroup. The I-SPY 1 population is enriched for tumors with a poor prognosis but is still heterogeneous in terms of rates of pCR and RFS. The ability of pCR to predict RFS is better by subset than it is for the whole group. Molecular markers improve prediction of RFS by identifying additional patients with excellent prognosis within the no pCR group.

Introduction

Molecular and genetic studies demonstrate that breast cancer is a heterogeneous disease. Several classifiers are available for distinguishing tumor types based on prognosis and prediction of response to chemotherapy and hormonal therapy [1–3]. Molecular features are associated with substantially different outcomes [4] and with wide variability in response to standard therapies [5, 6]. Symptomatic tumors that tend to be large and palpable on presentation have substantially higher risk of recurrence than tumors detected by screening [7]. For these larger tumors, neoadjuvant, or preoperative, chemotherapy makes it possible to assess response to treatment and may provide insights to the tumor's biology. Studies examining the degree to which pathologic complete response (pCR) to therapy is predictive of recurrence-free survival (RFS) or overall survival (OS) have given mixed results in relatively unselected populations [8–12].

The I-SPY 1 TRIAL (investigation of serial studies to predict your therapeutic response with imaging and molecular analysis) is a multicenter neoadjuvant study of women with histologically confirmed invasive breast cancers. This report describes associations between molecular markers assessed in pretreatment tumor biopsy samples and response to neoadjuvant chemotherapy at the time of surgery, longer-term disease outcomes, and the relationship between response and RFS.

Methods

Study design and patient selection

The I-SPY 1 TRIAL methods have been described in detail elsewhere [13, 14] and was a collaboration of the American College of Radiology Imaging Network (ACRIN), Cancer and Leukemia Group B (CALGB), and Specialized programs of research excellence (SPORE). All patients gave written consent and had histologically confirmed invasive breast cancers measuring at least 3 cm by clinical examination or imaging, with no evidence of distant metastatic disease. Patients' clinical stage 1 by exam was eligible if tumor size was >3 cm by imaging. Patients with T4 or inflammatory disease were eligible. The regimen of neoadjuvant chemotherapy included an initial anthracycline-based regimen after which patients either underwent surgery or received a taxane-based regimen prior to surgery.

Assays were conducted in nine laboratories. Data was integrated for central accession for analysis using NCICB's caINTEGRATOR application (https://caintegrator-stage.nci.nih.gov/ispy/index2.jsp)—I-SPY 1 data version dated February 2011.

Standard pathology biomarkers

Hormone and HER2 receptor expression were measured from pretreatment core biopsies. Estrogen and progesterone receptor status were determined by immunohistochemistry (IHC) and calculation of Allred scores [15] at the study sites. HER2 status was determined locally by IHC and/or fluorescence-in situ hybridization assays (FISH). HER2 testing (IHC and FISH) was also performed centrally at the University of North Carolina (UNC) [13, 16]. HER2 status was considered positive if either local or central assays were positive. Ki67 was recorded as low (<10%), medium (10–20%), or high (>20%) and described in detail in supplemental methods [17].

Evaluation of pathologic response

pCR is defined as no invasive tumor present in either breast or axillary lymph nodes. Residual Cancer Burden (RCB) [9] was assessed and included the primary tumor bed area, overall invasive cancer and in situ disease cellularity, number of positive lymph nodes and diameter of largest metastasis. I-SPY 1 TRIAL pathologists were centrally

C. Lehman · The I-SPY 1 TRIAL Investigators
University of Washington, Washington, USA

M. C. Liu · The I-SPY 1 TRIAL Investigators
Georgetown University, Washinton D.C., USA

O. I. Olopade · The I-SPY 1 TRIAL Investigators
University of Chicago, Chicago, USA

D. Chhieng · The I-SPY 1 TRIAL Investigators
Yale University, New Haven, USA

B. Singh · The I-SPY 1 TRIAL Investigators
New York University, New York, USA

trained; all cases were re-reviewed and scored for RCB as a dichotomous outcome (0, I vs. II, III) and by class (0, I, II, III). Data was recorded using NCI's Oracle Clinical Remote Data Capture version 4.5 electronic database.

RNA analysis

Tissue samples immediately frozen in OCT were assayed on catalog 44,000 feature Agilent Human oligonucleotide microarrays (catalog # G4112F). Total RNA purification and microarray hybridization were done as previously described [18]. The background was subtracted and Lowess normalized log2 ratio of Cy3 and Cy5 intensity values were calculated [19]. The primary microarray data presented in this study are available in the GEO database under accession number GSE22226.

Intrinsic subtype classification was determined by PAM50 50-gene assay as described [18]. The risk of recurrence score (ROR-S) classified patients as having high, medium, or low risk of relapse using predefined cutpoints as described previously [19].

The 70-gene prognostic profile was determined using representative probes and data normalization as previously described [20]. This profile classified patients as having high or low risk of relapse using the predefined threshold [20, 21].

Wound healing signature [22] was used to classify tumors as quiescent or activated. A gene-expression signature predictive of p53 genotype [23] was used to classify tumors as p53 wild type or mutated.

DNA analysis

DNA copy number abnormality was assessed by a molecular inversion probe (MIP) platform with focal amplification and high resolution ($\sim$10 K bp) as previously described [24–26]. Direct p53 genotyping was performed and mutations were detected by the Roche p53 AmpliChip beta test array [27, 28] and a combination of two approaches described in supplemental methods.

Statistical analysis

The primary endpoint for the trial was RFS according to the STEEP criteria [29]. Time to recurrence was computed from start of treatment; and RFS at 3 years was determined by Kaplan–Meier analysis.

Associations of signature classifications with pCR and RFS were assessed by logistic regression and proportional hazards modeling, respectively. Association of pCR with RFS within each risk subset was also tested. These analyses were conducted using JMP Version 8.0.1, SAS Institute Inc.

We took HR/HER2 categories to be standard and addressed the ability of other signatures to predict RFS and pCR assuming HR/HER2 category is given. We used multivariate Cox regression of RFS on molecular signature classification, adjusting for the predefined contribution of HR and HER2 (as derived from a Cox model fit of RFS on HR and HER2). Similarly, we employed multivariate logistic regression to evaluate whether other molecular signature classifications were independently predictive pCR when given HR/HER2. These analyses were conducted using Bioconductor R [30].

