Aldehyde dehydrogenase 1A1 expression in breast cancer is associated with stage, triple negativity, and outcome to neoadjuvant chemotherapy

Thaer Khoury[1], Foluso O Ademuyiwa[2], Rameela Chandraseekhar[3], Marah Jabbour[4], Albert DeLeo[5], Soldano Ferrone[5], Yangyang Wang[5] and Xinhui Wang[5]

[1]Department of Pathology, Roswell Park Cancer Institute, Buffalo, NY, USA; [2]Department of Medicine, Roswell Park Cancer Institute, Buffalo, NY, USA; [3]Department of Biostatistics, Roswell Park Cancer Institute, Buffalo, NY, USA; [4]Buffalo, NY, USA and [5]Department of Immunology, University of Pittsburgh Medical Center, Pittsburgh, PA, USA

Studies have shown that ALDH1A1 expression in the breast is associated with worse clinical outcome. ALDH1A1 inactivates cyclophosphamide, which is an integral agent in breast cancer chemotherapy regimens. The purposes of this study were to verify these results, to correlate ALDH1A1 expression with clinical outcome in patients treated with cyclophosphamide as part of the chemotherapy (adjuvant or neoadjuvant), and to evaluate ALDH1A1 as a useful marker to predict the clinical outcome of breast cancer subsets. A total of 513 primary breast cancers were studied. Tissue microarrays of the studied cases were stained with ALDH1A1. Key clinicopathological information was obtained. Disease-free survival and overall survival were calculated. Patients with neoadjuvant therapy who had substantial residual cancer burden (RCB) were included in the study. Fisher's exact test and Kaplan–Meier methods were used for statistical analysis. ALDH1A1 was expressed in 53 (10%) patients, with a higher frequency in triple negative, followed by HER2 +, and finally hormonal receptor +/HER2− ($P < 0.0001$). Tumors with advanced stage, node-positive, or larger tumor size were correlated with ALDH1A1 expression ($P = 0.006$, $P < 0.0001$, and $P = 0.05$, respectively). ALDH1A1 expression was also correlated with worse disease-free survival ($P < 0.006$) and overall survival ($P < 0.01$) in patients who were treated with neoadjuvant chemotherapy. In all, 8 of 22 (36%) received neoadjuvant chemotherapy and died of disease-expressed ALDH1A1 ($P = 0.008$). Similarly, 8 of 23 (35%) who received neoadjuvant chemotherapy and had tumor recurrence expressed this marker ($P = 0.002$). The risk of recurrence was fivefold greater than negative ALDH1A1 tumors. The risk of recurrence became 11-fold greater when cyclophosphamide but not trastuzumab was part of the regimen. Our results are consistent with previous studies. Moreover, we found that ALDH1A1 could be a useful marker to predict worse clinical outcome after chemotherapy in the neoadjuvant setting with substantial RCB. However, a larger cohort is required to verify our results.

Breast cancer is not a single disease. It has multiple histological subtypes or entities (ductal, lobular, etc).[1] More recently, based on the molecular profiling, breast cancer has been divided into three subtypes, namely basal-like, luminal, and HER2 +.[2,3] Even

within each group, there is considerable variation in clinical outcome and response to therapy.

One of the suggested reasons for this variation and therapy resistance is cancer stem cells. There is increasing evidence that human breast cancers are driven by a tumor-initiating cancer stem cell component that may contribute to tumor metastasis and therapeutic resistance.[4–8] They are characterized as CD44 +/CD24− and/or aldehyde dehydrogenase1A1 (ALDH1A1)-positive cells.[9] Ginestier *et al*[10] found that ALDH1A1 is a better marker of breast cancer stem cells as fewer ALDH1A1-positive tumor cells

ALDH1A1 in breast cancer
T Khoury *et al*
389

than CD44 + /CD24− tumor cells are required to produce tumors in immunodeficient mice. However, Marcato *et al*[11] found ALDH1A3 and not ALDH1A1 to be the primary contributor of ALDEFLUOR activity in breast cancer stem cells. Nonetheless, ALDH1A1 expression is found to correlate with high histological grade, estrogen and progesterone receptor negativity, and HER2 positivity with higher relative risk of poor overall survival.[4–6,9,10,12–18]

ALDHs are a family of NAD(P)$^+$-dependent enzymes involved in detoxifying a wide variety of aldehydes to their corresponding weak carboxylic acids.[19] ALDH1 mainly functions as a retinoic acid enzyme, catalyzing the conversion of vitamin A (retinol) to retinoic acid. ALDH1A1 and ALDH2 participate in alcohol metabolism.[20] It was found that the high expression of ALDH1A1 in a tumor may provide a route for tumors to resist chemotherapy, particularly cyclophosphamide.[21–23] It is known that the activation of cyclophosphamide to 4-hydroxycyclophosphamide is catalyzed by hepatic cytochromes. 4-Hydroxycyclophosphamide rapidly interconverts with its tautomer aldophosphamide. Both of them leave hepatic cells and enter other cells. Aldophosphamide yields phosphoramide mustard and acrolein through non-enzymatic elimination reaction. Phosphoramide mustard is believed to be the important metabolite for the therapeutic effect of cyclophosphamide. ALDH1A1, with lesser extent ALDH3A1 and ALDH5A1, leads to major detoxification of aldophosphamide to inactive the metabolite carboxyphosphamide instead of phosphoramide mustard.[24]

There have been multiple studies linking ALDH1A1 positivity to clinical outcome and breast cancer phenotypes.[4–6,9,10,12–18] However, little is known about the prognostic significance of this marker in breast cancer subsets, or response to systemic or radiation therapy.

Materials and methods

Patients

A total of 513 breast cancer patients were identified from the files of Roswell Park Cancer Institute between 1995 and 2007. Clinicopathological information obtained included patients' age at diagnosis, menopausal status, race, tumor size, modified SBR grade, histological type, node status, stage, biomarkers status (hormonal receptors and HER2), therapy modality (adjuvant and neoadjuvant), hormonal treatments, and receipt of radiation. We also obtained survival data including date of death or last follow-up for calculating disease-free survival and overall survival.

