Clinical Cancer Research

Stromal CD8$^+$ T-cell Density—A Promising Supplement to TNM Staging in Non–Small Cell Lung Cancer

Tom Donnem[1,2], Sigurd M. Hald[2], Erna-Elise Paulsen[1,2], Elin Richardsen[3,4], Samer Al-Saad[3,4], Thomas K. Kilvaer[1], Odd Terje Brustugun[5,6], Aslaug Helland[5,6], Marius Lund-Iversen[6], Mette Poehl[7, 8, 9], Karen Ege Olsen[9,10], Henrik J. Ditzel[8,11], Olfred Hansen[8,9], Khalid Al-Shibli[12], Yury Kiselev[4,13], Torkjel M. Sandanger[14], Sigve Andersen[1], Francesco Pezzella[15], Roy M. Bremnes[1,2], and Lill-Tove Busund[3,4]

Abstract

Purpose: Immunoscore is a prognostic tool defined to quantify *in situ* immune cell infiltrates, which appears to be superior to the tumor–node–metastasis (TNM) classification in colorectal cancer. In non–small cell lung cancer (NSCLC), no immunoscore has been established, but *in situ* tumor immunology is recognized as highly important. We have previously evaluated the prognostic impact of several immunological markers in NSCLC, yielding the density of stromal CD8$^+$ tumor-infiltrating lymphocytes (TIL) as the most promising candidate. Hence, we validate the impact of stromal CD8$^+$ TIL density as an immunoscore in NSCLC.

Experimental Design: The prognostic impact of stromal CD8$^+$ TILs was evaluated in four different cohorts from Norway and Denmark consisting of 797 stage I–IIIA NSCLC patients. The Tromso cohort ($n = 155$) was used as training set, and the results were further validated in the cohorts from Bodo ($n = 169$), Oslo ($n = 295$), and Denmark ($n = 178$). Tissue microarrays and clinical routine CD8 staining were used for all cohorts.

Results: Stromal CD8$^+$ TIL density was an independent prognostic factor in the total material ($n = 797$) regardless of the endpoint: disease-free survival ($P < 0.001$), disease-specific survival ($P < 0.001$), or overall survival ($P < 0.001$). Subgroup analyses revealed significant prognostic impact of stromal CD8$^+$ TIL density within each pathologic stage (pStage). In multivariate analysis, stromal CD8$^+$ TIL density and pStage were independent prognostic variables.

Conclusions: Stromal CD8$^+$ TIL density has independent prognostic impact in resected NSCLC, adds prognostic impact within each pStage, and is a good candidate marker for establishing a TNM-Immunoscore. *Clin Cancer Res; 21(11); 2635–43. ©2015 AACR.*

[1]Department of Oncology, University Hospital of North Norway, Tromso, Norway. [2]Institute of Clinical Medicine, The Arctic University of Norway, Tromso, Norway. [3]Department of Clinical Pathology, University Hospital of North Norway, Tromso, Norway. [4]Institute of Medical Biology, The Arctic University of Norway, Tromso, Norway. [5]Department of Oncology, Oslo University Hospital, The Norwegian Radium Hospital, Oslo, Norway. [6]Institute of Clinical Medicine, University of Oslo, Oslo, Norway. [7]Department of Oncology, Rigshospitalet, Copenhagen, Denmark. [8]Department of Oncology, Odense University Hospital, Odense, Denmark. [9]Institute of Clinical Research, University of Southern Denmark, Odense, Denmark. [10]Department of Pathology, Odense University Hospital, Odense, Denmark. [11]Institute of Molecular Medicine, University of Southern Denmark, Odense, Denmark. [12]Department of Pathology, Nordland Hospital, Bodo, Norway. [13]Department of Pharmacy, The Arctic University of Tromso, Tromso, Norway. [14]Department of Community Medicine, The Artic University of Tromso, Tromso, Norway. [15]Nuffield Department of Clinical Laboratory Sciences, University of Oxford, John Radcliffe Hospital, Oxford, United Kingdom.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

R.M. Bremnes and L.-T. Busund contributed equally to this article.