Differences in rates of pCR and rates of RCB class 0 or 1 within molecular signature-defined subgroups were assessed by χ^2 tests.

Patient were categorized as low- or high-risk by each molecular signature HR+/HER2− versus others, luminal (A and B) versus others (HER2-enriched, basal, and normal-like) [31], p53 wild type versus mutation, 70-gene prognosis signature low versus high [20], and wound-quiescent versus wound-activated [22], and stage 3 (including inflammatory) versus earlier stage to ensure comparable degrees of freedom. The pairwise concordances of the molecular signatures were compared by Kendall's rank correlation and Fisher's exact test.

All statistical analyses were performed (CY) and verified by a second statistician (DB) to confirm the results.

Results

Patient characteristics

Between May 2002 and March 2006, 237 patients were enrolled at nine institutions as described previously [13]. As shown in the CONSORT diagram (Fig. 1), 221 patients received an anthracycline as initial neoadjuvant chemotherapy, with 95% receiving a taxane, and were considered evaluable. Of these, 215 (97%) underwent surgical resection and had pathologic data available for analysis; IHC receptors were available for 210 and RCB was available for 196 (93%). We attempted microarrays on all 210 patients but could generate high quality gene-expression arrays for just 149 (Fig. 1). In addition, 171 patients had p53 gene mutation chip data and 153 had copy number variation by MIP arrays. Patients with available gene-expression arrays, who did and did not receive trastuzumab, were not significantly different from the overall cohort (Table 1).

Among the 215 patients, 20 of 67 (30%) HER2+ patients received neoadjuvant trastuzumab, as previously described [13]. Of the 46 HER2+ patients who did not, 17 (36%) received adjuvant trastuzumab. Radiation and hormonal adjuvant therapy were also given at physician discretion as clinically indicated (Table 1). Analysis of RFS was limited to patients who did not receive trastuzumab.
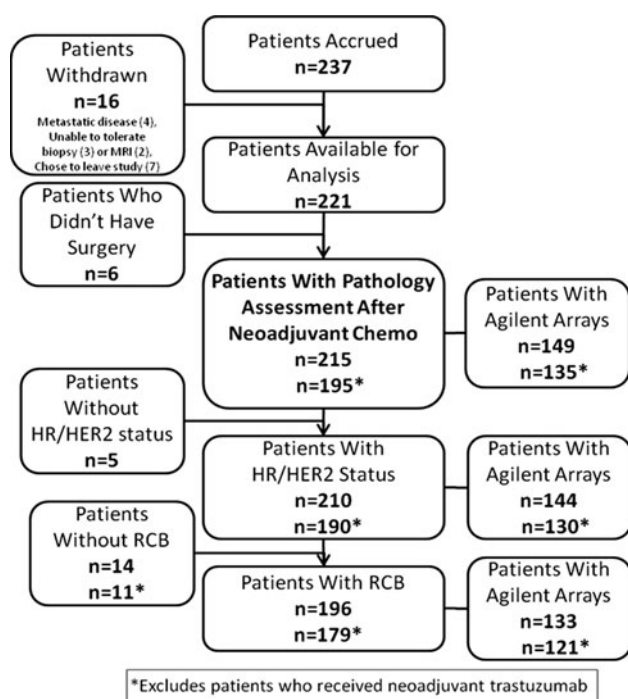
Fig. 1 CONSORT diagram: patients available for analysis. Of the 237 patients enrolled in the study, 16 patients withdrew. Of the 221 patients available for analysis, six decided not to undergo surgery after completing neoadjuvant chemotherapy, leaving 215 patients available for pathologic response analysis

Most (65%) patients had clinically or pathologically confirmed axillary lymph node involvement at diagnosis and 90% had tumors of intermediate or high histologic grade. Median follow-up for survival and RFS was 3.9 years.

Early outcomes: residual disease measured at the time of surgical resection

The overall rate of pCR was 27%, and the rate of RCB scores of 0 or I was 37%.

As shown in Table 2, the response to therapy varied considerably by marker subset. For the subsets defined by clinical assay (IHC markers), pCR rates were lowest for the HR+/HER2− subset (9%) and highest for the HR−/HER2+ subset (54%). Of the PAM50 intrinsic subtypes, pCR rates were lowest for luminal A (3%) and highest for HER2-enriched population (50%). For Ki67, the pCR rates varied from 5% for the low group to 35% for the high group. Other biomarkers were associated with high rates of pCR, including p53 null mutations by Gene Chip (47%) and amplification at 17q as measured by MIP array (45%).

The rates of pCR and RCB scores of 0 or I for four gene-expression prognostic classifiers—70-gene prognosis signature, ROR-S, wound healing, and p53 mutation signatures—were all low (Table 2). In this population of patients treated with neoadjuvant chemotherapy, a minority of patients had good prognosis profiles: 9% were classified as 70-gene low risk, 27% as ROR-S low, 25% as wound healing quiescent, and 49% as p53 wild type. The respective pCR rates were 0, 6, 7, and 9%. Rates of pCR were higher for poor prognosis signatures, including 70-gene high risk (24%), wound healing signature activated (26%), ROR-S moderate risk (17%) and high risk (36%) and p53 mutation predicted by expression profile (34%). Clinical outcomes were better when pCR or an RCB score of 0 or 1 was achieved.

Recurrence-free and OS

Three-year RFS and OS for the entire cohort were 78 and 85%, respectively. When RFS for the population was stratified by molecular signatures, outcomes by subtype differed substantially (Table 2).

