ORIGINAL ARTICLE

The prognostic significance of aldehyde dehydrogenase 1A1 (ALDH1A1) and CD133 expression in early stage non-small cell lung cancer

Muhammad Alamgeer,[1,2] Vinod Ganju,[1,2] Anette Szczepny,[2] Prudence A Russell,[3] Zdenka Prodanovic,[4] Beena Kumar,[4] Zoe Wainer,[5] Tracey Brown,[6] Michal Schneider-Kolsky,[7] Matthew Conron,[8] Gavin Wright,[5] D Neil Watkins[2]

[1]Department of Medical Oncology, Monash Medical Centre, East Bentleigh, Melbourne, Australia
[2]Monash Institute of Medical Research, Monash University, Clayton, Victoria, Australia
[3]Department of Pathology, St Vincent's Hospital, Fitzroy, Melbourne, Australia
[4]Department of Pathology, Monash Medical Centre, Clayton, Melbourne, Australia
[5]Department of Surgery, University of Melbourne, St Vincent's Hospital, Fitzroy, Melbourne, Australia
[6]Department of Biochemistry and Molecular Biology, Faculty of Medicine, Nursing and Health Sciences, Monash University, Clayton, Melbourne, Australia
[7]Department of Medical Imaging and Radiation Science, Faculty of Medicine, Nursing and Health Sciences, Monash University, Clayton, Melbourne, Australia
[8]Department of Respiratory Medicine, St Vincent Hospital, Fitzroy, Melbourne, Australia

Correspondence to
Professor D Neil Watkins, Monash Institute of Medical Research, Monash University, 27-31 Wright St, Clayton, VIC 3168, Australia; neil.watkins@monash.edu and Dr Gavin Wright, Department of Surgical Oncology, St Vincent's Hospital, 55 Victoria Parade, Fitzroy, VIC 3065, Australia; gavin.wright@svhm.org.au

MA and VG contributed equally.

ABSTRACT

Background Expression of aldehyde dehydrogenase 1A1 (ALDH1A1) and CD133 has been functionally associated with a stem cell phenotype in normal and malignant cells. The prevalence of such cells in solid tumours should therefore correlate with recurrence and/or metastasis following definitive surgical resection. The aim of this study was to evaluate the prognostic significance of ALDH1A1 and CD133 in surgically resected, early stage non-small cell lung cancer (NSCLC).

Methods A retrospective analysis of ALDH1A1 and CD133 expression in 205 patients with pathologic stage I NSCLC was performed using immunohistochemistry. The association between the expression of both markers and survival was determined.

Results We identified 62 relapses and 58 cancer-related deaths in 144 stage 1A and 61 stage 1B patients, analysed at a median of 5-years follow-up. Overexpression of ALDH1A1 and CD133, detected in 68.7% and 50.7% of primary tumours, respectively, was an independent prognostic indicator for overall survival by multivariable Cox proportional hazard model (p=0.017 and 0.039, respectively). Overexpression of ALDH1A1, but not of CD133, predicted poor recurrence-free survival (p=0.025). When categorised into three groups according to expression of ALDH1A1/CD133, patients with overexpression of both ALDH1A1 and CD133 belonged to the group with the shortest recurrence-free and overall survival (p=0.015 and 0.017, respectively).

Conclusions Expression of ALDH1A1 and CD133, and coexpression of ALDH1A1 and CD133, is strongly associated with poor survival in early-stage NSCLC following surgical resection. These data are consistent with the hypothesis that expression of stem cell markers correlates with recurrence as an indirect measure of self-renewal capacity.

INTRODUCTION

Poor outcomes for patients with lung cancer are associated with limited opportunities for early detection and the lack of response to chemotherapy and radiotherapy. Although curative surgical resection is the current treatment of choice for stage 1 non-small cell lung cancer (NSCLC), the risk of loco-regional and distant relapse in stage 1 lung cancer remains high at 22%–30%,[1] with a 5-year overall survival (OS) rate of 73% for stage 1A and 58% for stage 1B NSCLC.[2] Cisplatin-based

Key messages

What is the key question?
► If the cancer stem cell hypothesis is correct, then expression of stem cell markers in early-stage non-small cell lung cancer (NSCLC) should predict recurrence following curative surgery.

What is the bottom line?
► Coexpression of two recognised stem cell markers, aldehyde dehydrogenase 1A1 (ALDH1A1) and CD133, are associated with a markedly increased risk of recurrence in early-stage NSCLC, with significant implications for individualised medicine, and the biology of cancer stem cells.

Why read on?
► Our study examines a large, well-characterised cohort of early-stage NSCLC for the prognostic significance of expression of both ALDH1A1 and CD133 and shows that the interpretation of these expression patterns is dependant on histological types.

adjuvant chemotherapy has provided a further 5% increase in survival for resected stages 2 and 3 but only for a small subset of stage 1 NSCLC.[3] In stage 1 NSCLC, adjuvant treatments have either no benefit,[3] or could potentially be detrimental.[4] The identification of biomarkers that predict recurrence in early stage NSCLC independent of tumour/node/metastasis (TNM) stage may help identify patients who might benefit from adjuvant chemotherapy and also shed light on the potential drivers of recurrence and metastasis.

Published studies have described gene expression profiles, the expression of molecules involved in DNA repair (ERCC1, BRCA1),[5 6] or tumour invasiveness (RRM1),[7] as potential prognostic biomarkers. However, the clinical use of these molecular markers is currently limited. Therefore, the identification of robust biomarkers, which predict a high risk of relapse, may allow a more targeted approach to adjuvant therapies for stage 1 NSCLC.

