Tailoring adjuvant therapy for hormone receptor-positive breast cancer

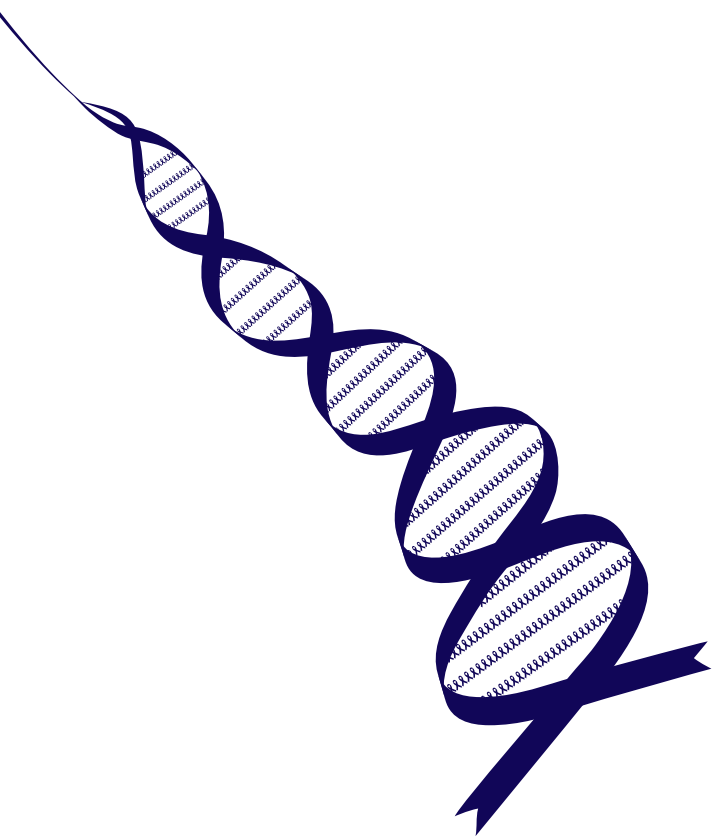
Blok, E.J.

Chapter 6

ER pathway activity as a predictive marker during neo-adjuvant endocrine therapy in early breast cancer; Results of the TEAM IIA trial

E.J. Blok
M. Alves de Inda
P.J.K. Kuppen
A.C. Charehbili
E. den Biezen-Timmermans
S. Fruytier
W.M. Meershoek-Klein Kranenbarg
A. van Brussel
S.C. Linn
W. Verhaegh
A. van de Stolpe
C.J.H. van de Velde
J.R. Kroep

Manuscript as submitted to
Cancer Research



Abstract

Endocrine therapy is an important asset in the management of estrogen receptor (ER) positive breast cancer. Currently, patients are selected for endocrine therapy based on immunohistochemical expression of ER and progesterone receptor (PR). However, this does not necessarily imply an active ER pathway, which is the target of endocrine therapy. The aim of this study was to validate a recently described computational model that infers ER pathway activity based on the expression of its target genes.

Two cohorts were analyzed: a cohort of 61 patients treated with 3 months of neo-adjuvant letrozole for which a public dataset containing gene expression and ultrasound-based tumor-response data is available, and the TEAM-IIA trial cohort in which 102 patients were treated with up to 6 months of neo-adjuvant exemestane. The original ER pathway model was adapted to process Affymetrix HGU133A data from the public dataset, and qPCR data obtained from TEAM-IIA samples.

Mean ER pathway activity decreased significantly during therapy. Furthermore, in the public dataset, both baseline activity and decrease in ER pathway activity was significantly higher in responding patients. In the TEAM-IIA cohort, palpation-based progressive disease and radiology-based non-response after therapy were associated with lower levels of baseline ER pathway activity.

Our results indicate, in two independent cohorts, that low baseline ER pathway activity is associated to an inferior response to endocrine therapy. Upon prospective validation, this model could be used in a clinical setting to predict response to endocrine therapy and thereby better select patients who will benefit from this treatment.

Introduction

Endocrine therapy is one of the mainstays in the treatment of both early and metastatic breast cancer. Especially the use of tamoxifen and aromatase inhibitors (AI) has resulted in increased survival rates.[1-3] Patients are currently selected for endocrine therapy using immunohistochemical analysis of estrogen receptor (ER) and progesterone receptor (PR) expression, which was developed more than a decade ago.[4] Both the American Society for Clinical Oncology (ASCO) and the European Society for Medical Oncology advise a threshold of 1% ER positive tumor cells.[5, 6] In practice, many clinicians and countries choose a threshold of 10%.[7-9] More quantifiable analyses like the Allred scoring and H-score have been developed and suggested for clinical application, but are currently not routinely used.[5]

Despite the success of endocrine therapy in ER-positive breast cancer, there are still patients that do not respond to endocrine therapy, regardless of the presence of ER or PR in the investigated tissue sample. In addition to cancer tissue heterogeneity, several mechanisms have been proposed to explain lack of therapy response, like emergence of activating mutations in ESR1 or activation of other signal transduction pathways upon pharmacological inhibition of the ER pathway.[10, 11] Standard immunohistochemical analysis only detects presence of ER and PR; however, the issue of how well a positive nuclear ER staining, accompanied or not by a positive nuclear PR staining, actually indicates an active ER pathway, as an alternative explanation, has not yet been satisfactorily addressed. The development of tests to predict response to endocrine therapy based on measuring actual activity of the ER pathway will provide additional information that might be useful in the decision to treat with endocrine therapy and whether to add additional targeted therapy.

A possible approach towards predicting the response of endocrine therapy would be to assess activity of the ER pathway by measuring expression of downstream ER pathway target genes. After all, it is likely that if this pathway is highly active, endocrine therapy will be more effective than when the pathway is barely active or inactive, irrespective of the presence of the estrogen receptor itself. So far, no test has been developed for assessing ER pathway activity, although numerous studies have been performed.[12-15] Recently, Verhaegh *et al* have developed a computational model for the ER pathway, enabling assessment of this pathway in tumor tissue. This computational Bayesian network model uses mRNA expression of 27 genes which are proven target

genes of the ER pathway.[16, 17] The aim of the current work is to evaluate this model in a clinical setting. We hypothesize that ER pathway activity at baseline as measured by this computational model is capable of predicting response to therapy and that a decrease in pathway activity during treatment represents a successful treatment.