We dichotomized the classifiers, using low versus medium and high for the analyses shown in Tables 3 and 4. Although developed using heterogeneous methods and populations, dichotomized tumor classifiers and signatures were highly correlated, with concordances generally >70% (Table 3). As shown in Table 4, for the breast cancer subtypes known to be associated with a better outcome (e.g., luminal, 70-gene low risk, wound healing quiescent, and p53 wild type), the RFS was relatively high for patients who did not achieve pCR, but pCR was associated with a better outcome, regardless of subtype. Poor prognosis tumors tend to be more sensitive to chemotherapy, and thus have a higher rate of pCR. The hazard ratio for the prediction of pCR versus not was better within the dichotomized molecular classifications than for the population as a whole. Associations between RFS and pCR differed the most for HR+/HER2− versus not (hazard ratio 0.17, [95% CI 0.04–0.51]), which was also where rates of pCR differed the most (9 vs. 38%). For the dichotomized breast cancer molecular classifiers (Table 4), the difference between highest and lowest rates of pCR varied from 19 to 29%.

In a multivariate model, the factors that added to HR and HER2 in predicting RFS were clinical stage (stage 3 vs. not), wound healing (activated vs. quiescent), ROR-S, and predicted p53 mutation signature (Table 5). The only factor that added to HR and HER2 in predicting pCR was Ki67 low and medium versus high (data not shown). The analysis shown was performed using the dichotomized groupings for the molecular signatures, but the results were qualitatively the same when the original groupings were used (data not shown).

When receptor status and pCR were fixed in a multivariate model, most molecular signatures and clinical stage improved the ability to predict RFS (data not shown). The most likely reason is because the molecular classifiers can identify additional patients in the "no pCR" group who have excellent outcomes, the majority of whom are in the

Table 1 Demographics and characteristics of patients in the I-SPY 1 TRIAL

Characteristics	I-SPY trial evaluable (n = 221)	Profiled with agilent microarray (n = 149)	Profiled with agilent microarray without any trastuzumab (n = 120)
Age (years)			
Median (range)	49 (26–68)	48 (27–65)	47 (28–65)
Premenopausal	48% (106)	49% (72)	47% (56)
Race			
Caucasian	75% (165)	76% (114)	77% (93)
African American	19% (42)	18% (27)	18% (21)
Asian	4% (9)	5% (7)	4% (5)
Other	2% (5)	1% (1)	1% (1)
Clinical tumor size (cm)			
Median (range)	6.0 (0–25)	5.5 (0–25)	5.5 (0–18)
Tumor longest diameter on baseline MRI (cm)			
Median (range)	6.8 (0–18.4)	6.5 (0–18.4)	6.45 (2.0–16.6)
Clinically node positive at diagnosis	65% (143)	66% (99)	63% (76)
Histologic grade (baseline)			
Low	8% (18)	7% (10)	7% (9)
Intermediate	43% (96)	42% (63)	40% (48)
High	47% (103)	50% (75)	52% (62)
Indeterminate	2% (4)	1% (1)	1% (1)
Clinical stage (baseline)			
I	1% (3)	2% (3)	2% (3)
IIA	19% (43)	21% (32)	25% (30)
IIB	28% (61)	26% (38)	26% (31)
IIIA	35% (78)	34% (51)	33% (40)
IIIB	5% (11)	5% (8)	7% (8)
IIIC	3% (7)	3% (5)	2% (2)
Inflammatory	8% (17)	7% (11)	4% (5)
Indeterminate	<1% (1)	1% (1)	1% (1)
Hormone receptors (baseline)			
ER-positive	56% (124)	55% (82)	58% (70)
PR-positive	46% (102)	44% (66)	49% (59)
HR-positive (ER or PR)	59% (130)	58% (86)	62% (74)
Her-2 positive (baseline)	30% (67)	30% (45)	15% (18)
HR-negative/Her-2 negative (baseline) (triple negative)	24% (53)	25% (37)	30% (36)
Neoadjuvant treatment			
AC only	5% (11)	3% (4)	3% (4)
AC + T	85% (187)	87% (129)	95% (114)
AC + T + trastuzumab	9% (20)	9% (14)	–
AC + T + other	1% (3)	1% (2)	1% (2)
Surgery type			
Mastectomy	56% (123)	57% (84)	54% (65)
Lumpectomy	41% (92)	40% (60)	43% (51)
No Surgery	3% (6)	3% (5)	3% (4)
Post-operative adjuvant therapy	58% (128)	56% (84)	45% (54)
Any hormonal therapy	34% (75)	34% (52)	38% (46)
Tamoxifen	43% (95)	44% (61)	43% (52)
Aromatase inhibitor	12% (27)	12% (20)	12% (15)

Table 1 continued

Characteristics	I-SPY trial evaluable (n = 221)	Profiled with agilent microarray (n = 149)	Profiled with agilent microarray without any trastuzumab (n = 120)
Ovarian suppression or ablation	3% (7)	3% (6)	4% (5)
Trastuzumab	16% (35)	16% (25)	–

Institutions participating in the I-SPY trial

Total accrual 237

Institution name	Accrual	Biomarker tests performed*
Georgetown University Hospital	4	
Memorial Sloan-Kettering Cancer Center	20	
National Cancer Institute/George Mason University	N/A	Reverse phase phosphoprotein arrays
University of Alabama at Birmingham	50	
University of California, San Francisco	67	Frozen tissue, DNA and RNA processing, cDNA arrays
University of Chicago	2	
University of North Carolina	38	Paraffin-based IHC; p53 gene chip; expression profiling
University of Pennsylvania Medical Center	36	Paraffin-based IHC
University of Texas, Southwestern	14	
University of Washington, Seattle	6	

ER estrogen receptor, *PR* progesterone receptor, *AC* anthracycline, *T* trastuzumab

* All sites performed testing for estrogen receptor (ER), progesterone (PR), and human epithelial growth factor receptor 2 (HER2). Frozen tissue was stored at the University of California, San Francisco; formalin-fixed, paraffin-embedded at the University of North Carolina; and blood at CALGB Pathology Central Office at Ohio State University

HR+/HER2− receptor subtypes. The Kaplan–Meier plots for the hormone receptor positive patients without a pCR, stratified by four of the molecular signatures, and by clinical stage, showed that the low risk classification identified patients with a much better outcome than the high risk classification (Fig. 2).

A visual representation or unsupervised cluster of the clinical features of the patients in I-SPY 1 who did not receive trastuzumab is shown in Fig. 3. Like Table 3, this figure illustrates the high degree of overlap among clinical variables and molecular signatures. There is a cluster of tumors with many high risk features illustrated by the cluster in red in the lower left, some with pCR and good outcomes and without pCR and poor outcomes, but the molecular features of these tumors are the same and do not appear to provide information to discriminate between good and poor outcomes.