Figure 1 Correlation in staining between EP1933Y (rabbit monoclonal 1:100) and 44/ALD (mouse monoclonal, 1:200) antibodies for stage 1 non-small cell lung cancer. (A) Quantitative assessment of staining intensities for 30 cases (positive and negative) for aldehyde dehydrogenase 1A1 (ALDH1A1). (B) Photomicrograph of case no. 3 (a and b) and case no. 16 (c and d) stained with EP1933Y antibody (a and c) and 44/ALD antibody (b and d).
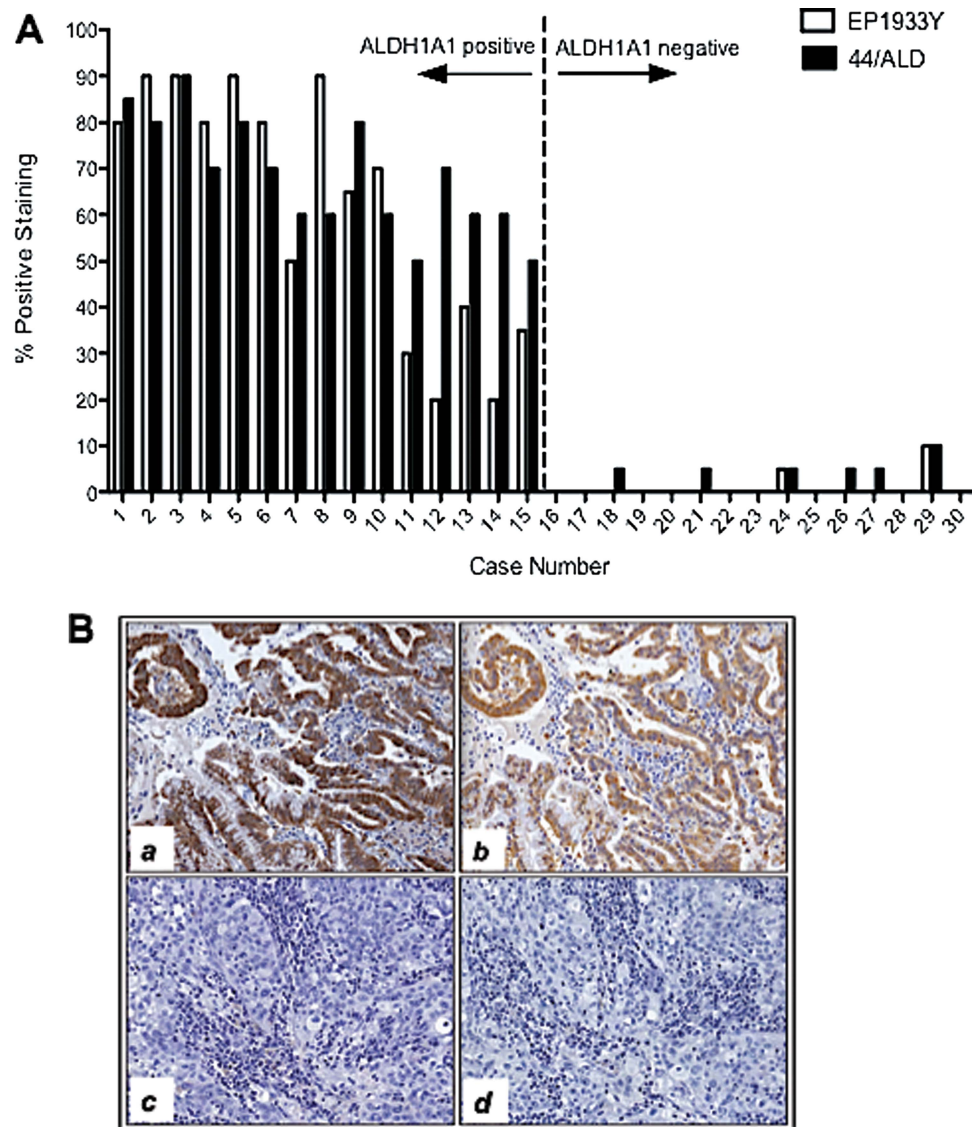
According to the cancer stem cell hypothesis, most solid tumours contain a small subset of phenotypically distinct cells with the properties of unlimited self-renewal, innate chemoresistance and enhanced clonogenic potential.[8] Among the most consistently identified cancer stem cell markers are the cytosolic enzyme aldehyde dehydrogenase 1 (ALDH1), its isoform aldehyde dehydrogenase 1A1 (ALDH1A1) and the transmembrane glycoprotein CD133.[9 10] However, the importance of these markers in early-stage lung cancer is yet to be established. If the stem cell hypothesis is correct, then the prevalence of cancer stem cells in resected early-stage NSCLC should strongly associate with the incidence of recurrent disease. Tumours with relatively high stem cell population are believed to have an aggressive phenotype, leading to recapitulation of the entire tumour after initial therapy due to their high proliferative potential.[11] In order to test this hypothesis, we investigated the expression of both ALDH1A1 and CD133 in a large retrospective cohort of stage I NSCLC patients undergoing surgical resection with curative intent. The primary objective was to determine the association of ALDH1A1 and CD133 expression in the tumour and survival.

MATERIALS AND METHODS
Patient population
From August 1999 until August 2010, a total of 267 consecutive patients undergoing surgical resection for stage 1 (according to TNM7 classification) NSCLC at either the Monash Medical Centre or at St Vincent's Hospital were reviewed. All histopathological information was systematically reviewed from corresponding haematoxylin and eosin slides. Patients with clear diagnosis of adenocarcinoma (ADC) or squamous cell carcinoma (SCC) on histology report were included. Large cell carcinoma was further designated as ADC or SCC by staining with either TTF-1 or p63. ADC was subclassified according to the new IASLC/ATS/ERS International Multidisciplinary Lung Adenocarcinoma Classification.[12] Other inclusion criteria were (i) no adjuvant chemotherapy or radiotherapy, (ii) minimum 18 months follow-up data available and (iii) adequate paraffin block available for analysis. Exclusion criteria were (i) patients with surgical mortality (defined as in-hospital death within 30 days after surgery) and (ii) ADC in situ (AIS) (previously pure bronchioalveloar carcinoma) or minimally invasive adenocarcinoma (MIA). A total of 205 patients met the inclusion

Table 1 Association between ALDH1A1 and CD133 expression and clinicopathological variables for patients with non-small cell lung cancer (N=205)

Variable	No. (%)	ALDH1A1 expression			CD133		
		Low score No. (%)	High score No. (%)	p Value†	Low score No. (%)	High score No. (%)	p Value†
Age							
≤70	98 (47.8)	32 (31.9)	66 (68.1)	0.761	46 (46.4)	52 (53.6)	0.571
>70	107 (52.2)	32 (29.2)	75 (70.8)		55 (50.9)	52 (49.1)	
Sex							
Male	125 (61)	42 (33.3)	82 (66.6)	0.442	60 (47.2)	65 (52.8)	0.660
Female	80 (39)	22 (26.3)	59 (73.7)		41 (51.2)	39 (48.8)	
Smoking							
Yes	182 (88)	54 (29.4)	128 (70.6)	0.365	90 (48.9)	92 (51.1)	0.414
No	23 (12)	10 (43.5)	13 (56.5)		11 (48)	12 (52)	
Stage							
1a	144 (70)	54 (37)	90 (63)	**0.003***	63 (43.4)	81 (56.6)	**0.023***
1b	61 (30)	10 (15)	51 (85)		38 (61.7)	23 (38.3)	
Histological type							
ADC	121 (59)	48 (33.9)	74 (66.1)	**0.002***	48 (39.7)	73 (60.3)	**0.001***
SCC	84 (41)	16 (18.3)	67 (81.7)		53 (62.2)	31 (37.8)	
Nodal dissection							
<10 nodes	121 (59)	35 (27.6)	86 (72.4)	0.279	60 (50)	60 (50)	0.847
≥10 nodes	84 (41)	30 (36.6)	54 (63.4)		43 (51.6)	42 (49.4)	
Lymphatic invasion							
Absent	174 (85)	55 (30.8)	119 (69.2)	0.841	86 (48.8)	88 (51.2)	1.000
Present	31 (15)	9 (29)	22 (71)		15 (48.4)	16 (51.6)	
Vascular invasion							
Absent	149 (72.6)	49 (32.7)	100 (67.3)	0.489	70 (46.3)	79 (53.7)	0.341
Present	56 (27.4)	15 (25)	42 (75)		31 (55.4)	25 (44.6)	
Surgery type							
Lobectomy	178 (87)	56 (30.7)	122 (69.3)	0.967	86 (47.7)	92 (52.3)	0.563
Wedge resection	23 (11)	7 (30.4)	16 (69.6)		12 (52.2)	11 (47.8)	
Pneumonectomy	4 (2)	1 (25)	3 (75)		3 (75)	1 (25)	

*Bold indicates statistical significance.
†χ² test.
ADC, adenocarcinoma; ALDH1A1, aldehyde dehydrogenase 1A1; SCC, squamous cell carcinoma.

criteria and were included in the analysis. All patients were followed up every 3 months for the first 2 years, then biannually thereafter.