To test our hypothesis, we investigated publicly available Affymetrix datasets containing baseline and outcome data from neoadjuvant studies on AIs in breast cancer patients.[18, 19] After initial confirmation of our hypothesis we proceeded to test it in the TEAM IIA clinical trial cohort of ER positive breast cancer patients treated with neo-adjuvant endocrine therapy. We assessed the value of baseline ER pathway activity to predict response to neo-adjuvant therapy with an AI and whether the decrease in activity could be used to assess clinical response to neoadjuvant therapy.

Materials & Methods

Publicly available dataset

Search of the GEO repository yielded one dataset appropriate for our proof-of-principle analysis: the GSE20181 dataset.[18, 19] This dataset contains Affymetrix HGU133A gene expression and outcome data of 61 post-menopausal patients that underwent neo-adjuvant letrozole treatment. Those patients were treated in a neo-adjuvant setting with 2.5 mg daily letrozole for three months. The GSE20181 dataset contains gene expression data from biopsies, containing at least 20% tumor collected at baseline, 14 days, and three months of treatment. Clinical response was determined by ultrasound assessment of the tumor size and patients with more than 50% tumor reduction were considered responders. Raw Affymetrix data from the GSE20181 dataset was retrieved from the GEO repository and normalized with fRMA.[20]

Study cohort

The Dutch TEAM IIA trial is a neo-adjuvant phase II trial for which the details have been previously described.[21] Briefly, 102 ER positive patients were randomized for either 3 or 6 months of neo-adjuvant exemestane. Due to unforeseen slow accrual, the study changed to a single arm design with 6 months of therapy. Standard clinicopathological baseline characteristics, including PR and HER2 status were known. Pre-treatment biopsies and post-treatment resection specimens were collected by the investigators. Change in tumor size assessed by palpation was the primary objective. Secondary

outcomes were clinical response rates measured by mammography, ultrasound and MRI. Assessment was performed according to RECIST 1.1 criteria.

Laser Capture Microdissection

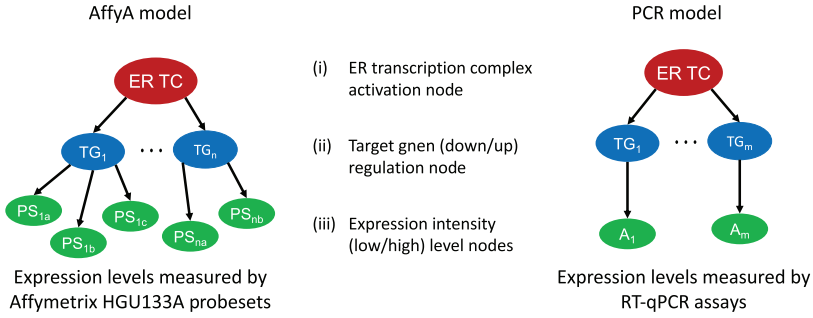
In order to accurately determine the activity of the ER pathway within tumor cells, those cells were separated from surrounding stromal tissue using Laser Capture Microdissection (LCM), as described by Espina *et al.*[22] Briefly, slides were cut in quintuplicate from tumor-containing formalin fixed paraffin embedded (FFPE) blocks from both pre-treatment biopsies as well as from post-treatment resection material. The slides were cut and placed onto PEN Membrane slides (Life Technologies), which contain a special membrane making it suitable for LCM. Slides were stained with hematoxylin and eosin using standard RNase free protocols. All slides were microscopically assessed and tumor regions (at least 1mm^2 per slide) were dissected from the slides using LCM. For each sample, the dissected tumor regions were collected in an RNase-free microfuge tube (Life Technologies) and were stored at -20 degrees Celsius.

ER pathway models

The methodology used to develop the ER pathway models used in this study was described in detail earlier.[16] The basic idea behind this methodology is to construct a Bayesian network model of the ER transcriptional program, which uses the pathway target genes' mRNA levels in a certain sample to infer the probability that the ER pathway is transcriptionally active in the respective sample. This Bayesian network structure is a simplified model of the transcriptional program of the ER signal transduction pathway, consisting of three types of nodes: (i) ER transcription complex activation node, (ii) target gene regulation node (with states 'down' and 'up'), and (iii) expression intensity level nodes (with states 'low' and 'high') each corresponding to a target gene. The final model describes (a) how target gene regulation depends on ER transcription complex activity and (b) how expression level intensities in turn depend on regulation of the respective target genes (Supplemental Figure 1).

The original paper described this methodology for mRNA expression levels measured using the Affymetrix HGU133Plus2.0 microarray platform.16 This methodology was adapted to be used with both RT-qPCR and Affymetrix HGU133A platforms. This adaptation consisted of modifying and re-calibrating the ER pathway Bayesian network model for each platform. For the RT-qPCR platform the original 27 target gene ER

pathway Bayesian network was reduced to a 12 ER target gene network. The genes included in this network were chosen based on literature evidence and discriminative power. RT-qPCR assays were developed and validated for each of these target genes plus 7 other genes used as reference genes. The model was calibrated using mRNA expression data of MCF7 cell cultures exposed to 1nM E2 for 16h or DMSO after being maintained in estradiol deprived (charcoal treated FBS, phenol red free) medium for 48h, as described before.[23] For the Affymetrix HGU133A platform the network was adapted by selecting Affymetrix probe sets representing the original 27 ER target genes that were available in the HGU133A platform and the resulting network model was calibrated using fRMA transformed Affymetrix HGU133A mRNA expression data from publicly available dataset GSE9936. These Affymetrix HGU133A calibration samples were from MCF7 cell cultures exposed to 6nM E2 for 24h or vehicle control after being maintained in estradiol deprived (charcoal treated FBS, phenol red free) medium for at least 48h.[24]



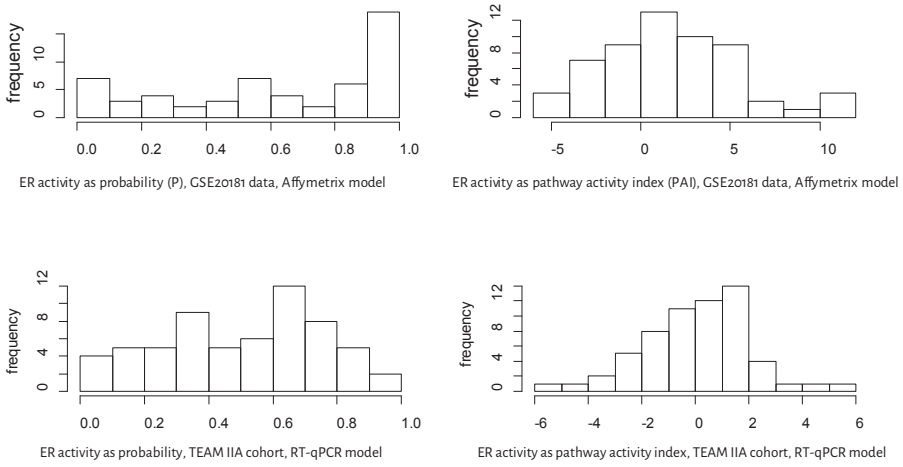
Suppl Figure 1: schematic representation of ER pathway network. Left: RT-qPCR model consisting of 12 target genes with expression levels measured by RT-qPCR assays. Center: Node type description. Right: Affimetrix model consisting of 27 target genes with expression levels measured by Affymetrix hgu133a probesets.