Discussion

The I-SPY 1 collaboration demonstrates that standards for imaging, data and tissue collection can be followed and molecular profiling from small specimens is achievable. Molecular profiles were generated for over 65% of all patients (improving as the trial proceeded), and these patients are representative of the entire data set.

Patients who present with large breast tumors, as exemplified by the I-SPY 1 cohort, have biologically poor-risk cancers, as evidenced by 91% having 70-gene high risk profile and the fact that many are interval cancers [32]. Even within this clinically high risk population, response to therapy was heterogeneous. HER2 positivity and HR negativity were associated with a greater rate of pCR, as were four poor prognosis molecular signatures: wound-activated signature, ROR-S high risk, 70-gene poor-risk, and p53 predicted mutation. Patients with good prognosis signatures had a lower chance of short-term (pCR, RCB) response to chemotherapy, but had better long-term (RFS, OS) outcomes, even when their tumors did not respond to therapy. These findings support the emerging consensus that patients with good risk signatures (wound healing quiescent, 70-gene low, and ROR-S low) have low rates of early recurrence in spite of large tumor size. The molecular profiles vary by the percent of the population they classify as low risk, the fraction that respond to therapy, and the outcomes among those without pCR, even though the data set was not sufficiently large to show a statistical difference.

The International Breast Cancer Study Group (IBCSG), NSABP, and MD Anderson Cancer Center [33] have found that pCR rates are much higher in patients with HR-negative tumors than in those with HR-positive tumors. These

Table 2 Distribution and response rates by molecular phenotypes and profiles

	Overall distribution available for analysis, % (n)	pCR, % (n)	pCR vs. not P value	RCB 0, I, % (n)	RCB 0, I vs. II, III P value	3 year RFS Overall	pCR	NO pCR
Entire population	$n = 215$	27% (58/215)		37% (74/201)		78% (168/215)	86% (50/58)	74% (118/157)
IHC/FISH	$n = 210$	27% (56/210)		37% (72/196)				
HR+/HER2−	44% ($n = 93$)	9% (8/93)	<0.001	17% (15/87)	<0.001	87% (81/93)	100% (8/8)	86% (73/85)
HR+/HER2+ (w or w/o T)	16% ($n = 33$)	33% (11/33)		50% (15/30)		75% (25/33)	61% (7/11)	82% (18/22)
HR−/HER2+ (w or w/o T)	16% ($n = 33$)	58% (19/33)		72% (23/32)		75% (25/33)	89% (17/19)	57% (8/14)
HR−/HER2−	24% ($n = 51$)	35% (18/51)		40% (19/47)		66% (34/51)	89% (16/18)	54% (18/33)
HER2+ (w/o any T†)	14% ($n = 29$)	41% (12/29)		55% (16/29)		75% (7/29)	91% (1/12)	65% (6/17)
HR+/HER2+ (w/o any T†)	55% ($n = 16$)	31% (5/16)		44% (7/16)		81% (13/16)	80% (4/5)	82% (9/11)
HR−/HER2+ (w/o any T†)	45% ($n = 13$)	54% (7/13)		69% (9/13)		67% (9/13)	100% (7/7)	33% (2/6)
Population without any trastuzumab								
Intrinsic subtype	$n = 120$	$n = 116$		$n = 108$				
Luminal A	30% ($n = 36$)	3% (1/36)	0.001	9% (3/32)	0.001	94% (34/36)	100% (1/1)	94% (33/35)
Luminal B	22% ($n = 26$)	16% (4/25)		21% (5/24)		79% (20/25)	100% (4/4)	75% (16/21)
Her2-enriched	9% ($n = 11$)	50% (5/10)		60% (6/10)		90% (9/10)	80% (4/5)	100% (0/5)
Basal	36% ($n = 43$)	33% (14/42)		41% (16/39)		58% (25/42)	85% (12/14)	45% (13/28)
Normal-like	3% ($n = 4$)	33% (1/3)		66% (2/3)		100% (3/3)	100% (1/1)	100% (2/2)
ROR-S								
Low	27% ($n = 32$)	6% (2/32)	0.006	18% (5/28)	0.06	97% (31/32)	100% (2/2)	96% (29/30)
Medium	35% ($n = 42$)	17% (7/40)		24% (9/38)		81% (33/40)	83% (6/7)	81% (27/33)
High	38% ($n = 46$)	36% (16/44)		43% (18/42)		61% (27/44)	88% (14/16)	46% (13/28)
70-Gene prognosis								
Low risk	9% ($n = 11$)	0% (0/11)	0.12	18% (1/9)	0.28	100% (11/11)	NA (0/0)	100% (11/11)
High risk	91% ($n = 109$)	24% (25/105)		31% (31/99)		75% (80/105)	88% (22/25)	72% (72/80)