Follow-up information was obtained from patients' records. Clinicopathological data routinely collected included age, sex, smoking history, tumour subtype (ADC vs SCC), tumour stage (1a vs 1b), lymphatic invasion, vascular invasion and type of surgery (lobectomy, pneumonectomy or wedge resection). Loco-regional recurrence was defined as tumour recurrence at the site of initial resection, ipsilateral hilar or mediastinal nodes. Any other site of recurrence was considered distant, including contralateral and supraclavicular nodes.

Specimen characteristics
Archived paraffin blocks were retrieved from 205 eligible patients. Individual sections of 4–5 μm were cut and mounted on aminopropylethoxysilane precoated glass slides. Sections from normal human liver and human colon were used as controls for ALDH1A1 and CD133, respectively.

Immunohistochemistry
All sections were stained with primary antibodies for ALDH1A1 and CD133. For ALDH1A1, we used two commercially available and previously well-used monoclonal antibodies (clone 44/

ALD, BD Transduction Laboratories, dil 1 : 200 and rabbit monoclonal IgG, clone EP1933Y, Abcam, dil 1 : 100). When comparing the immunohistochemical staining of the two antibodies in sequential sections of 55 cases in our cohort, they displayed similar pattern and intensity of staining in malignant and non-malignant cells (figure 1). Clone 44/ALD was used to stain the whole cohort. For CD133, we studied two antibodies (clone bs-0395R Bioss, dil 1 : 500 and clone Ab19898 Abcam, dil 1 : 500). Both antibodies recognise the same immunogen, located at intracellular C-terminus of human CD133 molecule, independent of glycosylation status of CD133. Initially, 80 cases were stained separately with both antibodies and 80% concordance was achieved. Then validation was performed against a normal colonic epithelium tissue with positive staining of the crypt cells. This has been shown to be a reliable method of identifying stem cells.[13] Clone bs-0395R was selected to stain the whole cohort.

The staining was performed using Vectastain *Ellite* ABC kit according to the manufacturer's recommendations. Briefly, sections were deparaffinised in xylene and rehydrated in ethanol. Antigen retrieval was performed by microwaving the slides in citrate buffer (pH 6.0) for 10 min. Endogenous peroxidase activity was inhibited by incubating the slides in 1% hydrogen peroxide for 15 min. A protein block with a 10% normal serum

was performed for 30 min. Incubation with primary antibodies was carried out at 4°C overnight. After washing with tris-buffered saline (TBS), the secondary antibody was applied for 30 min. Development of colour was achieved by 15 min incubation with diaminobezadine solution, followed by counterstaining with haematoxylin. All staining runs were accompanied by appropriate control slides.

Scoring of immunohistochemical stains

Two pathologists (BK and PR) independently evaluated all slides in a blinded manner and interobserver agreement was reached in all cases. Tissue sections were first examined at low power to characterise the overall staining pattern and to identify representative areas for precise quantitation. Immunostaining analysis was carried out using direct light microscopy in 5–10 different fields at 400× magnification. Approximately 500–1000 cells were counted per tumour, depending on the amount of tissue present. Only staining specific to cancer cells was taken as positive, while staining on stromal tissue, macrophages and cellular debris was considered as non-specific and was excluded from analysis. Patterns of staining, either membranous or cytoplasmic, were interpreted separately. CD133 immunoreactivity was evaluated within the neoplastic epithelial component where both cytoplasmic and membranous staining was quantitated. ALDH1A1 was quantitated in the cytoplasmic compartment but not on the pericellular membranes.

Scoring of ALDH1A1 and CD133 was performed according to the following criteria: (i) Proportion score (PS): to assess the total percentage of tumour cells showing staining (any intensity) with ALDH1A1 or CD133 and (ii) Intensity score (IS) to assess the intensity of staining in ALDH1A1 or CD133 stained cells. Each individual case was given an IS as follows; 0=no staining, 1+ = weak staining, 2+ = moderate staining and 3+=strong staining.
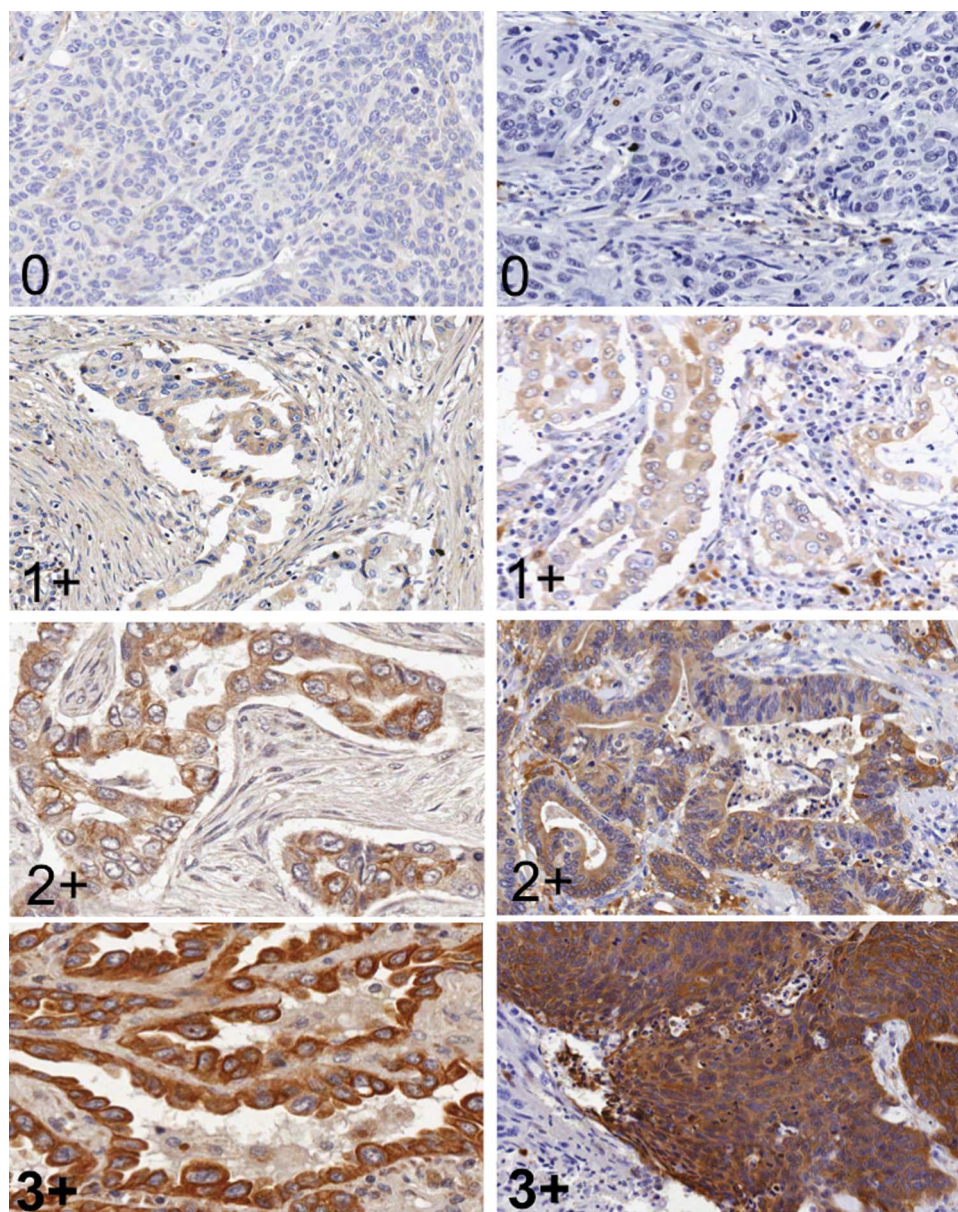