Calculating the Residual Cancer Burden in Neoadjuvant-Treated Breast Excisions

Measuring the extent of residual cancer burden disease (RCB) was calculated adopting a formula proposed by Symmans *et al.*[25] Pathological stage

after treatment was determined using the revised American Joint Committee on Cancer staging system for breast cancer.[26] An index of 3.28 was considered extensive disease (RCB-III), whereas an index of 1.36 was considered minimal disease (RCB-I), and an index between these two was considered intermediate (RCB-II).

Tissue Microarray, ALDH1A1 Immunohistochemistry, and Scoring

For each case, at least two core samples of tumor tissue were acquired from at least two different donor blocks. In all, 15 tissue microarrays were assembled using a Beecher tissue puncher and array system (Beecher Instruments, Silver Spring, MD). For each block, four 0.6-mm core samples were acquired and assembled in the receiving tissue microarray block. One hematoxylin and eosin slide from each block was prepared and reviewed to confirm the presence of the invasive tumor.

ALDH1A1 staining was carried out using a rabbit monoclonal antibody; clone EP1933Y was obtained from Abcam Company (Cambridge, MA) with dilution 1/100. Sides were microwaved for 20 min with target retrieval solution (Dako, Carpinteria, CA) for antigen retrieval. Known positive (liver) and negative (skin) controls were used. Slides were cooled to room temperature, deparaffinized in three changes of xylene, and rehydrated using graded alcohols. Endogenous peroxidase was quenched twice with aqueous 3% H_2O_2 for 10 min each time and washed with PBS/T. For antigen retrieval, slides were heated in the microwave for 20 min in target retrieval solution (Dako S1699), followed by a 15 min cool down and a PBS/T wash. Slides were then loaded on the Dako Autostainer and blocked for 30 min with 0.03% Casein (in PBS/T). Casein was blown off and the primary rabbit monoclonal antibody ALDH1A1 (Abcam no. ab52492) was applied at a concentration of 1.1 μg/ml for 1 h. An isotype-matched control (1.1 μg/ml Rab IgG) was used on a duplicate slide in place of the primary antibody as a negative control. A PBS/T wash was followed by anti-mouse Envision + reagent (Dako) for 30 min. PBS/T was used as a wash and the chromagen DAB + (Dako) was applied for 10 min. After immunohistochemical staining, slides were counterstained with hematoxylin, dehydrated, cleared, and coverslipped.

The ALDH1A1 score was recorded as 3 + ($\geq$50% positive tumor cells), 2 + (<50% but $\geq$10%), 1 + (<10%), or negative (0%). Scores of 2 and 3 were considered positive.[10,17] Special attention was paid to cases that had a histiocytic reaction, as these cells cross-react with ALDH1A1. Therefore, all cores that had positive cells were compared with the hematoxylin and eosin sections. Cores that had a histiocytic reaction and dispersed positive cells were considered negative. Owing to the nature of the study (tissue microarray of 0.6-mm cores), ADH1A1 expression in the stroma was not evaluated.

Statistical Analysis

Statistical analyses for comparing groups with regard to categorical variables were performed using Fisher's exact test. Overall survival was defined as the time from the date of diagnosis to the date of death or the date of last follow-up. Disease-free survival was defined as the time from the date of diagnosis to the date of first recurrence, date of death, or date of last follow-up. Overall survival curves were generated using the Kaplan–Meier method and compared using the log-rank test. Proportional hazards model was used to compare survival as a function of patient and treatment factors. Values for categorical data are specified as frequency (percentage). Statistical analysis was performed using SAS Statistical Analysis Software Version 9.2 (SAS Institute, Cary, NC, USA). A nominal significance level of 0.05 was used.

Results

Patients

A total of 513 patients were included in the study. Chemotherapy was refused by the patient and/or was not given to 29 patients, whereas 18 had missing data. Of the remaining 466, 40 (9%) patients received chemotherapy in the neoadjuvant setting and 426 (91%) in the adjuvant setting. Adjuvant chemotherapy with hormonal therapy was given to 145 of 426 (34%) patients, adjuvant chemotherapy alone to 126 (23%), and adjuvant hormonal therapy alone to 155 (36%). Adjuvant radiation therapy was administered to 381 (74%) patients. Estrogen receptor was positive in 333 (65%), whereas HER2 was positive in 138 (27%). There were 218 (42%) stage 1, 239 (47%) stage 2, and 56 (11%) stage 3. A total of 132 (26%) patients died of disease, whereas 154 (30%) had local recurrence (Table 1).

Cyclophosphamide was part of the neoadjuvant therapy in 36 of 40 (90%) patients. Similarly, it was part of the adjuvant chemotherapy with or without hormonal therapy in 377 (89%) patients. When cyclophosphamide was part of the adjuvant therapy, it was given with methotrexate and fluorouracil, or with adriamycin with or without taxane.