Table 2 continued

	Overall distribution available for analysis, % (n)	pCR, % (n)	pCR vs. not P value	RCB 0, I, % (n)	RCB 0, I vs. II, III P value	3 year RFS		
						Overall	pCR	NO pCR
Wound healing								
Quiescent	25%	7%	0.04	27%	0.81	100%	100%	100%
	(n = 30)	(2/29)		(7/26)		(29/29)	(2/2)	(27/27)
Activated	75%	26%		30%		71%	86%	65%
	(n = 90)	(23/87)		(25/82)		(62/87)	(20/23)	(42/64)
p53 Mutation signature								
Wildtype	49%	9%	0.001	19%	0.02	91%	100%	90%
	(n = 59)	(5/58)		(10/53)		(53/58)	(5/5)	(48/53)
Mutation	51%	34%		40%		65%	84%	54%
	(n = 61)	(20/58)		(22/55)		(38/58)	(17/20)	(21/38)
P53 Gene chip	n = 144	n = 140		n = 132				
Wild type	59%	16%	0.04	23%	0.08	85%	85%	85%
	(n = 85)	(13/83)		(18/79)		(71/83)	(11/13)	(60/70)
Any mutation	41%	32%		38%		71%	94%	61%
	(n = 59)	(18/57)		(20/53)		(41/57)	(17/18)	(24/39)
Missense Zn-binding	10%	27%		29%		72%	75%	72%
	(n = 15)	(4/15)		(4/14)		(11/15)	(3/4)	(8/11)
Null	12%	47%		47%		80%	100%	62%
	(n = 16)	(8/16)		(7/15)		(13/16)	(8/8)	(5/8)
MIP arrays	n = 129	n = 125		n = 116				
Amplification 17q	17%	45%	0.03	50%	0.03	95%	100%	92%
	(n = 21)	(8/20)		(10/20)		(19/20)	(8/8)	(11/12)
No amplification 17q	83%	15%		25%		76%	88%	74%
	(n = 108)	(16/105)		(23/96)		(80/105)	(14/16)	(66/89)
Ki67 IHC	n = 154	n = 150		n = 141				
Low	30%	7%	<0.001	14%	0.005	82%	100%	81%
	(n = 47)	(3/45)		(6/43)		(37/45)	(3/3)	(34/42)
Medium	32%	10%		18%		79%	75%	79%
	(n = 49)	(5/48)		(8/44)		(38/48)	(4/5)	(34/43)
High	38%	35%		43%		73%	95%	61%
	(n = 58)	(20/57)		(23/54)		(42/57)	(19/20)	(23/37)

IHC immunohistochemistry, *FISH* fluorescence-in situ hybridization assay, *HR* hormone receptor, *ROR-S* risk of relapse subtype score, *T* trastuzumab, *RCB* is residual cancer burden. RCB 0, 1 refers to absence of any invasive cancer (RCB 0) or minimal residual disease (RCB 1), *ROR-S* risk of recurrence-S. The first four rows of the IHC results show the entire evaluable population, whereas the IHC rows with *italic values* show the population broken down by the use of neoadjuvant trastuzumab (w/T) or without (w/o T)

observations are consistent with our results and with adjuvant studies that show patients with HR-negative disease benefit more from chemotherapy [34] than do patients with HR-positive disease.

Molecular profiles may provide the opportunity to identify, beyond HR and HER2 status, what might be driving tumor behavior and outcomes. In a multivariate model, when receptor types were fixed, the factors that added to RFS included clinical stage, wound healing signature, ROR-S, and p53 predicted mutation. When pCR was also fixed, most of the dichotomized molecular markers added some additional predictive value, likely because of the ability to identify patients in the "no pCR" group who have excellent outcomes, largely the HR+/HER2− subgroup, though not exclusively. Given that the low proliferative HR+ subset is at risk for late recurrence, longer follow-up and additional studies will be required to validate this observation.

Table 3 Correlations among the molecular signatures

Correlation coefficient

Concordance (*P* value)

	p53 wt	70-Gene low	Luminal	Wound-quiescent	ROR-S low
HR+, HER2−	0.44	0.26	0.62	0.34	0.40
	0.72	0.63	0.81	0.68	0.71
	(<0.001)	(0.002)	(<0.001)	(<0.001)	(<0.001)
p53 wt	⇓	0.32	0.65	0.55	0.56
	⇒	0.59	0.83	0.73	0.74
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
70-Gene low		⇓	0.33	0.41	0.54
		⇒	0.61	0.82	0.83
			(<0.001)	(<0.001)	(<0.001)
Luminal			⇓	0.40	0.55
			⇒	0.68	0.75
				(<0.001)	(<0.001)
Wound-quiescent				⇓	0.63
				⇒	0.86
					(<0.001)

HR hormone receptor, *wt* wild type, *ROR-S* risk of relapse score. Signatures are shown as dichotomies: HR+/HER2− vs. not, luminal (A and B) vs. the other intrinsic subtypes (HER2 enriched, basal, and normal-like), p53 wild type vs. mutation, 70-gene prognosis low vs. high (only 13 (9%) patients are in the low risk prognosis group), and wound-quiescent vs. wound-activated. Concordance refers to the lower risk subsets

Table 4 Pathological complete response and recurrence-free survival by molecular subtypes

	Pathological complete response (pCR)		Recurrence-free survival	
	Rate of pCR (*n*)	Odds ratio (*P* value)	Hazard ratio (95% CI)	Hazard ratio, pCR vs. no PCR (95% CI within subgroup)
All patients with surgery	27% (215)	–	–	0.41* (0.18–0.82)
Population without any trastuzumab	22% (180)	–	–	0.23 * (0.06–0.63)
HR+/HER2− (Yes vs. other)	Yes: 9% (8/93)	0.15	0.50*	Yes: 0.00 (–)
	Other: 38% (30/79)	(<0.001)	(0.27–0.89)	Other: 0.17* (0.04–0.51)
p53 (Wt vs. Mut)	Wt: 9% (5/58)	0.18	0.34*	Wt 0.00 (–)
	Mut: 34% (20/58)	(<0.001)	(0.15–0.69)	Mut: 0.22* (0.05–0.65)
70-Gene (low vs. high)	Low: 0% (0/11)	0.00	0.00*	Low: 0.00 (–)
	High: 24% (25/105)	(0.02)	(–)	High: 0.29* (0.07–0.82)
Luminal PAM50 (luminal vs. other)	Luminal: 8% (5/61)	0.16	0.47*	Luminal: 0.00 (–)
	Other: 36% (20/55)	(<0.001)	(0.23–0.93)	Other: 0.26* (0.06–0.79)
Wound healing (quiescent vs. activated)	Quiescent: 7% (2/29)	0.21	0.16*	Quiescent: 0.00 (–)
	Activated: 26% (23/87)	(0.02)	(0.03–0.51)	Activated: 0.25* (0.06–0.71)
ROR-S (low vs. med/HIGH)	Low: 6% (2/32)	0.18	0.22*	Low: 0.00 (–)
	Med/high: 27% (23/84)	(0.006)	(0.05–0.61)	Med/high: 0.26* (0.06–0.74)