Figure 2 Representative immunohistochemical staining intensity of CD133 (left column) and aldehyde dehydrogenase 1A1 (ALDH1A1) (right column) from patients with stage 1 non-small cell lung cancer. In all cases, 0=no staining, 1+=mild, 2+=moderate and 3+=strong intensity of staining. Photographs were taken at magnification 200×.

Table 2 Univariable analysis of recurrence-free and overall survival of 205 patients with stage 1 non-small cell lung cancer

Variable	Recurrence-free survival			Overall survival		
	Median (months)	HR (95% CI)	p Value*	Median (months)	HR (95% CI)	p Value
Age						
≤70	111			NR		
>70	65.7	1.77 (1.11 to 2.82)	**0.028**	80.0	1.99 (1.16 to 3.42)	**0.015**
Stage						
1a	100			NR		
1b	75	1.43 (0.90 to 2.31)	0.070	NR	1.17 (0.70 to 2.00)	0.448
Sex						
Female	93			NR		
Male	75	1.41 (0.80 to 2.33)	0.185	100	1.55 (0.91 to 2.71)	0.129
Histology						
ADC	85			NR		
SCC	82	1.19 (0.75 to 1.92)	0.412	NR	1.15 (0.69 to 1.94)	0.489
Extent of lung resection						
WR	71.6			NR		
Lobectomy	82	1.00 (0.44 to 2.15)	0.990	NR	1.35 (0.55 to 3.25)	0.557
Lymphatic invasion						
Absent	82			NR		
Present	65.7	1.26 (0.73 to 2.39)	0.439	NR	1.42 (0.72 to 2.75)	0.294
Vascular invasion						
Absent	82			NR		
Present	68	1.65 (0.99 to 2.60)	0.051	NR	1.45 (0.80 to 2.51)	0.186
ALDH1A1						
Low	NR			NR		
High	68	2.25 (1.22 to 4.00)	**0.005**	100	2.0 (1.03 to 3.93)	**0.027**
CD133						
Low	NR			NR		
High	64	1.22 (0.80 to 1.96)	0.362	100	1.43 (0.80 to 2.35)	0.102

Bold indicates statistical significance.
*Log rank test.
ADC, adenocarcinoma; ALDH1A1, aldehyde dehydrogenase 1A1; NR, not reached (median overall survival for these patients had 'not reached' during this follow-up period); SCC, squamous cell carcinoma.; WR, Wedge resection.

Cut-off point determination.

Patients' samples with at least 10% cells expressing ALDH1A1 in moderate-to-strong intensity were considered positive for ALDH1A1, while patients' samples with at least 5% of CD133 expression in moderate-to-strong intensity were considered positive for CD133. As no universally acceptable cut-off point for immunohistochemistry (IHC) detected stem cell markers has been described so far, we devised the following strategy: The cohort was randomly divided into a smaller 'training set' and a larger 'validation set'. Cut-off points were determined based on the results from the training set and were then applied to larger validation set. This strategy has previously been described by Hilsenbeck et al[14] to reduce the risk of type 1 error associated with multiple testing for optimal cut-off points.

Statistical analysis

The expressions of ALDH1A1 and CD133 were dichotomised into either 'low' or 'high' scores according to the criteria described above. The correlation between ALDH1A1 and CD133 expression and clinicopathological characteristics were then analysed using a χ^2 test. OS was defined as duration (in months) between date of surgery and date of death due to any cause. Recurrence-free survival was defined as duration (in months) between date of surgery and date of first recurrence or death due to any cause. Patients alive and showing no recurrence at the last follow-up were censored. The survival curves

were plotted using the Kaplan–Meier method and log-rank test was used to assess the statistical difference between the groups. Variables with p value 0.1 or less were entered into multivariable analysis and the Cox proportional hazard model was used to carry out group comparisons. The assessment of the proportional hazards assumption was done graphically by plotting cumulative hazards functions for the covariates. Statistical significance was set at probability value of <0.05, with two-tailed p values. All analyses were performed using SPSS for windows V.20 (SPSS Inc, Chicago, Illinois, USA).

RESULTS
Patient characteristics

The characteristics of 205 patients with stage 1 NSCLC are shown in table 1. The median age was 70 years (range 34–85). With a median follow-up period of 60 months (range 18–140 months), the 5-year OS was 75.3% (77.4% in stage 1a and 70% in stage 1b). Tumour recurrence was recorded in 62 (30%) patients.