Corresponding Author: Tom Donnem, Department of Oncology/Institute of Clinical Medicine, University Hospital of North Norway/The Arctic University of Norway, 9038 Tromso, Norway. Phone: 47-91145107; Fax: 47-77626779; E-mail: tom.donnem@uit.no

doi: 10.1158/1078-0432.CCR-14-1905

Introduction

In non–small cell lung cancer (NSCLC), staging according to the tumor–node–metastasis (TNM) system and histologic subtype has been the most relevant clinicopathologic variables for routine prognostication and treatment (1). TNM summarizes data on the extent of tumor burden (T), spread to regional lymph nodes (N) and evidence of metastasis (M). The American Joint Committee of Cancer and the Union Internationale Contre le Cancer (UICC) define and periodically update the TNM classification system (2). In lung cancer, the current 7th edition of the lung cancer staging system is based on an initiative undertaken by the International Association for the Study of Lung Cancer (IASLC; ref. 1). The TNM staging system has shown its value, but provides incomplete prognostic information as clinical outcomes vary among patients in the same TNM stage.

Hence, new strategies to improve outcome prediction and treatment selection are warranted. Extensive literature has demonstrated highly prognostic impact by *in situ* immune cell infiltrates in tumors. In colorectal cancer, a combination of the density of tumor-infiltrating lymphocytes (TIL; CD3$^+$ T cells, pan-lymphocyte marker; CD45RO$^+$, memory T cells; CD8$^+$ T cells, cytotoxic T cells, CTL) and location (central tumor vs. invasive margins) has defined the Immunoscore as a new component for the classification of colorectal cancer, designated TNM-Immune

Translational Relevance

The tumor–node–metastasis (TNM) staging system provides the most reliable guidelines for the routine prognostication and treatment of non–small cell lung cancer (NSCLC). However, several *in situ* immunological studies show tumor-infiltrating lymphocytes to be of vital importance in suppressing NSCLC development. In colorectal cancer, the presence of $CD8^+$ cytotoxic T cells seems to be a useful prognostic marker, and large ongoing studies aim at implementing a combination of Immunoscore with TNM classification, TNM-I, in a routine clinical setting. This study validates one of the most promising *in situ* immunological markers in NSCLC, stromal $CD8^+$ density, in tumor tissues of four different cohorts of Norwegian and Danish NSCLC patients. We conclude that stromal $CD8^+$ density adds significant prognostic impact to the established TNM staging in stage I–IIIA NSCLC, hence is a promising candidate in establishing an NSCLC TNM-I.

(TNM-I). An international consortium has been initiated to validate and promote the Immunoscore in the routine clinical colorectal cancer setting (3–6).

The CD8 marker can be found on natural killer and dendritic cells, but is predominantly expressed on the surface on cytotoxic T cells. $CD8^+$ T cells are a crucial component of the cellular immune system and are pivotal in cell-mediated antitumor immune responses (7, 8). $CD8^+$ T cells undergo a period of massive expansion, activation, and differentiation to cytotoxic T cells with effector functions. Once the pathogenic process is resolved, most effector $CD8^+$ cells undergo apoptosis, leaving a subset of long-lived memory cells (9, 10). The association of immune cell infiltrates and prognosis has been reviewed in different malignancies, and in a majority of the studies, memory T cells and cytotoxic T cells are predictive of a favorable clinical outcome (8). In NSCLC, we and others have found $CD8^+$ T cells to have significant prognostic effect alone or in combination with other markers (11–16).

Although our group has studied more than 100 prognostic markers related to *in situ* tumor immunology (11, 13, 14,17, 18), angiogenesis (19–21), and epithelial–mesenchymal transition (22) in NSCLC, stromal $CD8^+$ score ($CD8^+$ TIL density in the tumor-related stroma) has been one of the markers with the strongest prognostic impact independent of other clinicopathologic variables including TNM classification. Furthermore, cytotoxic $CD8^+$ T cells are of increasing interest in lung cancer as novel targeted therapies are aiming to stimulate the immune defense by blocking mechanisms that may inhibit cytotoxic T lymphocytes, e.g., programmed death-1 (PD1), PD-ligand 1 (PD-L1), and cytotoxic T-lymphocyte–associated antigen 4 (CTLA-4; ref. 23).