Pathway analysis of PCR samples

For the analysis of the ER pathway activity on the samples of the TEAM IIA cohort, mRNA was extracted from the micro dissected samples (Siemens VERSANT Tissue Prep. Reagent kit, according to manufacturer's instructions) and the expression of the 12 target genes and 7 reference genes was determined using RT-qPCR. Those values were used as input to the RT-qPCR ER pathway model with which the probability and the odds that the ER pathway is active was calculated. These data were interpreted with respect to ER pathway activity in a blinded manner at Philips Research and returned to the Leiden University Medical Center, in order to be correlated to tumor size, as assessed by palpation and mammography.

Statistical analysis

Statistical analysis was performed using R[25] and SPSS 23.0 (IBM). ER pathway analysis was calculated in two ways: the probability (P) that the ER pathway is active and its *pathway activity index* (PAI, defined as $\log_2(P/(1-P))$, i.e., the $\log_2$ of the odds in favor that the ER pathway is active, PAI=0 indicates a 50% chance of ER pathway activation, PAI=2 and -1 indicate, respectively, a chance four times as large and twice as small of the ER pathway being active than of it being inactive). For the purpose of comprehensibility, the probability is used as a visual representation of the likelihood that the ER pathway is active. However, the activity index is used for all statistical calculations since these data are generally more appropriate for statistical computations and more normal-like distributed (Supplemental Figure 2). For simplicity, we use 'ER pathway activity', or simply, 'ER activity' as a short form of 'the inferred probability that the ER pathway is active'. Paired t-tests were used to assess differences between baseline and post-treatment ER pathway activity. Two sample t-tests were used to assess correlations between outcome categories and average ER pathway activity. Two sample t-tests and ANOVA were used to access ER activity and decrease in activity association with baseline parameters. Presented p-values are 2 sided and refer to PAI quantities, unless explicitly indicated otherwise. All tests assume unequal variance.



Suppl Figure 2: histograms of ER-pathway activity at baseline. Upper: GSE20181 data, using Affymetrix model; Lower: TEAM II A cohort, using RT-qPCR model; Left: activity presented as a probability (P); Right: activity presented as Pathway Activity Index, i.e.: $\log_2(P/(1-P))$.

Results

GSE20181 proof-of-principle dataset

We first assessed, in a public dataset, whether the ER pathway model based on Affymetrix HGU133A microarray data, was able to detect a decrease in ER pathway activity in ER positive patients following neoadjuvant endocrine therapy. ER pathway activity significantly decreased during letrozole therapy on ER positive patients of the GSE20181 dataset (table 1) presents the mean ER pathway activity at baseline, 2 weeks, and 3 months. Overall, ER pathway mean activity decreased significantly between baseline (P=0.62) and 2 weeks (P=0.32) of treatment (with PAI decrease from 1.6 to -1.6, paired t-test p-value <0.001, n=58) as well as between baseline and 3 months (P=0.36) of therapy (with PAI decrease from 1.8 to -1.3, p-value <0.001, n=56). No significant difference was seen between 2 weeks and 3 months of treatment, suggesting an early maintained response to letrozole. Both baseline ER pathway activity and decrease in ER activity from baseline was significantly higher in responders than in non-responders (Table 1 and Figure 1).

Table 1: Average activity and decrease in activity of ER pathway for GSE20181. ER pathway activity is presented as probability that the ER pathway is active (P) and as activity index (PAI). R: responders, NR: non-responders;; sd: standard deviation; pv: p-value for t-test on PAI for R vs. NR

	P				PAI			
	N	Mean	sd	Range	Mean	sd	Range	p-value*
Baseline								0.02
all	58	0.62	0.33	[0.02,1]	1.6	3.8	[-5.8,12]	
R	37	0.65	0.34	[0.02,1]	1.9	3.9	[-5.8,10]	
NR	15	0.46	0.31	[0.06,0.92]	-0.3	2.4	[-3.9,3.5]	
2 weeks								0.86
all	58	0.32	0.27	[0.01,0.96]	-1.6	2.6	[-6.7,4.5]	
R	37	0.33	0.29	[0.01,0.96]	-1.6	2.9	[-6.7,4.5]	
NR	15	0.32	0.23	[0.01,0.76]	-1.5	2	[-6.2,1.7]	
3 months								0.86
all	60	0.38	0.28	[0.01,1]	-1.1	2.6	[-6.6,8.2]	
R	36	0.36	0.26	[0.03,0.86]	-1.2	2	[-4.8,2.7]	
NR	14	0.40	0.33	[0.01,0.91]	-1.0	3	[-6.4,3.4]	
Base - 2 w								0.02
all	58	0.30	0.35	[-0.56,0.93]	3.2	3.5	[-3.9,10]	
R	37	0.31	0.33	[-0.33,0.93]	3.5	3.5	[-3.1,10]	
NR	15	0.14	0.35	[-0.56,0.68]	1.2	2.7	[-3.9,5.1]	

Table 1: continued

		P			PAI			
	N	Mean	sd	Range	Mean	sd	Range	p-value*
Base - 3 m								0.04
all	56	0.28	0.37	[-0.8,0.89]	3.1	3.8	[-6.4,12]	
R	36	0.31	0.31	[-0.46,0.88]	3.3	3.4	[-4.3,10]	
NR	14	0.07	0.43	[-0.8,0.82]	0.76	3.8	[-6.4,8.6]	
2 w - 3 m								0.96
all	56	-0.03	0.23	[-0.66,0.41]	-0.29	2.2	[-5.1,3.8]	
R	36	-0.02	0.23	[-0.63,0.41]	-0.35	2.1	[-5.1,3.7]	
NR	14	-0.07	0.25	[-0.66,0.23]	-0.39	2.5	[-5,3.8]	