HR hormone receptor, *Wt* wild type, *Mut* mutation, *ROR-S* risk of relapse score, *Med* medium. * Denotes significant proportional hazard ratio (likelihood ratio *P* < 0.05). A value of 0.00 indicates that there were no recurrences in this category among patients who had a pCR. The signatures were dichotomized for the purposes of comparison among the signatures. Note that the hazard ratio for the prediction of pCR vs. not is better within the dichotomized molecular classifications than it is for the population as a whole. The greatest difference in RFS is seen with HR+/HER2− vs. not (hazard ratio 0.17, [95% CI 0.04–0.51]) which is also where the greatest difference in the rates of pCR were observed (9% vs. 38%). For the dichotomized breast cancer molecular classifiers, the difference between highest and lowest rates of pCR varied from 19 to 29%

Table 5 Univariate vs. multivariate Cox analyses adjusting for predefined HR/HER2 contribution

	Univariate hazard ratio (95% CI)	Multivariate hazard ratio adjusting for predefined HR/HER2 contribution (95% CI)
P53 (Wt vs. Mt)	0.33*	0.46*
	(0.16–0.70)	(0.22–0.96)
70-Gene (low vs. high)	0.00[a]	0.00[a]
	(–)	(–)
Luminal PAM50 (luminal vs. other)	0.47*	0.73
	(0.23–0.94)	(0.37–1.48)
Wound healing (quiescent vs. activated)	0.16*	0.20*
	(0.04–0.65)	(0.05–0.78)
ROR-S (low vs. med/high)	0.22*	0.29*
	(0.07–0.72)	(0.09–0.96)
Clinical stage (<stage 3 vs. not)	0.20*	0.21*
	(0.10–0.42)	(0.10–0.43)
Ki67 (low/med vs. high)	0.77	1.14
	(0.42–1.42)	(0.62–2.10)

Analyses were restricted to the patient subset who did not receive any trastuzumab treatment with known HR/HER2 status, which accounts for the slight discrepancy between the univariate hazard ratios reported here and Table 4. All signatures were dichotomized (if there were more than two categories) as follows: HR+/HER2− vs. others, luminal (A and B) vs. other intrinsic subtypes, p53 wild type vs. mutation, 70-gene prognosis signature low vs. high and wound-quiescent vs.-activated, clinical stage 3 (including inflammatory) vs. earlier stage

* Wald test $P < 0.05$

[a] No events were observed among the 11 patients in the 70-gene low group, which renders the Wald test for significance unreliable

Molecular signatures are currently being used to identify low risk patients who are less likely to benefit from chemotherapy regardless of nodal status [35] or in the setting of HR+ node-negative disease [36]. Such patients have been shown to have low rates of response to chemotherapy and very low rates of early recurrence [37]. Confirmation of chemotherapy benefit in molecularly low risk patients will be forthcoming from the TAILORx [38] and MINDACT [39] trials. In the follow-on I-SPY 2 TRIAL, an adaptive-design neoadjuvant trial to test the ability of phase 2 agents in combination with chemotherapy to increase pCR, 70-gene low risk, HR-positive and HER2-negative patients are being excluded from randomization. In I-SPY 1, none of the 11 patients with a 70-gene prognosis profile had a pCR or a recurrence (Fig. 2).

In the I-SPY cohort, the wound healing signature identified the largest fraction of low risk patients (based on RFS) of any signature. The genes consistent with an activated wound environment characterize women with poor outcomes, in keeping with increasing evidence that supports targeting the inflammatory pathway in high risk cancers [40] and breast cancer in particular [41, 42]. The activated wound healing signature is associated with poor outcomes across multiple tumor types and may well reflect the importance of the microenvironment in tumor behavior.

Although pCR and RCB are very predictive of RFS among the poor prognosis molecular profiles, the profiles do not predict an individual patient's response to standard chemotherapy. A substantial fraction of tumors with the highest risk features have a complete response to therapy and do well, while others with that same signature have a poor response and poor outcome. Ongoing analysis is focusing on the I-SPY 1 patients who did not have a complete response to therapy and had early recurrence, using the described biomarkers as well as phosphoprotein profiles, to explore targets for future therapeutic intervention.

Our study is limited by the short follow-up time. Patients with HR-positive tumors continue to be at risk for recurrence for many years, and early recurrence data may not reflect the overall outcome [34]. However, in this select group of patients where almost all patients had grade 2 or 3 disease, recurrence risk is likely to be concentrated in the first 5 years [43]. The Oxford Overview Analysis of the early breast cancer trials strongly suggests that the benefit of chemotherapy is reflected by distant disease-free survival at 5 years, where the survival curves for patients with chemotherapy versus not initially diverge but are then parallel, so any survival benefit from chemotherapy is likely to be manifest in the first 5 years [44]. The median follow-up period of 3.9 years in the I-SPY cohort should reflect the benefits in HER-positive