Expression of ALDH1A1 and CD133 in NSCLC tumours

Clinicopathological characteristics according to ALDH1A1 and CD133 are summarised in table 1. Of all 205 samples, 141/205 (68.7%) and 104/205 (50.7%) were considered high (positive) for ALDH1A1 and CD133, respectively. A few cases also showed ALDH1A1 staining on normal bronchial epithelium at variable intensity. A representative case of high and low scores
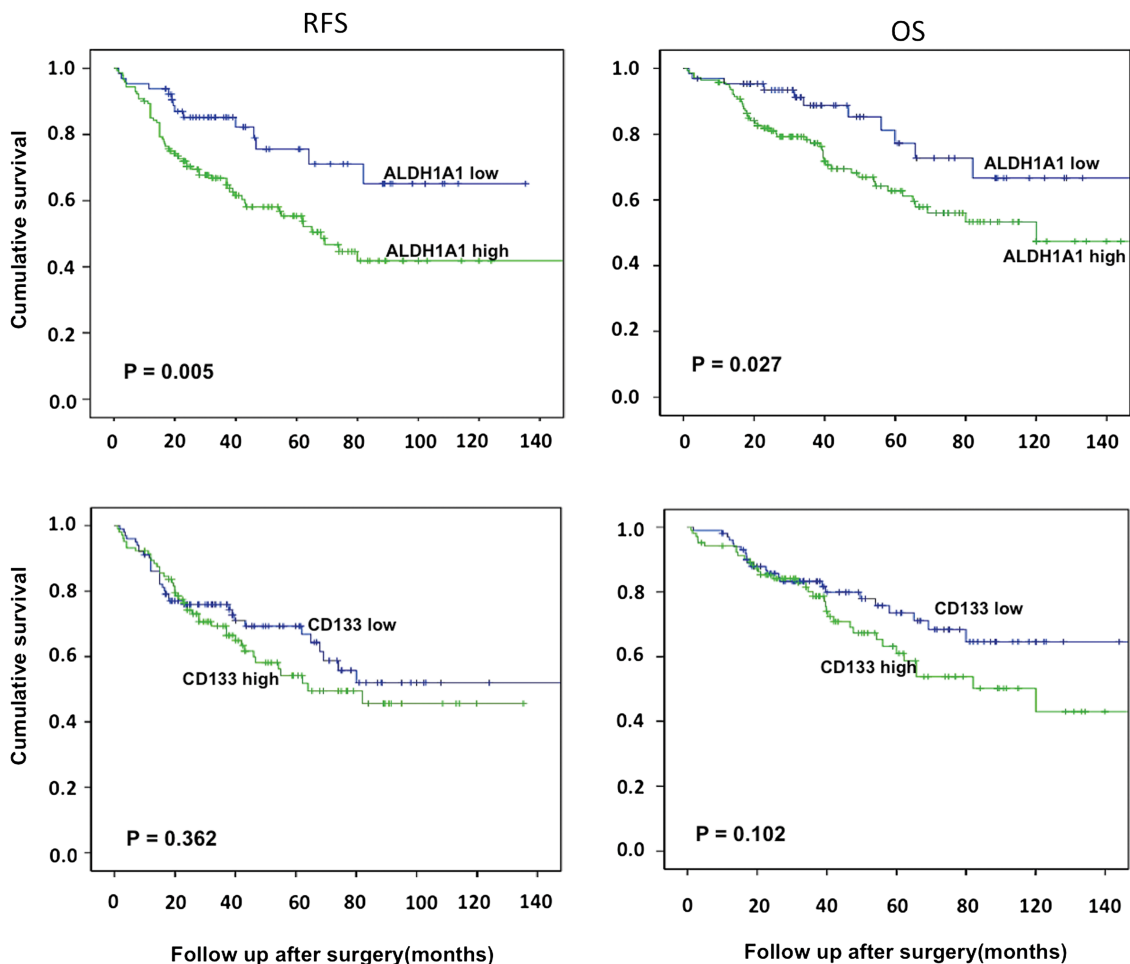
Figure 3 Kaplan–Meier analysis in patients with stage 1 non-small cell lung cancer according to aldehyde dehydrogenase 1A1 (ALDH1A1) and CD133 expression. (A) Recurrence-free survival in ALDH1A1 low versus high expression, (B) overall survival in ALDH1A1 low versus high expression, (C) recurrence-free survival in CD133 low versus high and (D) overall survival in CD133 low versus high expression.

of each marker is shown in figure 2. ALDH1A1 expression was strongly associated with TNM stage of 1b (p=0.003) and histological type of SCC (p=0.002). CD133 was strongly associated with histological type of ADC (p=0.001) and was more prevalent in stage 1a (p=0.023).

Expression of ALDH1A1 and CD133 as prognostic factors in patients with NSCLC

Kaplan–Meier survival analysis was carried out to investigate the prognostic value of individual marker in the whole cohort. Univariable analysis for all other variable is shown in table 2.

The results show that ALDH1A1 expression had significant impact on both recurrence-free survival (p=0.005, HR 2.25 95% CI 1.2 to 4.0) and OS (p=0.027, HR 2.0, 95% CI 1.03 to 3.9), while age (dichotomised at median 70) years also impacted the survival significantly. In case of CD133, there was a direction effect towards poor OS with CD133 high scores but it was not statistically significant. (figure 3).

Variables with p value of 0.1 or less were entered in Cox regression model for multivariable analysis. As shown in table 3, both ALDH1A1 and CD133 were independent predictors of survival. ALDH1A1 expression was associated with worse

Table 3 Multivariable analysis of recurrence-free and overall survival of 205 patients with stage 1 non-small cell lung cancer

Variables	Recurrence-free survival		Overall survival	
	HR (95% CI)	p Value*	HR (95% CI)	p Value
Age, years	1.21 (0.80 to 2.10)	0.209	1.03 (0.91 to 1.07)	0.135
Stage, 1b	1.66 (1.01 to 2.90)	0.049	–	
Sex	–		1.47 (0.80 to 2.51)	0.170
Vascular invasion (Present)	1.42 (0.80 to 2.34)	0.145	–	
ALDH1A1 (High)	1.88 (1.33 to 3.47)	**0.025**	2.00 (1.14 to 3.56)	**0.017**
CD133 (High)	–		8.50 (1.22 to 61.12)	**0.039**

Bold indicates statistical significance.
*Cox regression.
ALDH1A1, aldehyde dehydrogenase 1A1.

recurrence-free survival (p=0.025) and OS (p=0.017), while CD133 was associated with worse overall mortality (p=0.039). Stage 1b was associated with significantly worse recurrence-free survival (p=0.049). None of the other variables had a significant impact on survival (table 3).

Prognostic prediction using combined ALDH1A1 and CD133 staining

We studied the accumulative prognostic effect of ALDH1A1 and CD133 in stage 1 NSCLC. We divided 205 patients into three subgroups according to expression of ALDH1A1 and CD133. Group 1=(ALDH1A1^{low}/CD133^{low}) (n=32), group 2= (ALDH1A1^{high}/CD133^{low} or ALDH1A1^{low}/CD133^{high}) (n=107) and group 3=(ALDH1A1^{high}/CD133^{high}) (n=66). Kaplan–Meier survival curves were generated and differences between the three groups were examined. The results showed that patients with high expression of both CD133 and ALDH1A1 (group 3=double positive) had significantly shorter recurrence-free survival (p=0.015) and OS (p=0.017) compared with group 1 (double negative). Groups 2 (any marker positive) demonstrated a shorter progression-free survival (PFS) (p=0.046) and a trend towards shorter OS (p=0.077) compared with group 1 (figure 4).

Correlation between marker expression and NSCLC histology (ADC vs SCC)

Each marker was further studied according to the histological subtype. There was a significant association of each marker with histological subtype. ALDH1A1 was strongly expressed in SCC (p=0.002), while CD133 in was strongly expressed in ADC (p=0.001). We further studied the prognostic roles of both markers in each histological subtype. The results showed that a high ALDH1A1 score predicted shorter OS in ADC histological subtype (p=0.004), but not in SCC (p=0.743), while CD133 high scores did not show any significant difference in both histological groups (p=0.258 and 0.205 for SCC and ADC, respectively) (figure 5).