*p-value represents a t-test comparing responders (R) with non-responders (NR)

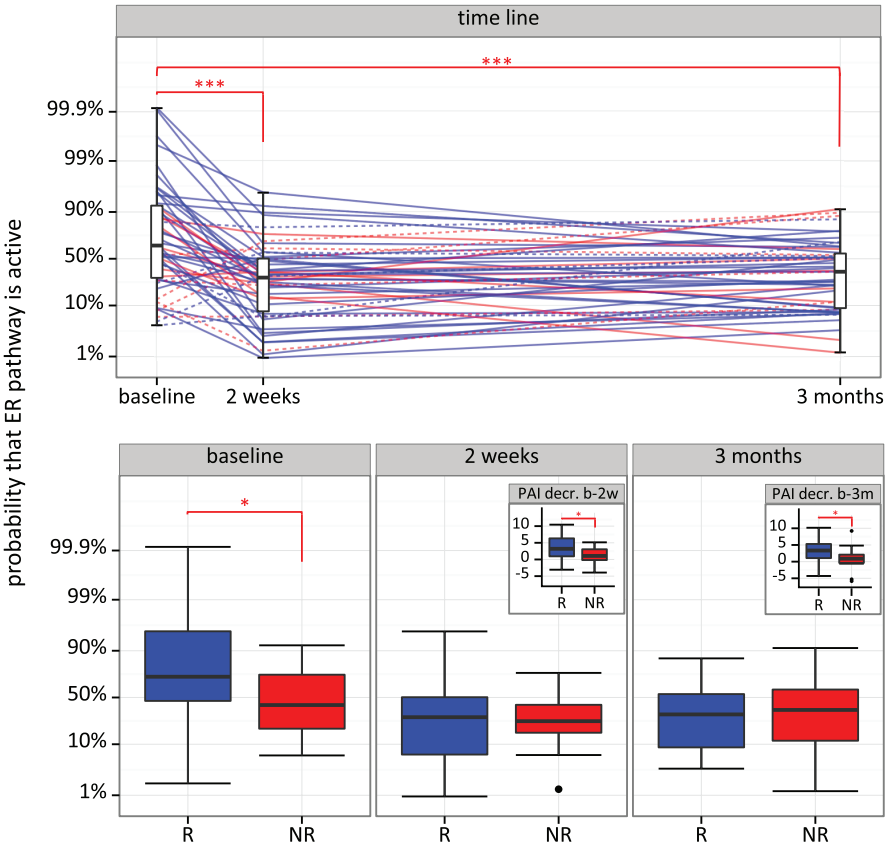


Figure 1: ER-pathway activity as a function of treatment time for GSE20181 dataset. Upper: line plot of activity per sample. Lower: box plots of activity in responders and non-responders per time point. Blue: responders, red: non-responders. Solid lines: activity went down during treatment, dashed lines: activity went up during treatment.

Subsequently we analyzed the tissue samples from the TEAM IIA clinical trial, using the RT-qPCR ER pathway model.

TEAM IIA cohort characteristics

The TEAM IIA cohort consists of 102 patients, all of which received at least 3 months of neoadjuvant endocrine therapy with exemestane. The majority of patients (n=83) received 6 months of therapy. Both biopsies and resection samples were collected and analyzed retrospectively. Thirty-three biopsies and 38 resection samples could not be retrieved, leading to a cohort of 69 patients from whom a biopsy sample was available and a cohort of 64 patients from whom a resection specimen was available. During LCM another 16 resection samples and 11 biopsies were excluded, mainly due to absence of tumor tissue in the specimen. Upon subsequent RT-qPCR, 9 biopsy samples yielded amounts of RNA too low to perform the required set of qPCRs, making them ineligible for further analysis, resulting in 49 biopsy and 48 resection samples eligible for analysis, of which 28 samples were matched cases from the same patient (Figure 2).

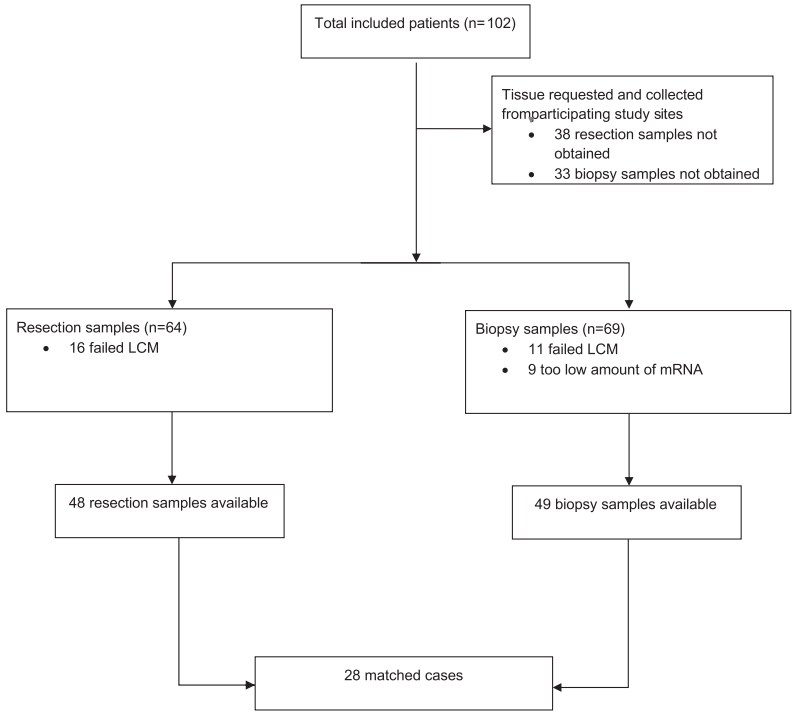


Figure 2 : consort diagram of tissue amples collected in the TEAM IIA trial

Baseline and post treatment ER pathway activity TEAM IIA

At baseline, the average probability of the ER pathway being active was 0.48, ranging from 0.02 to 0.89. After therapy, the average ER activity decreased to 0.32, ranging from 0.01 to 0.86 (Table 2). Of the 28 matched cases, the mean activity of ER decreased by 35% (absolute difference of -0.16), ranging from a decrease of 0.70 to an increase of 0.31 (mean PAI decrease from -0.3 to -1.8, paired t-test p-value <0.001).