ALDH1A1 Expression and Correlation with Clinicopathological Variables

ALDH1A1 was expressed in 53 of 513 cases (10%) (Figure 1). In these 53 cases, 24 (45%) were triple negative, 17 (32%) HER2 +, and 12 (23%) hormonal receptor + /HER2 − ($P < 0.0001$). ALDH1A1 was found most frequently in stage 3 disease (10 of 56 cases, 18%), followed by stage 2 (32 of 207 cases, 13%), and stage 1 (11 of 207 cases, 5%) ($P = 0.0006$). Tumors with positive lymph nodes expressed ALDH1A1 more frequently than did node-negative tumors ($P = 0.006$). There was no significant correlation

Table 1 Univariate analysis between ALDH1A1 and clinicopathological parameters of studied cases

Variables	ALDH1A1		P-value
	Negative (n = 460)	*Positive* (n = 53)	
Demographics			
Race			
Black	67 (14)	6 (11)	NS
Caucasian	381 (83)	45 (85)	
Others	12 (3)	2 (4)	
Menopausal status			
Post	275 (65)	33 (67)	NS
Pre	146 (35)	16 (33)	
Age (years)			
>50	294 (64)	29 (55)	NS
≤50	166 (36)	24 (45)	
Pathology			
mSBR grade			
I	23 (5)	1 (2)	NS
II	97 (21)	9 (18)	
III	333 (74)	40 (80)	
ER			
Negative	145 (32)	34 (64)	<0.0001
Positive	314 (68)	19 (36)	
PR			
Negative	222 (48)	40 (75)	0.0002
Positive	236 (52)	13 (25)	
HER2			
Negative	337 (74)	36 (68)	NS
Positive	121 (26)	17 (32)	
Histological type			
Ductal	428 (93)	48 (91)	NS
Lobular	27 (6)	4 (7)	
Others	5 (1)	1 (2)	
Stage[a]			
1	207 (45)	11 (21)	0.001
2	207 (45)	32 (60)	
3	46 (10)	10 (19)	
Nodes			
Negative	262 (62)	21 (41)	0.006
Positive	162 (38)	30 (59)	
Tumor size (cm)[b]			
≤2.0	265 (59)	22 (44)	0.05
> 2.0	188 (41)	28 (56)	
Types			
Triple negative	95 (21)	24 (45)	<0.0001
HR+/HER2−	240 (53)	12 (23)	
HER2+	121 (27)	17 (32)	
Radiation			
Yes	348 (78)	33 (62)	NA
No	100 (22)	20 (38)	
Subsets by therapy modality			
Adjuvant—chemo and hormonal			
Yes	136 (35)	9 (25)	NA
No	255 (65)	27 (75)	
Adjuvant—chemo only			
Yes	106 (27)	20 (56)	NA
No	285 (73)	16 (44)	
Adjuvant—hormonal only			
Yes	148 (38)	7 (19)	NA
No	243 (62)	29 (81)	
Neoadjuvant—all			
Yes	31 (7)	9 (17)	NA
No	427 (93)	44 (83)	
Cyclophosphamide—neoadjuvant			
Yes	28 (11)	8 (24)	NA
No	218 (89)	26 (76)	
Cyclophosphamide—adjuvant chemo			
Yes	217 (100)	25 (100)	NA
No	0 (0)	0 (0)	
Cyclophosphamide—adjuvant chemo and hormonal			
Yes	128 (59)	7 (28)	NA
No	88 (41)	18 (72)	

NA, not applicable; NS, not significant.
In many of the variables, the total number of cases does not add up to 513 because of missing data.
[a]Stage IV cases were excluded from the analysis (five cases).
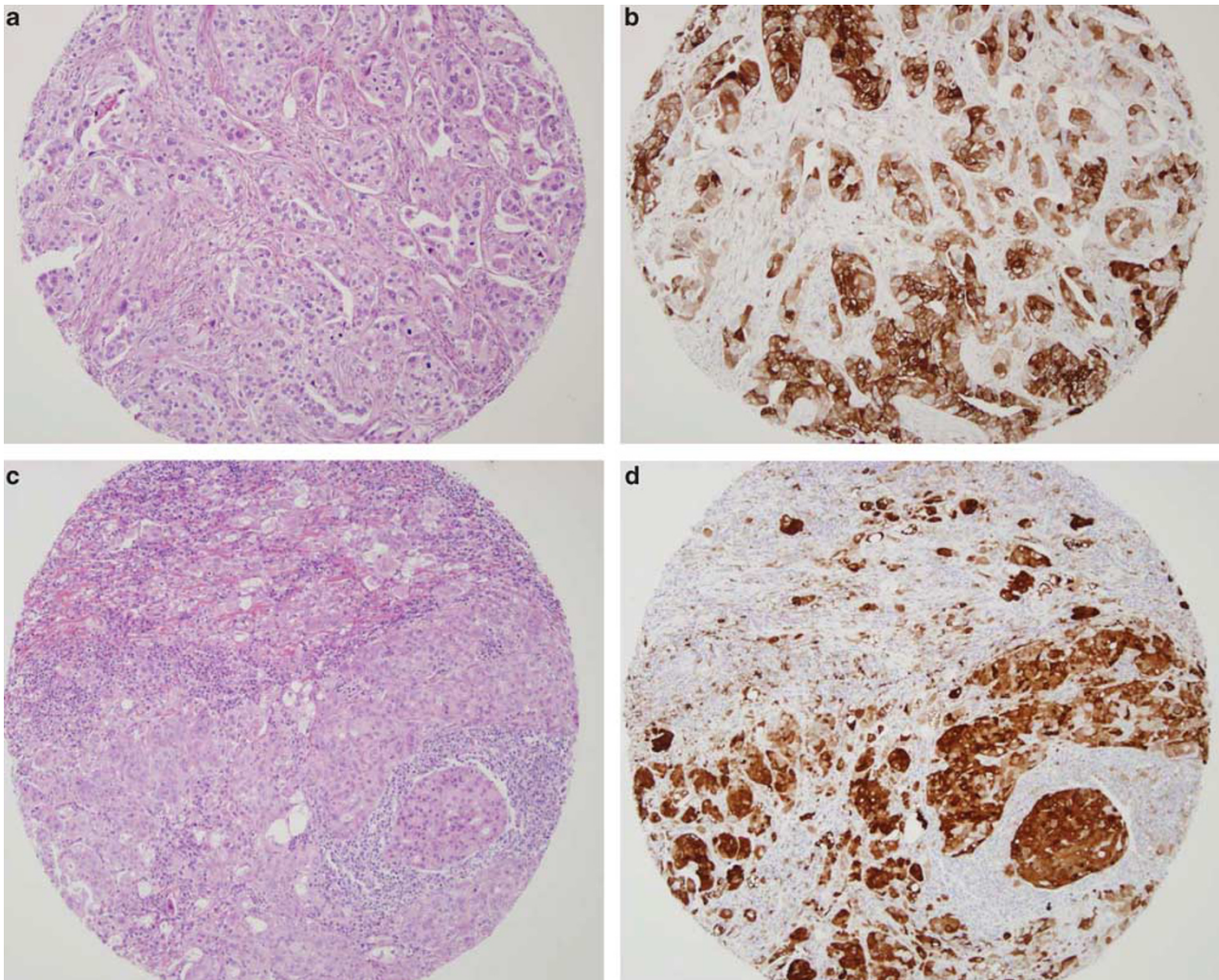[b]2.0 cm is the median.