Adenocarcinoma subtypes

The proportions of predominant ADC subtypes are shown in table 4. Acinar predominant tumours were the major subtype in our cohort, making up 60/121 (51% of all ADC, followed by solid 21/121 (18%) and papillary 18/121 (15%). High ALDH1A1 and CD133 scores in acinar tumour subtype were associated with shorter 5-year OS (p=0.017 and 0.030, respectively). There was no major difference in terms of survival based on expression of CD133 and ALDH1A1 in other subgroups.

DISCUSSION

Despite progress in surgical techniques, the proportion of patients relapsing and ultimately dying of pathological stage 1 NSCLC after curative resection remains substantial.[2] A number of prognostic factors associated with disease relapse and death have been described. Currently, tumour stage appears to be the best indicator of poor outcome. Although several studies have reported intratumour vascular invasion,[15] as a poor prognostic marker, it is yet to be incorporated into clinical practice. Similarly, patient factors such age and sex are not reliable indicators for disease relapse.

In this study, we investigated the prognostic value of ALDH1A1 and CD133 expression in patients with resected stage 1 NSCLC. Our results showed that the expression of ALDH1A1 and CD133 in primary NSCLC was significantly associated with shorter OS. In multivariable model, ALDH1A1 and CD133 were
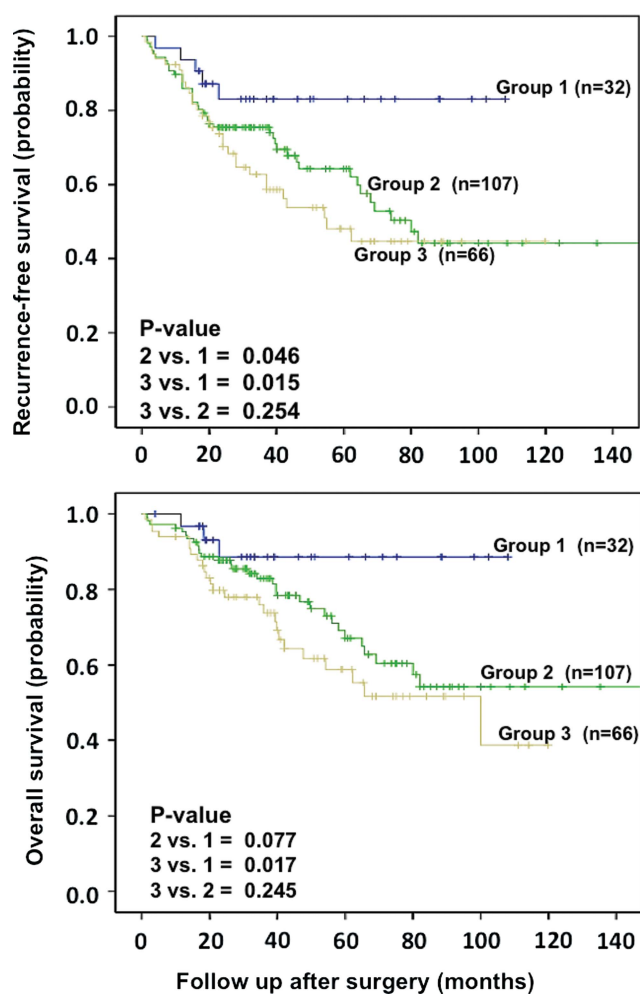

Figure 4 Kaplan–Meier survival analysis in patients with stage 1 non-small cell lung cancer according to number of markers expressed on individual patients. Patients in group 1 (ALDH1A1^{low}/CD133^{low}) had the best survival, group 2 (ALDH1A1^{high}/CD133^{low} or ALDH1A1^{low}/CD133^{high}) had the intermediate, while group 3 (ALDH1A1^{high}/CD133^{high}) had the worst survival. ALDH1A1, aldehyde dehydrogenase 1A1.

an independent factor for worse OS. Tumours expressing both ALDH1A1 and CD133 were associated with the worst outcome. By contrast, tumours negative for both markers were in the best prognostic group, whereas expression of one or other stem cell marker conferred an intermediate prognosis.

ALDH1 is a cytosolic isoenzyme, a member of the aldehyde dehydrogenase family responsible for oxidisation of intracellular aldehydes to carboxylic acids. Increased ALDH1 activity has been found in haematopoietic stem cells[16–18] and has been reported as a surrogate marker of cancer stem cells in several malignancies.[11 19 20] In vitro experiments suggest that isolated lung cancer cells with high ALDH1 activity are associated with cancer stem cell characteristics, including capacities of proliferation, self-renewal and resistance to chemotherapy.[9] ALDH1 positive cells have also showed enhanced engraftment capacity in nude mice.[9 21] CD133 (PROM-1 or AC133) is a transmembrane glycoprotein originally described in human haematopoietic stem cells.[22] Its expression is used for the isolation of normal stem cells from numerous tissues such as bone marrow,[22] brain[23] and kidney.[24] Recent data have provided evidence of strong association between expression of CD133 and stem cell characteristics in malignant tumours of brain,[25] prostate,[26] liver,[27] pancreas,[28] lung[10] and colon.[29]
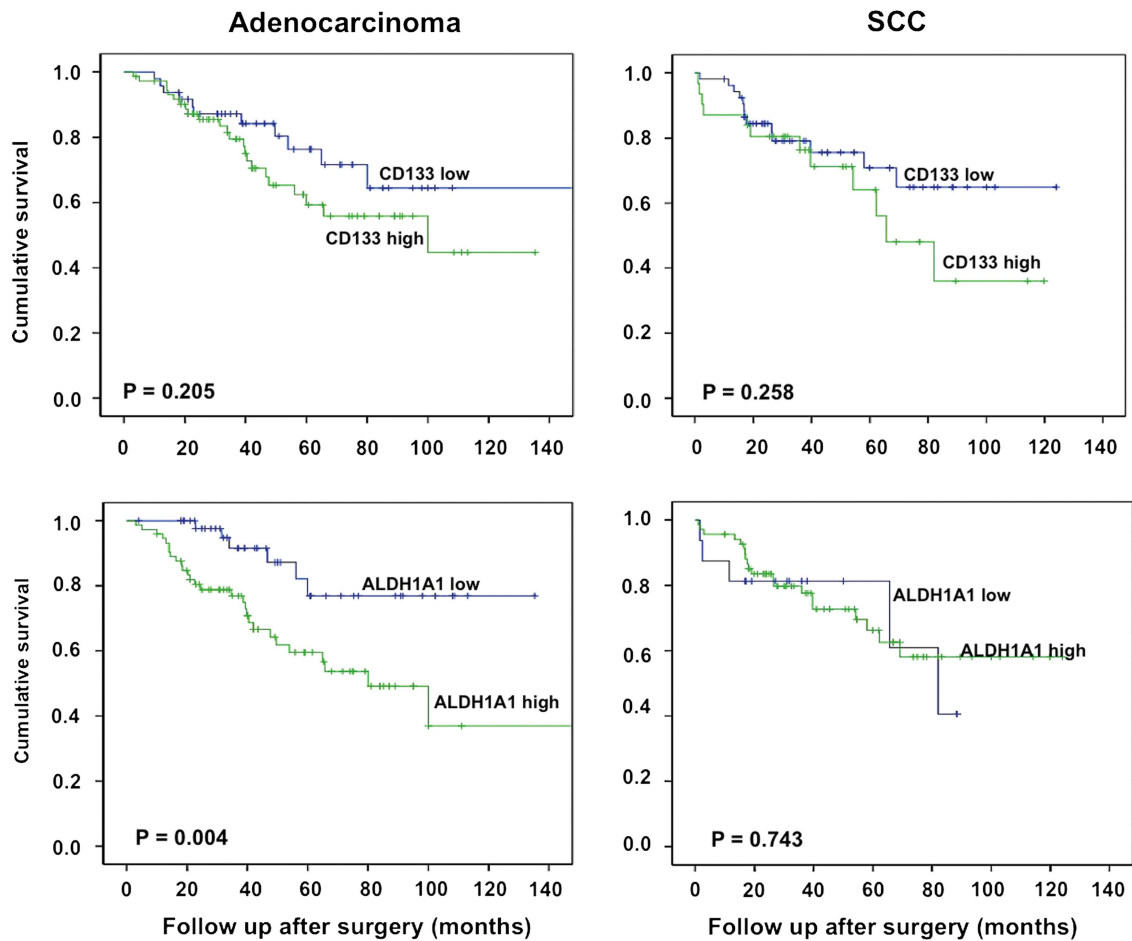
Figure 5 Kaplan–Meier survival curve according to expression of aldehyde dehydrogenase 1A1 (ALDH1A1) and CD133 in two histologic subtypes (adenocarcinoma (ADC) and squamous cell carcinoma) in patients with stage 1 non-small cell lung cancer.