ER pathway activity at baseline and decrease in activity during therapy were correlated with clinicopathological baseline characteristics (Table 3). Baseline mean ER pathway activity was significantly higher in PR positive samples (P=0.55) than PR negative samples (P=0.29), with PAI=0.3 vs. -1.8, two sample t-test p-value=0.003; and in patients with higher BMI (P=0.53) than in patients with lower BMI (P=0.38), with PAI=0.2 vs. -1, p-value=0.04. There was no significant difference in baseline activity between age categories, HER2-status, tumor type, tumor grade and tumor or nodal status.

Interestingly, the mean decrease in ER pathway activity in the available paired cases was neither significantly higher in baseline PR postive patients than in baseline PR negativepatients nor in patients with higher baseline BMI than in patients with lower baseline BMI (Table 3).

Correlation with TEAM IIA outcome assessed by palpation

To test the predictive value of the ER pathway model, the computed ER pathway activities were correlated to primary and secondary outcome measures of the TEAM IIA trial. The primary outcome of the trial was response to therapy based on palpation. Analysis showed that, at the end of therapy, patients with a palpation-based progressive disease (PD) had a significant lower ER pathway activity at baseline (P=0.16) compared to patients with complete remission (CR; P=0.59), partial remission (PR; P=0.48) or stable disease (SD; P=0.53); mean PAI PD=-2.5 vs non-PD=0.1, p-value=0.03 (Figure 3). Similar prognostic effects of ER pathway activity were shown at three months, as a measure of early response; p-value=0.038.

Table 2: ER pathway at baseline and resection for the TEAMII A cohort. ER pathway activity is presented as probability that the ER pathway is active (P) and as pathway activity index (PAI). CR: complete remission, PR: partial remission, SD: stable disease, PD: progressive disease; av: average, sd: standard deviation.

	Baseline							Resection						
	N	P			PAI			N	P			PAI		
		Mean	sd	Range	Mean	sd	Range		Mean	sd	Range	Mean	sd	Range
Total	49	0.48	0.25	[0.02, 0.89]	-0.26	1.9	[-5.6, 3.1]	50	0.31	0.21	[0.01, 0.86]	-1.65	2.3	[-6.3, 2.7]
Stratified by response assessed by palpation at last measurement														
CR	11	0.59	0.19	[0.36, 0.89]	0.63	1.3	[-0.8, 3]	9	0.20	0.18	[0.03, 0.57]	-2.5	1.7	[-5,0.38]
PR	12	0.48	0.28	[0.06, 0.81]	-0.32	2.1	[-4.1,2.1]	15	0.30	0.30	[0.01,0.86]	-2.0	2.7	[-6.3, 2.6]
SD	12	0.52	0.23	[0.11,0.80]	0.13	1.5	[-3,2]	13	0.40	0.23	[0.06, 0.77]	-0.8	1.7	[-3.9,1.8]
PD	3	0.16	0.11	[0.09, 0.29]	-2.50	1.1	[-3.3.-1.3]	3	0.08	0.05	[0.03, 0.12]	-3.8	1.2	[-5.2,-2.9]
Stratified by response assessed by mammography at last measurement														
CR/PR	0/11	0.50	0.31	[0.07,0.89]	-0.12	2.3	[-3.7,3]	1/9	0.29	0.26	[0.01,0.86]	-1.8	2.4	[-6.3, 2.6]
SD/PD	12/1	0.47	0.25	[0.6,0.8]	-0.28	1.8	[-4.1,2]	11/1	0.28	0.26	[0.3,0.73]	-2.0	2.2	[-5.2,1.4]

Table 3 Association of baseline parameters of the TEAM II-A cohort with ER pathway activity at biopsy and its decrease after therapy (difference between baseline biopsy and post-therapy resection sample; positive values represents a decrease in activity during therapy). ER pathway activity is presented as probability that ER is active (P) and as pathway activity index (PAI). %*: percentage excluding missing; sd: standard deviation; p-value ANOVA test

Total N=102			Baseline N=49							Decrease baseline - resection N=28						
	N	(%*)	N	(%*)	P Mean	Range	PAI Mean	Range	p-value	N	(%*)	P Mean	Range	PAI Mean	[range]	p-value
Histological grade									0.43							0.009
G1	12	(11.8)	8	(16.3)	0.37	[0.07,0.71]	-1.0	[-3.7,1.3]		2	(7.1)	0.21	[0.18,0.25]	2.9	[1.5,4.2]	
G2	27	(26.5)	13	(26.5)	0.50	[0.18,0.76]	0.0	[-2.2,1.6]		11	(39.3)	0.20	[-0.31,0.51]	0.0	[-2.2,1.5]	
G3	7	(6.9)	2	(4.1)	0.24	[0.11,0.37]	-1.9	[-3,-0.8]		2	(7.1)	-0.22	[-0.25,-0.20]	-1.9	[-3,-0.8]	
Gx	56	(54.9)	26	(53.1)	0.52	[0.02,0.89]	0.0	[-5.6,3]		13	(46.4)	0.20	[-0.20,0.70]	-0.3	[-4.1,2.4]	