Figure 1 ALDH1A1 with matching hematoxylin and eosin staining in breast cancer: (**a** and **b**) triple negative type cancer with 3+ staining (×10); (**c** and **d**) HER2+ type with 3+ staining (×10).

with patients' age, race, menopausal status, histological type, or mSBR grade. The rest of the results are presented in Table 1.

The multivariate odds ratio of a tumor being ALDH1A1+ for hormonal receptor+/HER2− subtype *vs* triple-negative tumors was 0.261 (95% CI; 0.120, 0.570). Similarly, for HER2 subtype *vs* triple-negative tumors, the odds ratio of a tumor being ALDH1A1+ was 0.517 (95% CI; 0.246, 1.083) ($P = 0.003$). Similarly, the odds of a tumor being ALDH1A1+ in stage 1 tumors was 0.346 (95% CI; 0.162, 0.738) compared with stage 2 or 3 ($P = 0.006$) (Table 2).

ALDH1A1 Correlation with Disease-Free and Overall Survival

All patients
Overall, ALDH1A1-positive cells were more often encountered in patients who died of breast cancer

(18 of 53 (34%) *vs* 114 of 460 (25%)) ($P = 0.04$) (Figure 2). There was no significant correlation between ALDH1A1 and disease-free survival (Table 3). ALDH1A1 expression was correlated with disease-free survival and overall survival for each tumor subtype (hormonal receptor+/HER2−, HER+ or triple negative). There was no significant correlation.

Neoadjuvant therapy subset
In all, 40 patients had neoadjuvant therapy; of those 38 had all slides to review (2 cases had tissue in the tissue microarray but no available slides to evaluate RCB). All 38 patients had either RCB-II ($n = 14$) or RCB-III ($n = 24$) based on the definition by Symmans *et al.*[25] Patients who had complete pathological response or RCB-I were excluded from the study, because of the absence or small residual tumor and not being able to construct in tissue microarrays. In all, 12 cases were hormone receptor+/HER2−, 8 HER2+, and 20 triple negative. The neoadjuvant

Table 2 Multivariate analysis to evaluate factors that predict ALDH1A1+

Variables	OR (95% CI)	P-value
Race		
Black	0.498 (0.191, 1.302)	NS
Others	1.092 (0.212, 5.631)	
Caucasian		
Menopausal status		
Post	2.169 (0.832, 5.655)	NS
Pre		
Age (years)		
≤50	1.843 (0.732, 4.643)	NS
>50		
Stage		
1	0.346 (0.162, 0.738)	0.006
2 and 3		
Status		
HER2+	0.517 (0.246, 1.083)	0.003
HR+/HER2−	0.261 (0.120, 0.570)	
Triple negative		

NS, not significant; HR, hormonal receptor.
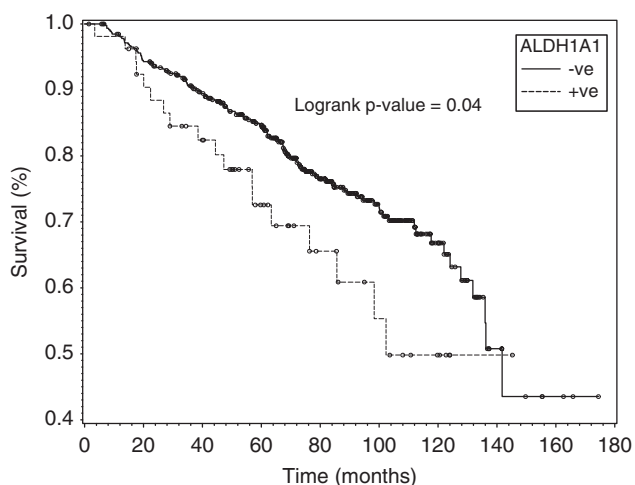
Figure 2 Kaplan–Meier curve showing ALDH1A1 correlation with overall survival in all patients.

therapy was composed of doxorubicin and cyclophosphamide with or without a taxane. Two patients had no cyclophosphamide as part of the chemotherapy. Three of eight patients who were HER2+ received trastuzumab.

ALDH1A1 was expressed more frequently in patients who developed local recurrence (8 of 9 (89%) *vs* 15 of 31 (48%), $P = 0.002$) or died of disease (8 of 9 (89%) *vs* 14 of 31 (45%), $P = 0.008$). When excluding patients who were treated with trastuzumab, the results remained significant. ALD-H1A1 was expressed more frequently in patients who developed local recurrence (7 of 8 (88%) *vs* 14

of 29 (48%), $P = 0.005$) or died of disease (7 of 8 (88%) *vs* 13 of 29 (45%) ($P = 0.01$)). The significance was similar after excluding two patients who were treated with chemotherapy without cyclophosphamide (Table 3, Figure 3).