In the interest of establishing an Immunoscore in resected NSCLC patients, we evaluated our most promising candidate marker, stromal $CD8^+$ TIL density, in four different NSCLC cohorts (three Norwegian and one Danish). In the Danish validation set, the impact of tissue microarray (TMA) core localization (central vs. peripheral tumor areas) has been explored (24), allowing us to evaluate the impact of stromal $CD8^+$ TILs localization.

Materials and Methods

Study design and patients

Patients surgically resected for stage I–IIIA NSCLC at the University Hospital of North Norway (UNN, Tromso) and Nordland Hospital (NH, Bodo) from 1990 through 2004 were included in this study (19). In the original cohort, 335 patients were included from these two hospitals and results on immunological markers reported in 2008 (11). However, patients have been restaged according to the revised 7th edition of UICC TNM classification and reclassified according to the new pathologic classification of lung cancer, excluding 11 previous bronchioalveolar carcinomas (25, 26). In addition, both cohorts now include follow-up data as of January 2011, hence leaving 155 patients from Tromso and 169 from Bodo. In this study, the Tromso cohort was used as a training cohort to define the optimal prognostic cutoffs for stromal CD8 density. In contrast with the exploratory approach in the Tromso cohort, the cutoffs from the Tromso cohort were set as a predefined scoring system in the validation cohorts (Bodo, Oslo, and Denmark).

Two additional validation cohorts of consecutive NSCLC stage I–IIIA patients from Oslo University Hospital, the Norwegian Radium Hospital (OUS; 2006–2011, $n = 295$), and Odense University Hospital, Denmark (1992–1999, $n = 178$), were used.

Exclusion criteria for patients in either the training or validation cohorts were as follows: (i) inadequate paraffin-embedded fixed tissue blocks; (ii) other malignancy within 5 years before NSCLC diagnosis; (iii) radiotherapy or chemotherapy before surgery; and (iv) or less than two cores with tumor stroma evaluable. All study cohorts were approved by the Regional Committee for Medical and Health Research Ethics (Tromso and Bodo: protocol ID: 2011/2503; Oslo: protocol ID 2009/1904; Denmark: protocol ID: 20080018). The reporting of clinicopathologic variables, survival data, and biomarker expression was conducted in accordance with the REMARK guidelines (27).

TMA construction

Lung cancer specimens were histologically reviewed by experienced pathologists. In all cohorts, representative areas of vital tumor tissue were carefully selected and marked on the hematoxylin and eosin stain (H&E) slide and sampled for the TMA blocks. Regardless of cohort, the large majority of the cores showed a mixture of epithelial tumor tissue and tumor-related stroma (Fig. 1). In the Danish cohort, designated cores from the tumor center and from the invasive margin were made (Fig. 1B). In the Norwegian cohorts, the localization of the cores (central tumor vs. invasive margin) was not noted. TMAs were made using a tissue arrayer instrument (Beecher Instruments). The detailed methodology has been previously reported (19, 24). Tromso and Bodo used 0.6-mm cores, whereas Oslo and Denmark used 1.0-mm cores. In all cohorts, multiple 4-μm sections were cut with a Micron microtome before antibody staining for immunohistochemical analysis (Table 1).

Immunohistochemistry

$CD8^+$ staining is well validated and in routine clinical use at all included hospitals.

In the Norwegian cohorts, Ventana Benchmark, XT automated slide stainer (Ventana Medical System), was used for immunohistochemistry (IHC). For the Tromso/Bodo cohorts, prediluted primary CD8 antibody from Ventana Medical Systems/Roche