Clinical nodal status								0.81						0.92
No	77	(75.5)	34	(69.4)	0.48	[0.02,0.84]	-0.3	[-5.6,2.4]	21	(75)	0.17	[-0.2,0.7]	1.5	[-1.9,5.2]
N1-3	25	(24.5)	15	(30.6)	0.48	[0.07,0.89]	-0.2	[-3.7,3]	7	(25)	0.14	[-0.31,0.47]	1.6	[-1.2,4.3]
Age category								0.32						0.45
50-59 y	8	(7.8)	4	(8.2)	0.27	[0.06,0.66]	-1.9	[-4.1,1]	3	(10.7)	0.14	[0,0.4]	1.2	[0,2.4]
60-69 y	32	(31.4)	13	(26.5)	0.47	[0.02,0.8]	-0.4	[-5.6,2]	5	(17.9)	0.01	[-0.25,0.51]	0.3	[-1.5,3.9]
≥ 70 y	62	(60.8)	32	(65.3)	0.51	[0.07,0.89]	0.0	[-3.7,3]	20	(71.4)	0.21	[-0.31,0.7]	1.9	[-1.9,5.2]
Tumor type								0.19						**
Ductal	66	(66.7)	34	(72.3)	0.50	[0.06,0.89]	-0.1	[-4.1,3]	21	(77.8)	0.18	[-0.31,0.7]	1.5	[-1.9,5.2]
Lobular	30	(30.3)	11	(23.4)	0.41	[0.02,0.8]	-0.9	[-5.6,2]	5	(18.5)	0.17	[0.03,0.34]	2.3	[0.2,4.3]
Other	3	(3.0)	2	(4.3)	0.72	[0.63,0.81]	1.4	[0.8,2.1]	1	(3.7)	0.16	-	1.2	-
PR status								0.003						0.42
Negative	32	(31.4)	13	(26.5)	0.29	[0.02,0.63]	-1.8	[-5.6,0.8]	6	(21.4)	0.06	[-0.2,0.47]	0.8	[-1.8,4.3]
Positive	70	(68.6)	36	(73.5)	0.55	[0.09,0.89]	0.3	[-3.3,3]	22	(78.6)	0.19	[-0.31,0.7]	1.7	[-1.9,5.2]
HER2 (IHC/FISH)								0.67						**
Negative	85	(88.5)	45	(91.8)	0.49	[0.02,0.89]	-0.2	[-5.6,3]	26	(92.9)	0.18	[-0.31,0.66]	1.7	[-1.9,5.2]
Positive	9	(9.4)	2	(4.1)	0.26	[0.11,0.41]	-1.8	[-3,-0.5]	1	(3.6)	-0.2	-	-1.8	-
Undetermined	2	(2.1)	2	(4.1)	0.46	[0.25,0.68]	-0.3	[-1.6,1.1]	1	(3.6)	0.03	-	0.2	-
Tumor size								0.97						0.41
< 3 cm	43	(43.4)	21	(44.7)	0.48	[0.02,0.89]	-0.3	[-5.6,3]	9	(33.3)	0.08	[-0.31,0.66]	0.7	[-1.9,4.6]
3-5 cm	41	(41.4)	21	(44.7)	0.48	[0.07,0.84]	-0.2	[-3.7,2.4]	16	(59.3)	0.23	[-0.2,0.7]	2.0	[-1.5,5.2]
> 5 cm	15	(15.2)	5	(10.6)	0.48	[0.19,0.8]	-0.1	[-2.1,2]	2	(7.4)	0.22	[0.18,0.27]	2.9	[1.6,4.2]
BMI								0.04						0.46
< 25	33	(32.4)	19	(38.8)	0.39	[0.02,0.81]	-1.0	[-5.6,2.1]	12	(42.9)	0.09	[-0.31,0.7]	1.2	[-1.9,5.2]
> 25	69	(67.6)	30	(61.2)	0.54	[0.09,0.89]	0.2	[-3.3,3]	16	(57.1)	0.22	[-0.25,0.66]	1.8	[-1.5,4.6]

** not enough sample to perform statistical analysis; ≠mean differs from other categories

ER pathway activity decreased in all six patients with complete remission, while 29% (two of seven) patients with partial remission and 38% (three of eight) of patients with stable disease showed an increase in ER activity. Most notably, the decrease in ER pathway activity was significantly higher in the combined CR/PR/SD group than in the PD group, whereas ER activity did not change significantly in the two patients with progressive disease (mean decrease in probability = 0.18 vs -0.002; PAI = -0.1 vs -3.2, paired t-test p-value=0.003) (Table 4). These observations are in accordance with the hypothesis that hormone therapy is only effective in patients with an active ER pathway and that success of response is associated with a decrease in ER pathway activity during treatment.

Table 4: Welsh two samples t-test for mean ER activity and mean decrease in ER activity during treatment for the TEAMIA cohort. ER pathway activity is presented as probability that the ER pathway is active (P) and as pathway activity index (PAI). Evaluation is performed on baseline biopsies, post-treatment resection samples, and the difference between those timepoints. The response is evaluated by palpation and mammography at 3 months and at last therapy. CR: complete remission, PR: partial remission, SD: stable disease, PD: progressive disease; p-values were computed based on PAI data.

	Group	N	P mean (sd)	PAI mean (sd)	Group	N	P mean (sd)	PAI mean (sd)	p-value
Stratified by response assessed by palpation at 3 months									
Baseline	CR/PR/SD	27	0.53 (0.23)	0.14 (1.7)	PD	3	0.17 (0.12)	-2.5 (1.2)	0.04
Stratified by response assessed by palpation at last measurement									
Baseline	CR/PR/SD	35	0.53 (0.11)	0.13 (1.7)	PD	3	0.16 (0.11)	-2.5 (1.1)	0.03
Resection	CR/PR/SD	37	0.31 (0.26)	-1.7 (2.2)	PD	3	0.08 (0.5)	-3.8 (1.2)	0.07
Baseline-Resection	CR/PR/SD	21	0.18 (2.28)	1.6 (2.2)	PD	2	-0.002 (0.003)	-0.02 (0.05)	0.003
Stratified by response assessed by mammography at 3 months									
Baseline	CR/PR	6	0.71 (0.17)	1.5 (1.2)	SD/PD	12	0.44 (0.24)	-0.49 (1.7)	0.015
Stratified by response assessed by mammography at last measurement									
Baseline	CR/PR	11	0.5 (0.31)	-0.12 (2.3)	SD/PD	13	0.48 (0.25)	-0.28 (1.8)	0.86
Resection	CR/PR	10	0.29 (0.26)	-1.8 (2.4)	SD/PD	12	0.28 (0.26)	-2 (2.2)	0.85
Baseline-Resection	CR/PR	6	0.18 (0.21	1.7 (2.4)	SD/PD	7	0.21 (0.34)	1.7 (2.4)	0.99

Correlation with TEAM IIA outcome assessed by mammography

Since response based on mammography served as a secondary outcome in the TEAM IIA trial, mammography was not mandatory for every patient. Due to the lower number of available observations in the CR and PD categories, CR and PR were combined into a category of responders, whereas SD and PD were combined into a category of non-responders. At three months, responders and non-responders could be clearly separated, based on the ER activity at baseline (responders P=0.71 vs non-responders P=0.44; PAI=1.5 vs 0.5, two samples t-test p-value=0.015). At 6 months

however, no distinction could be made (responders P=0.5 vs non-responders P=0.48; PAI=-0.1 vs -0.2, p-value=0.9).