In the multivariate analysis, a positive ALDH1A1 patient treated with neoadjuvant therapy when cyclophosphamide was part of the regimen without trastuzumab had more than 11 times higher risk of local recurrence than did a patient with negative ALDH1A1 tumor (95% CI; 2.13–62.86, $P = 0.005$). For overall survival, there was borderline significance of the hazard ratio with about 3 times higher risk of dying of disease with positive ALDH1A1 (95% CI; 0.85–10.93, $P = 0.09$) (Table 4). The risk of recurrence dropped to seven times and the risk of death became insignificant when all patients were included in the analysis (Table 5).

Variable subsets
The following analyses for subset population were performed: (1) tumor types, hormonal receptor+/HER2−, HER2+, and triple negative; (2) therapy modality, trastuzumab-based chemotherapy, adjuvant chemotherapy (with or without cyclophosphamide), and radiation therapy; (3) combining tumor types with therapy modality, like hormonal receptor+/HER2− type treated with chemotherapy, hormonal receptor+/HER2− type treated with radiation therapy, HER2+ type treated with chemotherapy (with and without trastuzumab), etc. No statistical significance was found. The results of adjuvant therapy with or without cyclophosphamide and with or without hormonal therapy are listed in Table 3. There were 25 patients treated with trastuzumab, 2 had local recurrence and 3 died of disease. These patients were followed up for at least 18.4 months. Given the small number of events, no statistical conclusion could be drawn.

Discussion

It is now known that ALDH1A3 and not ALDH1A1 is found to be expressed in breast cancer stem cells.[11] ALDH1A1 is an effective detoxifying enzyme. Its high expression can also provide a route for tumors to resist chemotherapy.[21–24] Moreover, multiple studies have shown that this marker correlates with worse clinical outcome. Therefore, we chose ALDH1A1 in this study to evaluate whether it can be used as a prognostic marker for various subsets of breast cancer, and response to systemic therapy.

In a meta-analysis that included 12 studies with 898 cases and 1853 controls, ALDH1A1 expression was significantly associated with high histological grade, estrogen and progesterone receptors negativity, and HER2 positivity. However, ALDH1A1 expression was not associated with tumor size or nodal status. The relative risk of poor overall

Table 3 Kaplan–Meier estimates for overall and disease-free survival for subset populations

Subset	ALDH1A1	OS			DFS		
		Dead N (%)	Alive N (%)	P-value	Yes N (%)	No N (%)	P-value
All patients							
With or without cyclophosphamide	+(n = 53)	18 (34)	35 (66)	0.04	18 (34)	35 (66)	NS
	−(n = 460)	114 (25)	346 (75)		136 (30)	324 (70)	
Cyclophosphamide only	+(n = 34)	11 (32)	23 (68)	0.086	11 (32)	23 (68)	NS
	−(248)	61 (25)	187 (75)		70 (28)	178 (72)	
Neoadjuvant							
All modalities	+(n = 9)	8 (89)	1 (11)	0.008	8 (89)	1 (11)	0.002
	−(n = 31)	14 (45)	17 (55)		15 (48)	16 (52)	
All modalities without trastuzumab	+(n = 8)	7 (88)	1 (12)	0.01	7 (88)	1 (12)	0.005
	−(n = 29)	13 (45)	16 (55)		14 (48)	15 (52)	
Cyclophosphamide all modalities	+(n = 8)	7 (88)	1 (12)	0.01	7 (88)	1 (12)	0.005
	−(n = 28)	12 (43)	16 (57)		12 (43)	16 (57)	
Cyclophosphamide without trastuzumab	+(n = 8)	7 (88)	1 (12)	0.01	7 (88)	1 (12)	0.006
	−(n = 27)	12 (44)	15 (56)		12 (44)	15 (56)	
Adjuvant							
Chemo alone	+(n = 18)	4 (22)	14 (78)	NS	4 (22)	14 (78)	NS
	−(n = 88)	25 (28)	63 (72)		28 (32)	60 (68)	
Chemo alone without trastuzumab	+(n = 18)	4 (22)	14 (78)	NS	4 (22)	14 (78)	NS
	−(n = 88)	25 (28)	63 (72)		28 (32)	60 (68)	
Chemo and hormonal	+(n = 7)	0 (0)	7 (100)	NS	0 (0)	7 (100)	NS
	−(n = 128)	23 (18)	105 (82)		28 (22)	100 (78)	
Chemo and hormonal without trastuzumab	+(n = 7)	0 (0)	7 (100)	NS	0 (0)	7 (100)	NS
	−(n = 128)	23 (18)	105 (82)		28 (22)	100 (78)	

NS, not significant.

survival for ALDH1A+ was 2.83.[27] These findings are in agreement with ours, although we found that in addition to correlation with biomarkers status, there was correlation with nodal status and overall tumor stage which are generally predictors of worse disease-free and overall survival.

Tanei et al[28] studied 108 patients who received neoadjuvant paclitaxel- and epirubicin-based chemotherapy. When ALDH1A1 and CD44/CD24 expressions were compared between core needle biopsy and subsequent excision, there was significant increase in ALDH1A1+ but not CD44+/CD24− cells. They concluded that ALDH1A1+ cells have a significant role in resistance to this regimen. It was speculated that the reason for this resistance was that these tumors contain a higher proportion of cancer stem cells. However, correlation with the clinical outcome of patients with incomplete pathological response including disease-free survival and overall survival was not studied.[28] We found that patients who were treated with neoadjuvant chemotherapy had higher frequency of ALDH1A1 expression than did adjuvant-treated and overall patients. This could be due to the increase in ALDH1A1-expressing cells after neoadjuvant therapy as Tanei et al[28] has suggested. The other reason is probably due to the selection bias, as neoadjuvant-treated patients had relatively advanced stage and more frequently had high-risk biomarkers (8 HER2+ and 20 triple negative).

Resetkova et al[16] studied 34 patients treated in the neoadjuvant setting. They found that although there was no increase in positive cells after neoadjuvant therapy, there was high degree of stromal expression which was associated with the best disease-free survival and a trend for overall survival.[16] We did not evaluate ALDH1A1 expression in stromal cells due to the study design of examining small tissue microarray cores (0.6 mm). These cores were predominantly tumor cells with no or very minimal stroma.