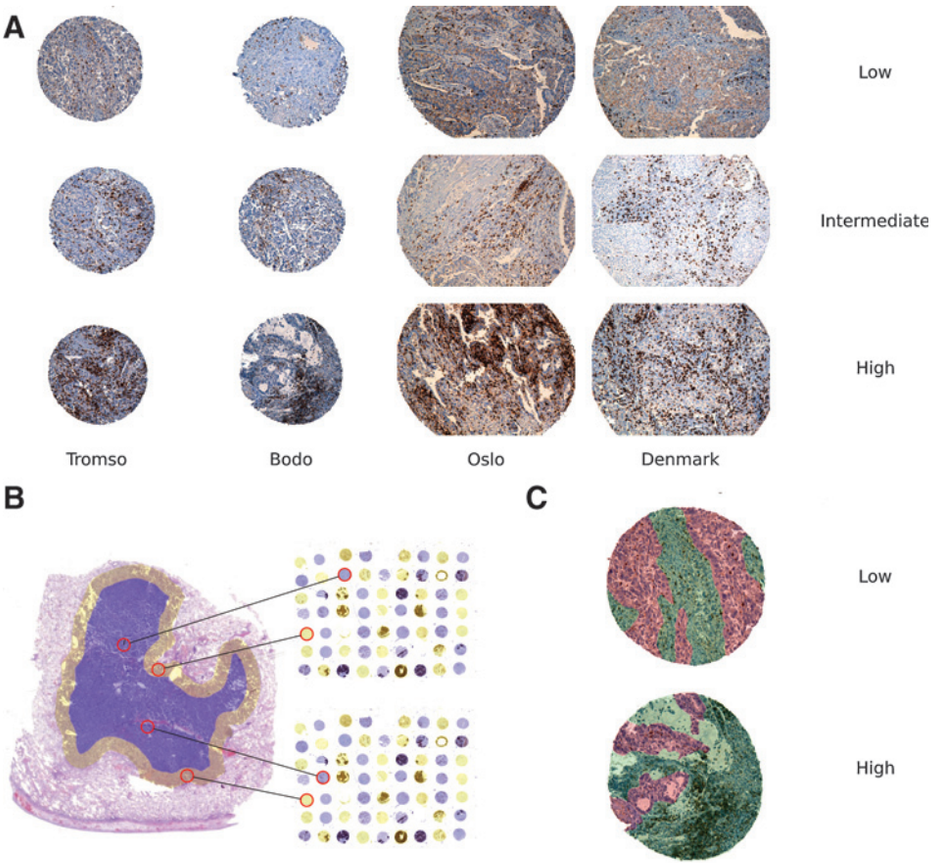

Figure 1.
A, low, intermediate, and high stromal CD8$^+$ density in the Tromso (6-mm cores), Bodo (6-mm cores), Oslo (1.0-mm cores), and the Danish cohort (1.0-mm cores). B, in the Danish cohort, designated cores from the tumor center (blue) and from the invasive margin (yellow) were made. In the Norwegian cohorts, the localization of the cores (central tumor vs. invasive margin) was not noted. C, in all cohorts, the large majority of the cores showed a mixture of epithelial tumor tissue (red) and tumor-related stroma (green). Within each core, the percentages of CD8$^+$ lymphocytes compared with the total amount of nucleated cells in the stromal compartments (green areas) were assessed and defined as stromal CD8$^+$ density. Low density: $\leq$25%; intermediate density: >25% to $\leq$ 50%; high density: >50%.

(clone C8/144, 760-4250) was used, and for the Oslo cohort, prediluted primary CD8 antibody from Roche (clone SP57, 790-4460) was used. The slides were baked for 60°C overnight, deparaffinized in EZ Prep, and submitted to heat-induced epitope retrieval. Cell conditioning-1 protocol (CC1 Mild) for 30 minutes at 95°C epitope retrieval was used for Tromso/Bodo cohorts and CC1 Standard for 60 minutes at 95°C for the Oslo cohort. The endogenous peroxidase activity was quenched using 3% hydrogen peroxide. The primary antibody was applied and the slides were incubated for 32 minutes at 36°C, followed by washing in buffer and visualization. Visualization method was carried out with the Iview DAB Detection Kit for Tromso/Bodo cohorts and the ultraview DAB Detection Kit for the Oslo cohort. 3,3′-diaminobenzidine was used as a chromogen for all cohorts.

In the Danish cohort, Ventana Benchmark, Ultra automated slide stainer (Ventana Medical System), was used. The primary antibody CD8 was from Dako M7103 (clone C8/144B). The slides were baked at 75°C for 4 minutes, deparaffinized in EZ Prep, and submitted to heat-induced epitope retrieval. CC1 Mild for 36 minutes at pH 8.5 and 99°C was used. The endogenous peroxidase activity was quenched using 3% hydrogen peroxide. The primary antibody dilution was applied and the slides were incubated for 32 minutes at 36°C, followed by washing in buffer and visualization using the OptiView DAB Detection Kit with 3,3′-diaminobenzidine as a chromogen.

Negative and positive staining controls were placed simultaneously with all TMAs. In all cohorts, slides were counterstained with hematoxylin to visualize the nuclei. For each cohort, staining was done in one single experiment.