Fig. 2 Stratification, by molecular classifier, of the hormone receptor positive HER2 negative subgroup that did not achieve a pathologic complete response. The patients in the HR+/HER2− subgroup that did not achieve a pathologic complete response are stratified by the molecular subtypes as shown: **a** 70-gene prognosis profile (*Blue line* low risk/*gold line* high risk); **b** wound healing signature (*Blue line* quiescent; *gold line* activated); **c** risk of relapse subtype score (ROR-S) (*Blue line* low risk/*Gold line* medium and high risk); **d** p53 predicted mutation (*Blue line* predicted wild type/*Gold line* predicted mutation); **e** clinical stage (*Blue line* clinical stage 2/ *Gold line* clinical stage 3). Stratification of the "no pCR" HR+/HER2− patient group by molecular signatures and clinical stage
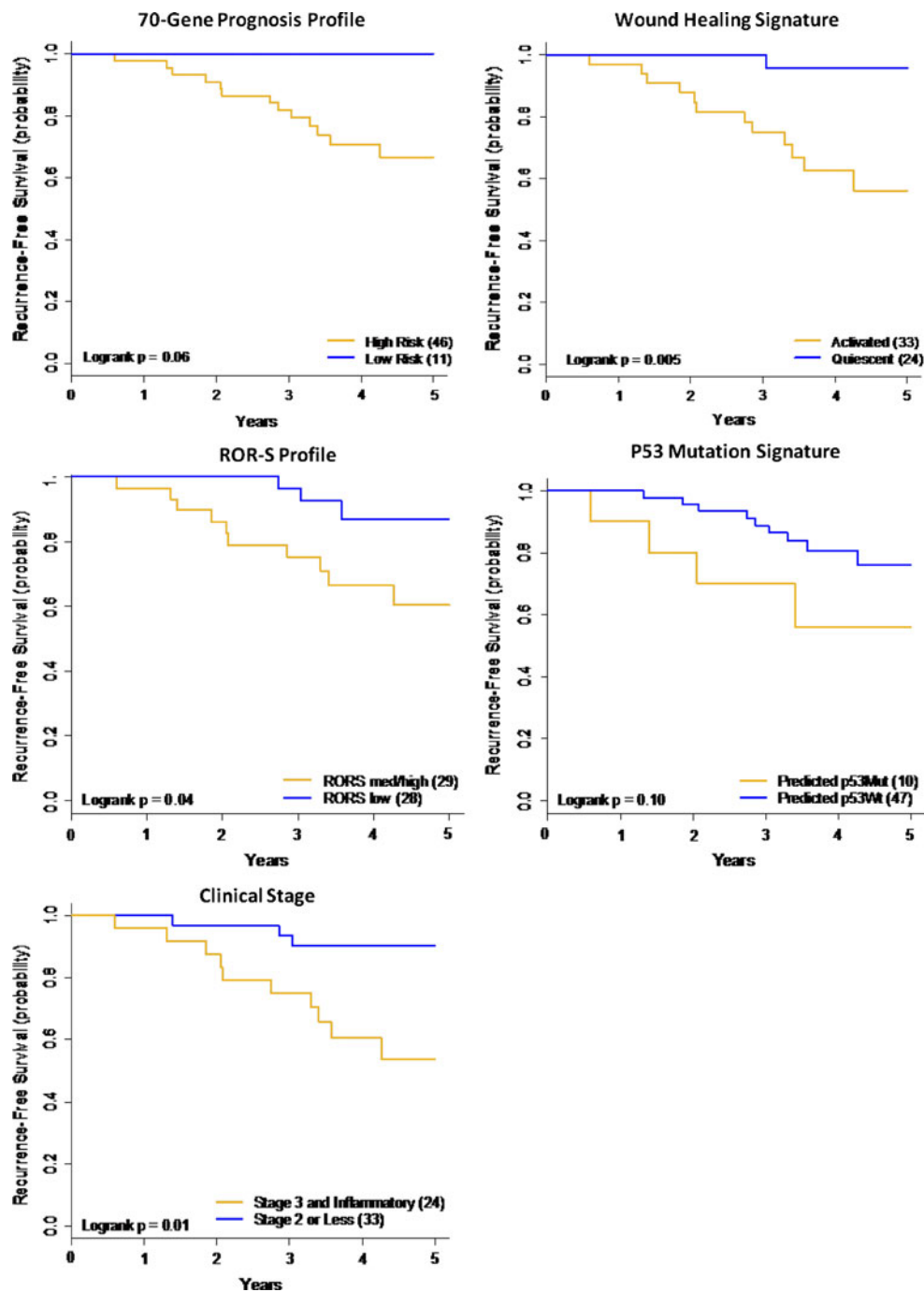
and triple negative disease, where the risk of recurrence is early [13].

Molecular and biological heterogeneity were substantial even within the high risk group of patients in the I-SPY TRIAL. In this patient cohort, HR and HER2 status were the most predictive of pCR, but the molecular signatures add to the ability of the receptors to predict RFS. The task that remains is to use current and emerging markers to identify optimal biological subsets for new therapeutic agents. Importantly, molecular marker data should be collected routinely in trials so that markers and imaging that are early predictors of outcome can be related to the target endpoint of RFS [45]. The I-SPY 1 database, with its rich resource of genomic and protein expression data, is an important resource to explore emerging and new biomarkers associated with resistance and response to standard therapy.

Fig. 3 Heat map by molecular features and outcomes for all patients

References

1. van de Vijver MJ, He YD, Van't Veer LJ et al (2002) A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 347:1999–2009
2. Paik S, Tang G, Shak S et al (2006) Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol 24:3726–3734
3. Cheang MC, Chia Sk, Boduc D et al (2009) Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. J Natl Cancer Inst 101:736–750
4. 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

5. Rouzier R, Perou CM, Symmans WF et al (2005) Breast cancer molecular subtypes respond differently to preoperative chemotherapy. Clin Cancer Res 11:5678–5685

6. Carey LA, Dees EC, Sawyer L et al (2007) The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. Clin Cancer Res 13:2329–2334

7. Esserman LJ, van't Veer LJ, Perou CM et al (2009) Biology of breast cancers that present as interval cancers and at young age should inform how we approach early detection and prevention. Cancer Res 69:2 (suppl 1; abstr 6034)

8. Fisher B, Brown A, Mamounas E et al (1997) Effect of preoperative chemotherapy on local-regional disease in women with operable breast cancer: findings from National surgical adjuvant breast and bowel project B-18. J Clin Oncol 15:2483–2493

9. Symmans WF, Peintinger F, Hatzis C et al (2007) Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. J Clin Oncol 25:4414–4422

10. Bear HD, Anderson S, Smith RE et al (2006) Sequential preoperative or postoperative docetaxel added to preoperative doxorubicin plus cyclophosphamide for operable breast cancer: National surgical adjuvant breast and bowel project protocol B-27. J Clin Oncol 24:2019–2027

11. Baselga J, Bradbury I, Eidtmann H et al (2010) First results of the NeoALTO Trial (BIG 01-06EGF 106903): a phase III, randomized, open label, neoadjuvant study of lapatinib, trastuzumab, and their combination plus paclitaxel in women with HER2-positivie primary breast cancer. Cancer Res 70:24 (suppl 2; abstr S3-3)

12. Untch M, Fasching PA, Konecny GE et al (2010) Pathologic complete response after neoadjuvant chemotherapy + trastuzumab treatment predicts survival and detects a patient subgroup at high need for improvement of anti-HER2 therapy. Three year medican follow-up data of the TECHNO Trial. Cancer Res 70:24 (suppl 2; abstr P1-11-03)