Clinical validation

In summary, we performed an ER pathway analysis in two distinct AI neoadjuvant settings: the first, a proof of principle, public Affymetrix dataset derived from a cohort of patients undergoing letrozole AI therapy for three months; the second, a trial cohort of patients that underwent 3 to 6 months exemestane AI therapy. Both cohorts where assessed using adapted ER pathway models based on the original Affymetrix ER pathway model [16]. While the proof-of-concept dataset analysis used Affymetrix HGU133A gene expression data from fresh-frozen tissue, the TEAM IIA trial cohort analysis used RT-qPCR data from FFPE tissue obtained by LCM. Furthermore the two models used a different number target genes (27 vs. 12) and were calibrated using data from MCF7 cell cultures stimulated different estradiol concentration (6 nM vs. 1 nM).

Despite the differences between cohorts and methods, both analyses detected a significant decrease in ER pathway activity during AI treatment and indicated that the decrease in activity, as well as baseline ER pathway activity were significantly higher in responders than in non-responders, though response was assessed differently in the two studies. The proof-of-concept study defined response as a reduction in tumor size of at least 50% measured by ultrasound, while the TEAMII A cohort adopted the RECIST 1.1 criteria (at least 30% tumor reduction) using response measured by palpation as primary outcome and measured using radiological modalities as secondary outcomes.

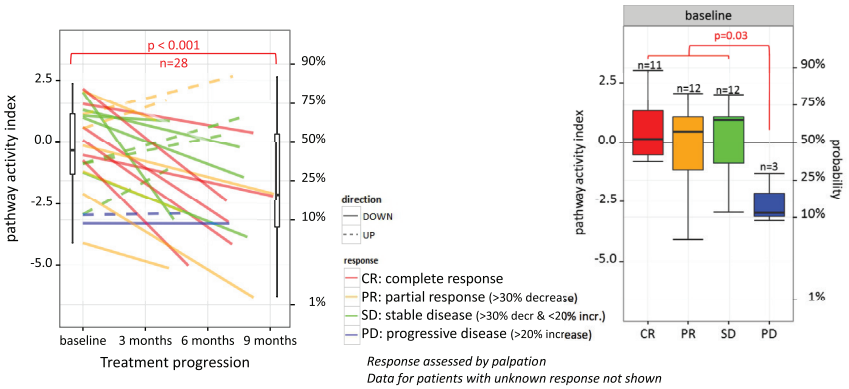


Figure 3: ER-pathway activity in patients of the TEAMIIA cohort stratified by response assessed by palpation at last measurement. Left: Line plots of ER-pathway activity at baseline and resection. Right: box plots of ER activity at baseline. Red: complete remission, orange: partial remission, green: stable disease, blue: progressive disease. Solid lines: activity went down during treatment, dashed lines activity went up during treatment.

Discussion

Duration of treatment and resistance

In the GSE20181 cohort, a significant difference in ER pathway activity between baseline and both two weeks and three months of letrozole therapy, but not between two weeks and three months, indicating an early and maintained response. A remarkable finding is the observation that baseline ER pathway activity was predictive for therapy response as measured by mammography at three months, but not at six months of therapy. This difference may be explained by mechanisms of treatment resistance. During the course of endocrine therapy, over 20% of breast cancers are known to acquire resistance against aromatase inhibitor treatment due to activating mutations in the estrogen receptor, resulting in reactivation of the ER pathway driving tumor growth. Alternatively other signal transduction pathways, like the PI3K pathway may take the lead in driving tumor growth.[11]

Another reason for loss of treatment effect during prolonged therapy, could be lower compliance at the end of therapy, resulting in reactivation of the ER pathway and the subsequent regrowth of the tumor. Although data on compliance are not available in this trial, it seems unlikely that non-compliance has occurred since patients were monitored closely during the trial.

Correlation with baseline parameters

The significantly higher baseline ER pathway activity in PR positive cases in the TEAM IIA cohort was expected, since PGR, the gene that codes for PR, is a target gene of the ER pathway. However, the decrease in ER pathway activity was not significantly higher in PR positive cases and the lack of correlation between baseline PR status and therapy response indicates that PR status by itself it is not sufficient to accurately infer pathway activation, or its deactivation in connection with AI therapy.

Study limitations

Drawbacks of this study are the retrospective setting, the limited availability of sample with sufficient quality, and the lack of a reliable read-out for the primary and secondary clinical outcome, that is, reduction in tumor size. Although the entire TEAM IIA cohort comprises 102 patients, due to described logistical reasons only 28 paired cases were available for analysis, limiting statistical power. Palpation was the primary outcome in this trial, whereas tumor reduction assessed by radiological imaging modalities was a secondary outcome, and therefore not performed in all patients. In addition, it is generally accepted that radiological evaluation using MRI in particular, may not be optimal in terms of accuracy for evaluation of therapy response in neo-adjuvant treatment of ER positive breast cancer.[26, 27] In this specific cohort, a comparison of tumor size at resection versus its estimated size by palpation, mammography, US, and MRI suggested that palpation and mammography were the most accurate methods, while in many cases, various methods provided discordant results.[28] Therefore only evaluations based on palpation and mammography were included in this analysis.

Another point of concern is tumor heterogeneity. ER activity as measured in the biopsy setting may not be representative of the whole tumor, which could be a source of error in the analysis, and could explain why some patients with low baseline ER activity showed some response to treatment.

Future perspectives

Having comparable results from two distinct ER positive breast cancer cohorts, a validated Affymetrix gene expression dataset and a trial cohort with daily practice FFPE tissue samples for mRNA qPCR analysis, this study offers a robust clinical validation of the ER pathway model, developed by Verhaegh *et al*[16], both when the full spectrum of target genes is measured using Affymetrix microarray on fresh tissue,

and when the selected subset of target genes is measured using multiple qPCR assays on FFPE material.

Further prospective validation of the ER pathway model is necessary before this approach can be used in a routine clinical setting to predict endocrine therapy response. A prospective study to validate the use of the ER pathway test to predict response to neo-adjuvant endocrine therapy is being planned as a side study in a CDK4/6 inhibitor clinical trial. This side study will investigate to which extent the ER pathway is capable of defining which patient benefits most from endocrine therapy alone, which patient should be offered a combination therapy with targeted agents such as CDK4/6 inhibitors, or which patient should be offered chemotherapy directly.