In these two previous studies, the authors concluded that the function of stem cells marked with ALDH1A1 is the reason for therapy resistance. However, in the light of new discovery of ALDH1A3 and not ALDH1A1 being a marker for stem cells, we believe that this is not the sole reason for this resistance. Bunting and Townsend[21] have suggested that human class 1 ALDH can oxidize aldophosphamide, a key aldehyde intermediate in the activation pathway of cyclophosphamide and other oxazaphosphorine in hamster v79 cell lines. Moreover, Dylla et al[29] found that cyclophosphamide treatment of human colonic xenografts enriches for CD44+ ALDH1A1+ cells, and that these double-positive cells are more tumorigenic than cells selected solely on the basis of CD44 positivity. Therefore, there is compelling evidence that ALDH1A1 has an important role in inactivating cyclophosphamide. In our study, we found that the risk of
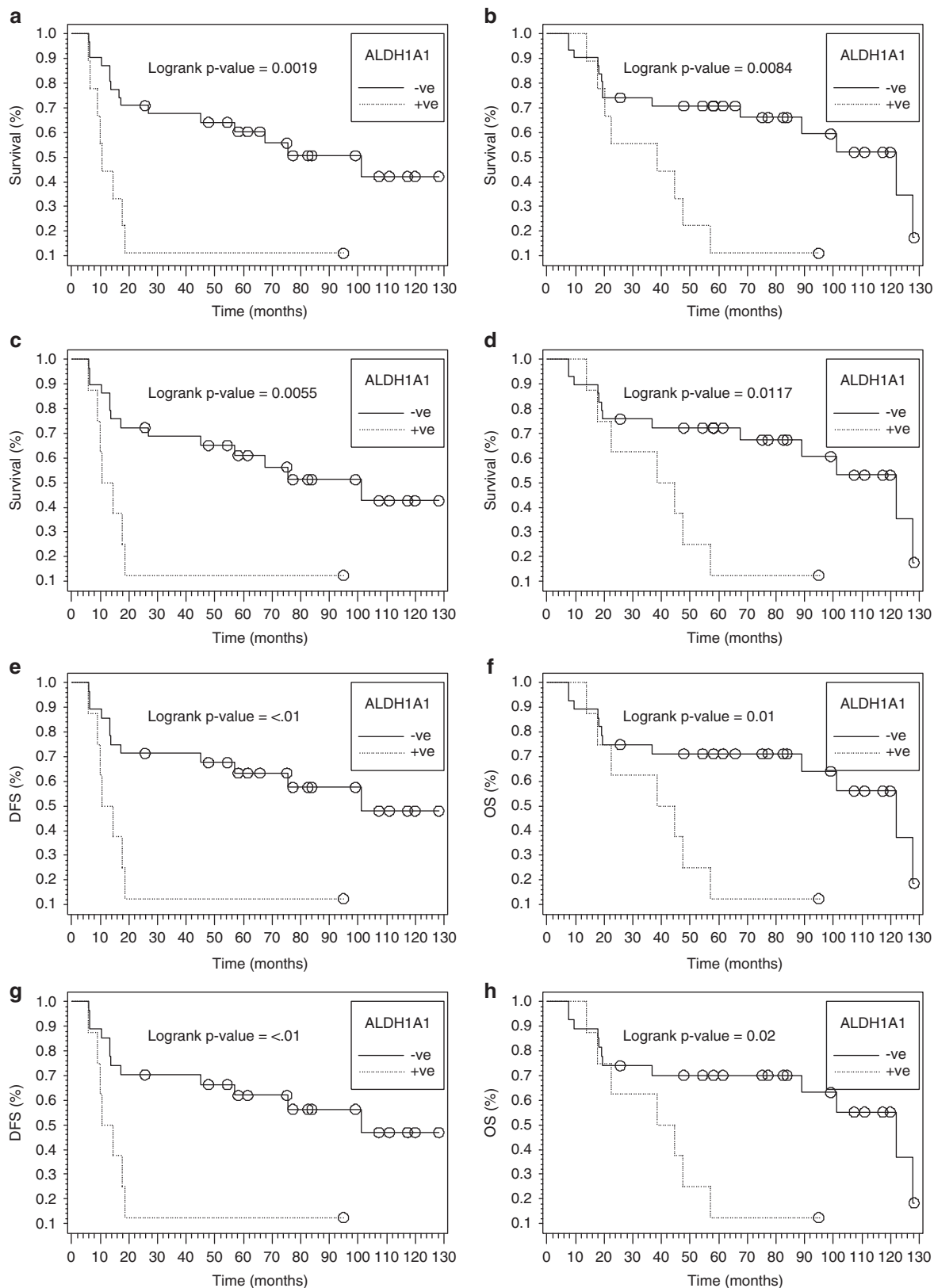
Figure 3 Kaplan–Meier curves showing ALDH1A1 correlation with disease-free survival and overall survival in patients treated in the neoadjuvant setting: (**a** and **b**) with or without trastuzumab and with or without cyclophosphamide; (**c** and **d**) without trastuzumab and with or without cyclophosphamide; (**e** and **f**) with cyclophosphamide with or without trastuzumab; (**g** and **h**) with cyclophosphamide without trastuzumab.