To validate different IHC procedures, new TMA slides from Denmark were stained and reevaluated with the same IHC procedure as for the Oslo cohort.

Scoring of IHC

In all cohorts, samples were anonymized and independently scored by two experienced pathologists in the Norwegian cohorts (K. Al-Shibli and S. Al-Saad in the Tromso and Bodo cohorts; S. Al-Saad and E. Richardsen in the Oslo cohort) and one experienced pathologist (K.E. Olsen) and one oncologist (M. Poehl) in the Danish cohort. The second scoring of the Danish cohort (validating the staining procedure/scoring) was done by S. Al-Saad and E. Richardsen. In case of disagreement, the slides were reexamined and the observers reached a consensus. By light microscopy, the tissue sections were scored for the degree of infiltration of CD8$^+$ lymphocytes. The percentages of CD8$^+$ lymphocytes compared with the total amount of nucleated cells in the stromal compartments were assessed. Based on experiences from one of our previous studies (11), scoring cutoff points at 5%, 25%, or 50% for each core according to the degree of cell densities were used in the training set (Tromso cohort): 0% to 5% = 0, 5% to 25% = 1, 25% to 50% = 2, and

Table 1. Clinicopathologic characteristics in each cohort and in the total material

	Training cohort Tromso, Norway (UNN)	**Validation cohort Bodo, Norway (NH)**	**Validation cohort Oslo, Norway (OUS)**	**Validation cohort Odense, Denmark**	**Total cohort**
Number of patients	155	169	295	178	797
Time of inclusion	1990–2005	1990–2005	2006–2011	1992–1999	1990–2011
Median age in years	66.6 (range, 38.8–84.7)	67.3 (range, 27.5–82.1)	66.5 (range, 39.1–84.1)	64.1 (range, 39.4–82.4)	65.2 (range, 27.5–84.7)
Last follow-up	January 2011	January 2011	March 2014	January 2010	January 2010–March 2014
Median follow-up time of the survivors in months	110.7 (range, 76.3–222.2)	102.8 (range, 72.9–234.0)	52.4 (range, 34.9–99.4)	161.7 (range, 120.6–210.8)	65.5 (range, 34.9–34.0)
Scoring system	*Exploratory* Different cutoff system tested, and both average and maximum score evaluated. Optimal prognostic impact cutoff defined and used in all cohorts.	*Predefined* Maximum score Low density: ≤25% Intermediate density: >25%–≤ 50% High density: >50%	*Predefined* Maximum score Low density: ≤25% Intermediate density: >25%–≤ 50% High density: >50%	*Predefined* Maximum score Low density: ≤25% Intermediate density: >25%–≤50% High density: >50%	*Predefined* Maximum score Low density: ≤25% Intermediate density: >25%–≤ 50% High density: >50%
Endpoints	DSS, DFS, OS	DSS, DFS, OS	DSS, DFS, OS	DSS, DFS, OS	DSS, DFS, OS
TMA core size	0.6 mm	0.6 mm	1.0 mm	1.0 mm	0.6–1.0 mm
$CD8^+$ IHC staining	Clinical routine	Clinical routine	Clinical routine	Clinical routine	Clinical routine
Number of cores from each patients	Four	Four	Variable, minimum of three	Four, two from central and two from invasive margins	Variable, but at least potentially three evaluable from each patient.
TMA thickness	4 μm	4 μm	4 μm	4 μm	4 μm
Distribution stromal $CD8^+$ density score (%) 1: low, 2: intermediate, 3: high	1: 35 (23) 2: 80 (52) 3: 39 (25) Missing: 1	1: 47 (28) 2: 90 (53) 3: 29 (17) Missing: 3	1: 77 (27) 2: 156 (53) 3: 56 (19) Missing: 6	1: 50 (28) 2: 77 (43) 3: 47 (26) Missing: 4	1: 209 (27) 2: 403 (52) 3: 171 (22) Missing: 14