Upon extended clinical validation, our findings could lead to a new diagnostic test to quantify ER pathway activity in a cancer tissue sample, with clearly defined clinical utility, to be implemented in stratification of ER positive breast cancer patients. For example, some studies suggest that neo-adjuvant hormonal therapy is a good alternative for chemotherapy for specific forms of breast cancer.[29, 30] Measuring ER pathway activity could be used to help decide on treatment course. Low baseline ER pathway activity wouldindicatea low chance of response to hormonal therapy, in which case chemotherapy could be indicated. In addition, in case hormonal therapy is selected, a second biopsy after one month of therapy to determine the remaining ER pathway activity could be implemented to assess therapy response. Persistent high levels of ER pathway activity could indicate resistance to therapy or non-compliance. This would lead to a tailored neo-adjuvant treatment with improved response monitoring, and a stepping stone towards personalized management of breast cancer.

References

1. EBCTCG. Aromatase inhibitors versus tamoxifen in early breast cancer: patient-level meta-analysis of the randomised trials. The Lancet 2015;386(10001):1341-1352.
2. Davies C, Pan H, Godwin J et al. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. The Lancet 2009;381(9869):805-816.
3. Blok EJ, Derks MG, van der Hoeven JJ et al. Extended adjuvant endocrine therapy in hormone-receptor positive early breast cancer: Current and future evidence. Cancer Treat Rev 2015;41(3):271-276.
4. Harvey JM, Clark GM, Osborne CK et al. Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. J Clin Oncol 1999;17:1474-1481.
5. Hammond ME, Hayes DF, Dowsett M et al. American society of clinical oncology/college of American pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J Clin Oncol 2010;28:2784-2795.
6. Senkus E, Kyriakides S, Ohno S et al. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology 2015;26(suppl 5):v8-v30.
7. NABON. Richtlijn Mammacarcinoom. 2012. 23-12-2015. Ref Type: Internet Communication
8. Yi M, Huo L, Koenig KB et al. Which threshold for ER positivity? a retrospective study based on 9639 patients. Ann Oncol 2014;25(5):1004-1011.
9. Harbeck N, Rody A. Lost in Translation? Estrogen Receptor Status and Endocrine Responsiveness in Breast Cancer. JCO 2012;30(7):686-689.
10. Jeselsohn R, Buchwalter G, De Angelis C et al. ESR1 mutations[mdash]a mechanism for acquired endocrine resistance in breast cancer. Nat Rev Clin Oncol 2015;12(10):573-583.
11. Ciruelos Gil EM. Targeting the PI3K/AKT/mTOR pathway in estrogen receptor-positive breast cancer. Cancer Treatment Reviews 2014;40(7):862-871.
12. Renoir JM, Marsaud V+, Lazennec G. Estrogen receptor signaling as a target for novel breast cancer therapeutics. Biochemical Pharmacology 2013;85(4):449-465.
13. Gross JM, Yee D. How does the estrogen receptor work? Breast Cancer Res 2002;4(2):62-64.
14. Pietras RJ, Marquez-Garban DC. Membrane-Associated Estrogen Receptor Signaling Pathways in Human Cancers. Clinical Cancer Research 2007;13(16):4672-4676.
15. Frasor J, Danes JM, Komm B et al. Profiling of Estrogen Up- and Down-Regulated Gene Expression in Human Breast Cancer Cells: Insights into Gene Networks and Pathways Underlying Estrogenic Control of Proliferation and Cell Phenotype. Endocrinology 2003;144(10):4562-4574.
16. Verhaegh W, van Ooijen H, Inda MA et al. Selection of Personalized Patient Therapy through the Use of Knowledge-Based Computational Models That Identify Tumor-Driving Signal Transduction Pathways. Cancer Research 2014;74(11):2936-2945.
17. Verhaegh W, de Stolpe AV. Knowledge-based computational models. Oncotarget; Vol 5, No 14 2014.
18. Miller WR, Larionov A. Changes in expression of oestrogen regulated and proliferation genes with neoadjuvant treatment highlight heterogeneity of clinical resistance to the aromatase inhibitor, letrozole. Breast Cancer Research 2010;12(4):1-9.
19. Miller WR, Larionov A, Anderson TJ et al. Sequential changes in gene expression profiles in breast cancers during treatment with the aromatase inhibitor, letrozole. Pharmacogenomics J 2012;12(1):10-21.
20. McCall MN, Bolstad BM, Irizarry RA. Frozen robust multiarray analysis (fRMA). Biostatistics 2010;11(2):242-253.
21. Fontein DBY, Charehbili A, Nortier JWR et al. Efficacy of six month neoadjuvant endocrine therapy in postmenopausal, hormone receptor-positive breast cancer patients: a phase II trial. European Journal of Cancer 2014;50(13):2190-2200.
22. Espina V, Wulfkuhle JD, Calvert VS et al. Laser-capture microdissection. Nat Protocols 2006;1(2):586-603.
23. Van de Wiel P, Verhaegh W, Alves de Inda M et al. 514: Assessing functional ER pathway activity using a computational pathway model. European Journal of Cancer50:S124.

24. Chang EC, Charn TH, Park SH et al. Estrogen Receptors a and b as Determinants of Gene Expression: Influence of Ligand, Dose, and Chromatin Binding. Molecular Endocrinology 2008;22(5):1032-1043.

25. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/. 2014.

Ref Type: Internet Communication

26. Charehbili A, Wasser MN, Smit VTHB et al. Accuracy of MRI for treatment response assessment after taxane- and anthracycline-based neoadjuvant chemotherapy in HER2-negative breast cancer. European Journal of Surgical Oncology (EJSO) 2014;40(10):1216-1221.

27. Chen JH, Feig B, Agrawal G et al. MRI evaluation of pathologically complete response and residual tumors in breast cancer after neoadjuvant chemotherapy. Cancer 2008;112(1):17-26.

28. Blok EJ, Charehbili A, Kroep JR et al. Abstract P6-01-05: Radiological evaluation of neo-adjuvant endocrine therapy in hormone-receptor positive early breast cancer. Cancer Research 2016;76(4 Supplement):6-01.

29. Purushotham A, Pinder S, Cariati M et al. Neoadjuvant Chemotherapy: Not the Best Option in Estrogen Receptor-Positive, HER2-Negative, Invasive Classical Lobular Carcinoma of the Breast? JCO 2010;28(22):3552-3554.

30. Dixon JM, Renshaw L, Dixon J et al. Invasive lobular carcinoma: response to neoadjuvant letrozole therapy. Breast Cancer Res Treat 2011;130(3):871-877.