Table 4 Multivariate analysis for overall and disease-free survival for patients who were treated in the neoadjuvant setting with cyclophosphamide without trastuzumab

Variables	OS		DFS	
	HR (95% CI)	P-value	HR (95% CI)	P-value
ALDH1A1				
+ *vs* −	3.05 (0.85, 10.93)	0.09	11.56 (2.13, 62.86)	0.005
Age (years)				
≤50 *vs* >50	4.815 (0.6, 38.7)	NS	1.459 (0.16, 13.62)	NS
RCB[a]	2.13 (1.09, 4.19)	0.03	2.879 (1.29, 6.44)	0.01
Race				
Black *vs* Caucasian	0.45 (0.09, 2.32)	NS	0.49 (0.09, 2.59)	NS
Others *vs* Caucasian	0		0	
Menopausal status				
Post *vs* pre	3.05 (0.38, 24.35)	NS	0.545 (0.058, 5.122)	NS
Tumor type				
HER2+ *vs* triple negative	0.1 (0.01, 0.71)	0.051	0.44 (0.09,2.17)	NS
HR+/HER2− *vs* triple negative	0.04 (0.002, 0.92)		0.045 (0.002, 1.0)	

NS, not significant; RCB, residual cancer burden.
Stage was not included in the model because of the small count for stage.
[a]RCB is used as a continuous variable not divided into classes.

Table 5 Multivariate analysis for overall and disease-free survival for patients treated in the neoadjuvant setting without trastuzumab

Variables	OS		DFS	
	HR (95% CI)	Multivariate P-value	HR (95% CI)	Multivariate P-value
ALDH				
+ *vs* −	2.57 (0.78, 8.52)	NS	7.08 (1.61 ,31.13)	0.01
Age (years)				
≤50 *vs* >50	3.79 (0.54, 26.84)	NS	1.846 (0.235, 14.529)	NS
RCB	1.703 (0.96, 3.03)	0.07	2.47 (1.24, 4.93)	0.01
Race				
Black *vs* Caucasian	0.54 (0.12, 2.4)	NS	0.442 (0.095, 2.05)	NS
Others *vs* Caucasian	—		—	
Menopausal status				
Post *vs* pre	1.77 (0.27, 11.53)	NS	0.58 (0.078, 4.33)	NS
Status				
HER2+ *vs* triple negative	0.181 (0.04, 0.93)	0.07	0.47 (0.115, 1.911)	NS
HR+/HER2− *vs* triple negative	0.07 (0.01, 1.09)		0.13 (0.012, 1.38)	

RCB, residual cancer burden.
Stage was not included in the model because of the small count for stage = 1. RCB is used as a continuous variable.

local recurrence increases from 7- to 11-fold when all neoadjuvant-treated patients included in the analysis *vs* only cyclophosphamide-treated patients. Taking all these factors into account, we speculate that in addition to stem cells marked with ALDH1A1 inducing therapy resistance, ALDH1A1 also functions as an enzyme that inactivates cyclophosphamide.

However, patients who were treated with cyclophosphamide as part of the adjuvant chemotherapy had no significant difference in disease-free or overall survival. There are at least three possible reasons for this lack of significance. The first reason could be due to the small number of events in the positive group (four cases). ALDH1A1 was tested in the specimens after chemotherapy for neoadjuvant

patients. As Tanei *et al* found, chemotherapy kills ALDH1A1-negative tumor cells, which subsequently increases the percentage of ALDH1A1-positive tumor cells. Therefore, the second reason could be that tumors tested in the adjuvant setting could have dispersed ALDH1A1-positive tumor cells that were not captured by our method of testing (immunohistochemistry on 0.6-mm cores assembled in tissue microarray blocks). The third reason could be the different chemotherapy regimens that were given. Cyclophosphamide was part of the adjuvant therapy with methotrexate and fluorouracil, or with adriamycin with or without taxane.

Given the fact that HER2 + breast cancer is a more aggressive disease and requires anti-HER2 therapy with trastuzumab,[30,31] two analyses were conducted in HER2 + patients, one with trastuzumab therapy and one without. We found that ALDH1A1 was significantly correlated with disease-free survival and overall survival in both analyses. Given the small sample size and the number of events for the trastuzumab-treated group, the analysis could not be interpreted. Another confounding factor was the lack of consistency in chemotherapy regimens. Most patients received cyclophosphamide plus doxorubicin with or without a taxane. These differences are due to the fact that some patients received chemotherapy before the widespread adoption of taxanes as part of adjuvant chemotherapy.

We retrospectively studied ALDH1A1 expression in post-neoadjuvant therapy of the excisional biopsies that showed minimal or no response to therapy. In these types of cases, there is no validated factor that can predict tumor recurrence or worse clinical outcome. Symmans *et al*[25] have shown that RCB was significantly correlated with clinical outcome in which complete pathological response or minimal residual disease had the same prognosis, whereas extensive residual disease was associated with poor prognosis. In our study, all patients had intermediate or extensive residual disease. When data were analyzed as continuous variable, patients who had higher RCB had higher risk of tumor recurrence. In analyzing ALDH1A1 expression in this group of patients, we found that a positive ALDH1A1 patient treated with neoadjuvant therapy with cyclphosphamide and without trastuzumab had 11-fold higher risk of local recurrence than did a negative ALDH1A1 patient. Therefore, we believe that for patients treated in the neoadjuvant setting who have substantial RCB, ALDH1A1 could be used as a biomarker to predict clinical outcome and prognosis. However, one limitation of this study was that five patients were HER2 + but not treated with trastuzumab. Owing to the small sample size, multivariate analysis could not be performed.

We conclude that ALDH1A1 correlates with triple-negative and HER2 + breast cancers. It is also correlated with advanced stage and lymph-node status. More importantly, ALDH1A1 could be used as a predictor biomarker for worse clinical outcome in patients treated with cytotoxic neoadjuvant therapy that includes cyclophosphamide and have substantial RCB. In the future, this marker may be used as an integral part of personalized therapy. However, the number of studied cases was relatively small. Therefore, a larger cohort is required to verify our results.

Disclosure/conflict of interest

The authors declare no conflict of interest.