Analysis from our study confirms that ALDH1A1 and CD133 overexpression is independently associated with poor survival in stage 1 lung cancer. Moreover, in our cohort, combined overexpression of ALDH1A1 and CD133 selects the patients from the worst prognostic group. Furthermore, our results also suggest that the prognostic role of these markers may be different in different histological subtypes. In ADC, ALDH1A1 expression was highly significant for poor survival, while CD133 showed a direction of effect towards poor survival. However, neither marker showed a significant difference in squamous histology. Further examination of ADC histologic subtypes suggested that both ALDH1A1 and CD133 are markers of poor survival in acinar

predominant histological subtype. Results with other subtypes were not significant, possibly due to fewer numbers of patients. We did not include AIS and MIA in our study due to their excellent prognosis.[30]

Work by Jiang et al[9] showed that ALDH1 overexpression is associated with poor prognosis in stage 1 NSCLC and that ALDH1 expression overlapped with CD133 in a small subset of patients. Similarly, Sullivan et al[31] reported that only ALDH1A1 (an isoform of ALDH1), but not CD133, was marker of poor prognosis in stage 1 NSCLC. Table 5 compares the methodologies and main results from previous studies with our findings. The consistencies in results support the importance of

Table 4 Five-year overall survival of adenocarcinoma subtypes according to marker expression (total n=110)

| Predominant subtype | N (%) | Five-year survival (%) | | | | | | | |
| | | Overall | ALDH1A1 | | | CD133 | | |
			High	Low	p Value	High	Low	p Value
Acinar	60 (51)	80	70.3	95.7	0.017*	60.7	79.8	0.030*
Solid	21 (18)	66.7	70.6	50.0	0.564	66.7	66.7	0.317
Papillary	18 (15)	72.2	55.6	88.9	0.219	63.6	85.7	0.582
Micropapillary	6 (5)	66.7	33.3	100	0.225	66.7	0	0.025*
Lepidic	5 (4)	80	100	66.7	0.380	66.7	100	0.380

*log rank (Mantel-cox).
ALDH1A1, aldehyde dehydrogenase 1A1.

Table 5 Comparison of major studies performed in non-small cell lung cancer (NSCLC) to investigate the prognostic significance of ALDH1/ALDH1A1 and CD133

Reference	Population	Number of patients	Specimen	Stem cell marker/s	Primary antibody	Cut-off point (%)	Positive cases (%)	Follow up in months Min	Median	Main results
Jiang et al[9]	Stage 1 NSCLC	148	TMA	ALDH1	Rabbit polyclonal (Santa cruz)	>10	29	NM		ALDH1 positive patients have worse OS (p=0.006)
Salnikov et al[32]	Stage I–III NSCLC	81	Whole section	CD133	Rabbit polyclonal (Abcam)	Any*	63	NM		CD133 expression did not correlate with survival.
Sullivan et al[31]	All stages NSCLC	282	TMA	ALDH1A1	Rabbit monoclonal (Abcam)	Median	51	NM		ALDH1A1 high expression was associated with poor survival (p=0.025)
				ALDH3A1	Santa Cruz	Median	50			No association between ALDH3A1 and CD133 expression and survival
				CD133	Mouse monoclonal (Miltenyi)	Any	27			
Woo et al[33]	Stage 1 ADC	177	Whole sections	CD133	Mouse monoclonal (Miltenyi)	>17.5	45.8	1.1	35.9	CD133 high expressers was associated with worse DFS (p=0.004)
Okudela et al[34]	Stage 1 ADC	177	Whole sections	ALDH1	Rabbit monoclonal (Abcam)	>85	20.3	1.1	35.9	High expression of ALDH1, CD133 and CD44 was associated with worse DFS (p=0.036, 0.021, 0.023 respectively)
				CD133	Mouse monoclonal (Miltenyi)	>17.5	45.8			
Li et al[35]	All stages NSCLC	179	TMA	ALDH1A1	Rabbit monoclonal (Abcam)	Any	45	6	49.2	ALDH1A1 positive patients had shorter survival
Current study	Stage 1 NSCLC	205	Whole sections	ALDH1A1	Rabbit monoclonal (Abcam) Mouse monoclonal (BD)	>10	69	18	60	ALDH1A1 and CD133 overexpression is associated with shorter OS (p=0.017 and 0.039, respectively)
				CD133	Rabbit polyclonal (Abcam) Rabbit polyclonal (Bioss)	>5	50.7			Patients expressing both ALDH1A1 and CD133 had worse survival (p=0.017)

*Any=any positive staining was considered positive, regardless of percentage or intensity.
ADC, adenocarcinoma; ALDH1, aldehyde dehydrogenase 1; ALDH1A1, aldehyde dehydrogenase 1A1; DFS, disease-free survival; Min, minimum follow-up; NM, not mentioned; OS, overall survival; TMA, tissue microarray.

ALDH1A1 as a prognostic marker in lung cancer. By contrast, our results differ with respect to CD133 expression. This could be best explained by the variation in the patient's population, or in the specificity of the antibodies used in IHC. Currently available commercial antibodies may not be able to adequately discriminate the most specific epitope of CD133, emphasising the need to investigate the multiple epitopes and their function in maintaining stem cell-like phenotype.