13. Esserman LJ, Berry DA, DeMichele A, Carey L, Davis SE, Buxton MB, Hudis C, Gray J, Perou C, Yau C, Livasy C, Krontiras H, Montgomery L, Tripathy D, Lehman C, Liu MC, Olopade OI, Rugo HS, Carpenter JT, Dressler L, Chhieng D, Singh B, Mies C, Rabban J, Chen Y-Y, Giri D, van 't Veer, Hylton N, on behalf of the I-SPY 1 TRIAL Investigators (2012) Pathologic complete response predicts recurrence-free survival more effectively by cancer subset: results from the I-SPY 1 TRIAL (CALGB 150007/150012; ACRIN 6657. J Clin Oncol (in press)

14. Hylton NM, Blume J, Bernreuter W et al Prediction of response to neoadjuvant chemotherapy for women with locally-advanced breast cancer: results from the ACRIN 6657/I-SPY Trial. Radiology (in press)

15. Allred DC, Harvey JM, Berardo M et al (1998) Prognostic and predictive factors in breast cancer by immunohistochemical analysis. Mod Pathol 11:155–168

16. Livasy C, Carey L, DeMichele A et al (2009) Biomarkers associated with pathologic complete response to neoadjuvant chemotherapy in women with locally advanced breast cancer: results from the I-SPY TRIAL (CALGB 150007/150012 & ACRIN 6657). Cancer Res 69:2 (suppl 1; abstr 5102)

17. Spyratos F, Ferrero-Pous M, Trassard M et al (2002) Correlation between MIB-1 and other proliferation markers: clinical implications of the MIB-1 cutoff value. Cancer 94(8):2151–2159

18. Hu Z, Fan C, Oh DS et al (2006) The molecular portraits of breast tumors are conserved across microarray platforms. BMC Genomics 7:96

19. 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

20. van't Veer LJ, Dai H, van de Vijer MJ et al (2002) Gene expression profiling predicts clinical outcome of breast cancer. Nature 415:530–536

21. Glas AM, Floore A, Delahaye JL et al (2006) Converting a breast cancer microarray signature into a high-throughput diagnostic test. BMC Genomics 7:278

22. Chang HY, Nuyten DS, Sneddon JB et al (2005) Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. Proc Natl Acad Sci USA 102:3738–3743

23. Troester MA, Herschkowitz JI, Oh DS et al (2006) Gene expression patterns associated with p53 status in breast cancer. BMC Cancer 6:276

24. Hardenbol P, Baner J, Jain M et al (2003) Multiplexed genotyping with sequence-tagged molecular inversion probes. Nat Biotechnol 21:673–678

25. Hardenbol P, Yu F, Belmont J et al (2005) Highly multiplexed molecular inversion probe genotyping: over 10,000 targeted SNPs genotyped in a single tube assay. Genome Res 15:269–275

26. Ji H, Kumm J, Zhang M et al (2006) Molecular inversion probe analysis of gene copy alterations reveals distinct categories of colorectal carcinoma. Cancer Res 66:7910–7919

27. Olshen AB, Venkatraman ES, Lucito R et al (2004) Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics 5(4):557–572

28. Conway K, Edmiston SN, Cui L et al (2002) Prevalence and spectrum of p53 mutations associated with smoking in breast cancer. Cancer Res 62:1987–1995

29. Hudis CA, Barlow WE, Constantino JP et al (2007) Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: the STEEP system. J Clin Oncol 25:2127–2132

30. Ihaka R, Gentleman R (1996) R, A language for data analysis and graphics. J Comput Graph Stat 5(3):299–314

31. 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(14):8418–8423

32. Lin C, Buxton MB, Moore D et al (2011) Locally advanced breast cancer are more likely to present as interval cancers: results from the I-SPY 1 TRIAL (CALGB 150007/150012, ACRIN 6657, InterSPORE Trial). Breast Cancer Res Treat. doi: 10.1007/s10549-011-1670-4

33. Kuerer HM, Newman LA, Smith TL et al (1999) Clinical course of breast cancer patients with complete pathologic primary tumor and axillary lymph node response to doxorubicin-based neoadjuvant chemotherapy. J Clin Oncol 17:460–469

34. Berry DA, Cirrincinoe C, Henderson IC et al (2006) Estrogen-receptor status and outcomes of modern chemotherapy for patients with node-positive breast cancer. JAMA 295:1658–1667

35. Buyse M, Loi S, van't Veer L et al (2006) Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. J Natl Cancer Inst 98(17):1183–1192

36. Paik S, Shak S, Tang G et al (2004) A multigene assay to predict recurrence to tamoxifen-treated, node-negative breast cancer. N Engl J Med 351(27):2817–2826

37. Knauer M, Mook S, Rutgers EJ et al (2010) The predictive value of the 70-gene signature for adjuvant chemotherapy in early breast cancer. Breast Cancer Res Treat 120(3):655–661

38. Sparano JA (2006) TAILORx: trial assigning individualized options for treatment (Rx). Clin Breast Cancer 7(4):347–350

39. Cardoso F, van't Veer L, Rutgers EJ et al (2008) Clinical application of the 70-gene profile: the MINDACT trial. J Clin Oncol 26(5):729–735

40. Condeelis J, Pollard JW (2006) Macrophages: obligate partners for tumor cell migration, invasion, and metastasis. Cell 124(2):263–266

41. Coussens LM, Pollard JW (2010) Leukocytes in mammary development and cancer. Cold Spring Harb Perspect Biol 2(10):a003202

42. Mukhtar RA, Nseyo O, Campbell MJ et al (2011) Tumor-associated macrophages in breast cancer as potential biomarkers for new treatments and diagnostics. Expert Rev Mol Diagn 11(1):91–100

43. Esserman LJ, Moore DH, Tsing PJ et al (2011) Biologic markers determine both the risk and the timing of recurrence in breast cancer. Breast Cancer Res Treat 129(2):263–266

44. 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

45. Berry DA (2010) The hazards of endpoints. J Natl Cancer Inst 102(18):1376–1377