References

1 Tavassoli FA, Devilee P, (eds) World Health Organization Classification of Tumors. Pathology and Genetics of Tumors of the Breast and Female Genital Organs. IARC Press: Lyon, France, 2003. pp 13.
2 Perou CM, Sørlie T, Eisen MB, *et al.* Molecular portraits of human breast tumours. Nature 2000;17: 747–752.
3 Sørlie T, Perou CM, Tibshirani R, *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA 2001;98:10869–10874.
4 Honeth G, Bendahl PO, Ringnér M, *et al.* The CD44+/CD24− phenotype is enriched in basal-like breast tumors. Breast Cancer Res 2008;10:R53.
5 Liu R, Wang X, Chen GY, *et al.* The prognostic role of a gene signature from tumorigenic breast-cancer cells. N Engl J Med 2007;356:217–226.
6 Buess M, Rajski M, Vogel-Durrer BM, *et al.* Tumor-endothelial interaction links the CD44(+)/CD24(−) phenotype with poor prognosis in early-stage breast cancer. Neoplasia 2009;11:987–1002.
7 Herschkowitz JI, Simin K, Weigman VJ, *et al.* Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol 2007;8:R76.
8 Creighton CJ, Li X, Landis M, *et al.* Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. Proc Natl Acad Sci USA 2009;106:13820–13825.
9 Neumeister V, Agarwal S, Bordeaux J, *et al. In situ* identification of putative cancer stem cells by multiplexing ALDH1, CD44, and cytokeratin identifies breast cancer patients with poor prognosis. Am J Pathol 2010;176:2131–2138.
10 Ginestier C, Hur MH, Charafe-Jauffret E, *et al.* ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell 2007;1:555–567.
11 Marcato P, Dean CA, Pan D, *et al.* Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis. Stem Cells 2011;29:32–45.
12 Abraham BK, Fritz P, McClellan M, *et al.* Prevalence of CD44+/CD24−/low cells in breast cancer may not be associated with clinical outcome but may favor distant metastasis. Clin Cancer Res 2005;11:1154–1159.
13 Mylona E, Giannopoulou I, Fasomytakis E, *et al.* The clinicopathologic and prognostic significance of CD44+/CD24(−/low) and CD44−/CD24+ tumor cells

in invasive breast carcinomas. Hum Pathol 2008;39: 1096–1102.

14 Aulmann S, Waldburger N, Penzel R, *et al.* Reduction of CD44(+)/CD24(−) breast cancer cells by conventional cytotoxic chemotherapy. Hum Pathol 2010;41:574–581.

15 Nalwoga H, Arnes JB, Wabinga H, *et al.* Expression of aldehyde dehydrogenase 1 (ALDH1) is associated with basallike markers and features of aggressive tumours in African breast cancer. Br J Cancer 2010;102:369–375.

16 Resetkova E, Reis-Filho JS, Jain RK, *et al.* Prognostic impact of ALDH1 in breast cancer: a story of stem cells and tumor microenvironment. Breast Cancer Res Treat 2010;123:97–108.

17 Morimoto K, Kim SJ, Tanei T, *et al.* Stem cell marker aldehyde dehydrogenase 1-positive breast cancers are characterized by negative estrogen receptor, positive human epidermal growth factor receptor type 2, and high Ki67 expression. Cancer Sci 2009;100:1062–1068.

18 Charafe-Jauffret E, Ginestier C, Iovino F, *et al.* Aldehyde dehydrogenase 1-positive cancer stem cells mediate metastasis and poor clinical outcome in inflammatory breast cancer. Clin Cancer Res 2010;16:45–55.

19 Sladek NE. Human aldehyde dehydrogenases: potential pathological, pharmacological, and toxicological impact. J Biochem Mol Toxicol 2003;17:7–23.

20 Vasiliou V, Nebert DW. Analysis and update of the human aldehyde dehydrogenase (ALDH) gene family. Hum Genomics 2005;2:138–143.

21 Bunting KD, Townsend AJ. *De novo* expression of transfected human class 1 aldehyde dehydrogenase (ALDH) causes resistance to oxazaphosphorine anticancer alkylating agents in hamster V79 cell lines. Elevated class 1 ALDH activity is closely correlated with reduction in DNA interstrand cross-linking and lethality. J Biol Chem 1996;271:11884–11890.

22 Moreb JS, Schweder M, Gray B, *et al. In vitro* selection for K562 cells with higher retrovirally mediated copy number of aldehyde dehydrogenase class-1 and higher resistance to 4-hydroperoxycyclophosphamide. Hum Gene Ther 1998;9:611–619.

23 Canuto RA, Muzio G, Salvo RA, *et al.* The effect of a novel irreversible inhibitor of aldehyde dehydrogenases 1 and 3 on tumour cell growth and death. Chem Biol Interact 2001;130–132:209–218.

24 Klein TE, Chang JT, Cho MK, *et al.* Integrating genotype and phenotype information: an overview of the PharmGKB project. Pharmacogenetics Research Network and Knowledge Base. Pharmacogenomics J 2001;1:167–170.

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

26 Singletary SE, Allred C, Ashley P, *et al.* Revision of the American Joint Committee on Cancer staging system for breast cancer. J Clin Oncol 2002;20:3628–3636.

27 Zhou L, Jiang Y, Yan T, *et al.* The prognostic role of cancer stem cells in breast cancer: a meta-analysis of published literatures. Breast Cancer Res Treat 2010;122:795–801.

28 Tanei T, Morimoto K, Shimazu K, *et al.* Association of breast cancer stem cells identified by aldehyde dehydrogenase 1 expression with resistance to sequential Paclitaxel and epirubicin-based chemotherapy for breast cancers. Clin Cancer Res 2009;15:4234–4241.

29 Dylla SJ, Beviglia L, Park IK, *et al.* Colorectal cancer stem cells are enriched in xenogeneic tumors following chemotherapy. PLoS One 2008;3:e2428.

30 Slamon DJ, Clark GM, Wong SG, *et al.* Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science 1987;235:177–182.

31 Slamon DJ, Leyland-Jones B, Shak S, *et al.* Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med 2001344:783–792.