According to the stem cell hypothesis, a small subgroup of specialised cells recapitulates the entire tumour after initial treatment, which eventually results in treatment failure. Clinical validation of stem cell hypothesis is best examined in a uniform population, reducing the impact of other prognostic factors like stage, nodal status or by adjuvant treatments. Although our results support this hypothesis, their variable prognostic role in different histological groups suggests that the expressions of CD133 and ALDH1A1 represent distinct markers of stem-like function rather than a single stem cell. Nevertheless, the expression of markers known to be functionally associated with stem-like phenotypes may well be a useful approach that can prospectively identify patients at high risk of recurrence following resection of early-stage lung cancer ADC.

REFERENCES

1 Kelsey CR, Marks LB, Hollis D, et al. Local recurrence after surgery for early stage lung cancer: an 11-year experience with 975 patients. Cancer 2009;115:5218–27.

2 Rami-Porta R, Ball D, Crowley J, et al. The IASLC lung cancer staging project: proposals for the revision of the T descriptors in the forthcoming (seventh) edition of the TNM classification for lung cancer. J Thorac Oncol 2007;2:593–602.

3 Pignon JP, Tribodet H, Scagliotti GV, et al. Lung adjuvant cisplatin evaluation: a pooled analysis by the LACE Collaborative Group. J Clin Oncol 2008;26:3552–9.

4 Arriagada R, Dunant A, Pignon JP, et al. Long-term results of the international adjuvant lung cancer trial evaluating adjuvant Cisplatin-based chemotherapy in resected lung cancer. J Clin Oncol 2010;28:35–42.

5 Jiang J, Liang X, Zhou X, et al. ERCC1 expression as a prognostic and predictive factor in patients with non-small cell lung cancer: a meta-analysis. Mol Biol Rep 2012;39:6933–42.

6 Rosell R, Skrzypski M, Jassem E, et al. BRCA1: a novel prognostic factor in resected non-small-cell lung cancer. PLoS One 2007;2:e1129.

7 Zheng Z, Chen T, Li X, et al. DNA synthesis and repair genes RRM1 and ERCC1 in lung cancer. N Engl J Med 2007;356:800–8.

8 Reya T, Morrison SJ, Clarke MF, et al. Stem cells, cancer, and cancer stem cells. Nature 2001;414:105–11.

9 Jiang F, Qiu Q, Khanna A, et al. Aldehyde dehydrogenase 1 is a tumor stem cell-associated marker in lung cancer. Mol Cancer Res 2009;7:330–8.

10 Eramo A, Lotti F, Sette G, et al. Identification and expansion of the tumorigenic lung cancer stem cell population. Cell Death Differ 2008;15:504–14.

11 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;1:555–67.

12 Travis WD, Brambilla E, Noguchi M, et al. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. J Thorac Oncol 2011;6:244–85.

13 Barker N, van Es JH, Kuipers J, et al. Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature 2007;449:1003–7.

14 Hilsenbeck SG, Clark GM, McGuire WL. Why do so many prognostic factors fail to pan out? Breast Cancer Res Treat 1992;22:197–206.

15 Miyoshi K, Moriyama S, Kunitomo T, et al. Prognostic impact of intratumoral vessel invasion in completely resected pathologic stage I non-small cell lung cancer. J Thorac Cardiovasc Surg 2009;137:429–34.

16 Hess DA, Meyerrose TE, Wirthlin L, et al. Functional characterization of highly purified human hematopoietic repopulating cells isolated according to aldehyde dehydrogenase activity. Blood 2004;104:1648–55.

17 Hess DA, Wirthlin L, Craft TP, et al. Selection based on CD133 and high aldehyde dehydrogenase activity isolates long-term reconstituting human hematopoietic stem cells. Blood 2006;107:2162–9.

18 Chute JP, Muramoto GG, Whitesides J, et al. Inhibition of aldehyde dehydrogenase and retinoid signaling induces the expansion of human hematopoietic stem cells. Proc Natl Acad Sci USA 2006;103:11707–12.

19 Wu RP, Li H, Rassenti LZ, et al. Increased aldehyde dehydrogenase activity in high-risk chronic lymphocytic leukemia. Leuk Lymphoma 2013;54:400–2.

20 Pearce DJ, Taussig D, Simpson C, et al. Characterization of cells with a high aldehyde dehydrogenase activity from cord blood and acute myeloid leukemia samples. Stem Cells 2005;23:752–60.

21 Liang D, Shi Y. Aldehyde dehydrogenase-1 is a specific marker for stem cells in human lung adenocarcinoma. Med Oncol 2012;29:633–9.

22 Miraglia S, Godfrey W, Yin AH, et al. A novel five-transmembrane hematopoietic stem cell antigen: isolation, characterization, and molecular cloning. Blood 1997;90:5013–21.

23 Uchida N, Buck DW, He D, et al. Direct isolation of human central nervous system stem cells. Proc Natl Acad Sci USA 2000;97:14720–5.

24 Sagrinati C, Netti GS, Mazzinghi B, et al. Isolation and characterization of multipotent progenitor cells from the Bowman's capsule of adult human kidneys. J Am Soc Nephrol 2006;17:2443–56.

25 Singh SK, Hawkins C, Clarke ID, et al. Identification of human brain tumour initiating cells. Nature 2004;432:396–401.

26 Collins AT, Berry PA, Hyde C, et al. Prospective identification of tumorigenic prostate cancer stem cells. Cancer Research 2005;65:10946–51.

27 Suetsugu A, Nagaki M, Aoki H, et al. Characterization of CD133+ hepatocellular carcinoma cells as cancer stem/progenitor cells. Biochem Biophys Res Commun 2006;351:820–4.

28 Hermann PC, Huber SL, Herrler T, et al. Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. Cell Stem Cell 2007;1:313–23.

29 O'Brien CA, Pollett A, Gallinger S, et al. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. Nature 2007;445:106–10.

30 Russell PA, Wainer Z, Wright GM, et al. Does lung adenocarcinoma subtype predict patient survival?: a clinicopathologic study based on the new International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multidisciplinary lung adenocarcinoma classification. J Thorac Oncol 2011;6:1496–504.

31 Sullivan JP, Spinola M, Dodge M, et al. Aldehyde dehydrogenase activity selects for lung adenocarcinoma stem cells dependent on notch signaling. Cancer Research 2010;70:9937–48.

32 Salnikov AV, Gladkich J, Moldenhauer G, et al. CD133 is indicative for a resistance phenotype but does not represent a prognostic marker for survival of non-small cell lung cancer patients. Int J Cancer 2010;126:950–8.

33 Woo T, Okudela K, Mitsui H, et al. Prognostic value of CD133 expression in stage I lung adenocarcinomas. Int J Clin Exp Pathol 2010;4:32–42.

34 Okudela K, Woo T, Mitsui H, et al. Expression of the potential cancer stem cell markers, CD133, CD44, ALDH1, and beta-catenin, in primary lung adenocarcinoma-their prognostic significance. Pathol Int 2012;62:792–801.

35 Li X, Wan L, Geng J, et al. Aldehyde dehydrogenase 1A1 possesses stem-like properties and predicts lung cancer patient outcome. J Thorac Oncol 2012;7:1235–45.