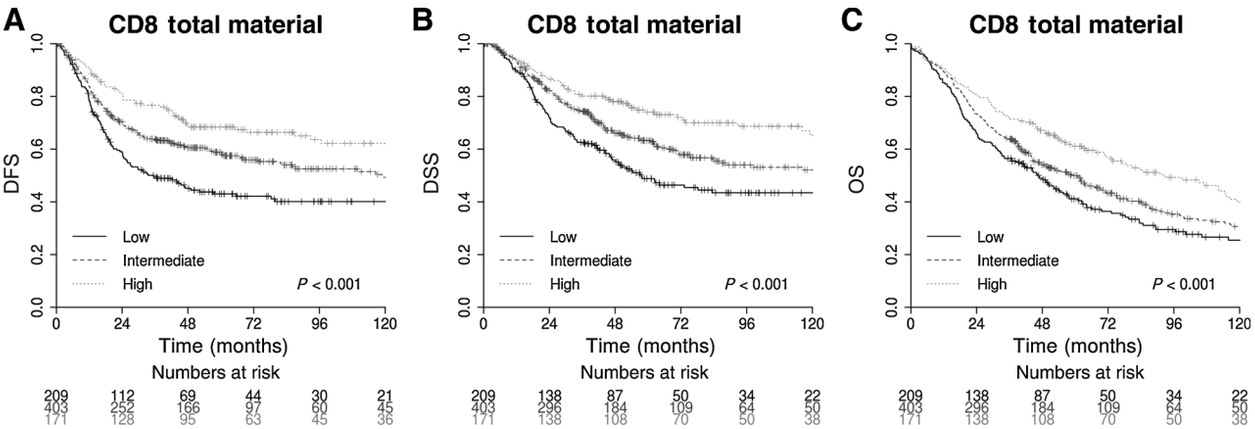

Figure 2.
DFS (A), DSS (B), and OS (C) curves according to low, intermediate, and high stromal CD8$^+$ T-cell density in the total material. The maximum score for each patient was used.

>50% = 3. In this training cohort, only 2 patients had an average or maximum score evaluated as 0; hence, scores 0 and 1 were merged and the following cutoffs were used in the validation sets: low density: ≤25%; intermediate density: >25% to ≤50%; high density: >50%. These percentages are also easy to follow and reproduce in daily practice. Both the average and the maximum score from each patient were assessed in the training set, resulting in an optimal significant prognostic impact ($P = 0.004$, endpoint DSS), when using a maximum score. Hence, the maximum score was used in the validation sets.

Statistical analysis

The Kaplan–Meier method was used to analyze the association between marker expression and survival endpoints. Disease-specific survival (DSS) was determined from the date of surgery to the time of lung cancer death, overall survival (OS) from the date of surgery to death, and disease-free survival (DFS) as the time from the date of surgery to time of first relapse. The statistical significance of differences between survival curves was assessed with the log-rank test. The χ^2 test was used to examine the association between molecular marker expression and various clinicopathologic variables. In the second IHC score of the Danish cohort, each observer was compared for interobserver reliability by use of a two-way random effect model with absolute agreement definition. The intraclass correlation coefficient (reliability coefficient, ICC) was obtained from these results. Variables with a P value of <0.25 from univariate analyses entered the multivariate analysis, using the Cox proportional hazards model (backward stepwise, probability for stepwise entry and removal was set at 0.05 and 0.10). P values < 0.05 were considered statistically significant.

Results

Patient characteristics

Demographic, clinical, and histopathologic variables for each cohort and in the total material are listed in Table 1 and Supplementary Tables S1–S3. Median age in the total material was 65.2 (range, 27.5–84.7) years (Table 1). The majority of patients were men (64%), but there was a higher rate of women in the cohort with the most recent inclusion period (Oslo, 49% women). The Oslo cohort also had the highest frequency of adenocarcinomas (63%) compared with 48% in the total population.

Marker expression and correlations

The stromal CD8$^+$ TIL density score distribution in each cohort is shown in Table 1. There was no significant difference in the stromal CD8$^+$ density score (low vs. intermediate vs. high) between the different cohorts ($P = 0.18$). In the total material, there was no significant correlation between stromal CD8$^+$ density and the clinicopathologic variables.

Univariate analyses

As shown in Supplementary Table S1, tumor differentiation ($P < 0.001$), pStage ($P < 0.001$), T stage ($P < 0.001$), and N stage ($P < 0.001$) were significant prognostic indicators for DSS in the total material (Supplementary Tables S2 and S3 show the results with DFS and OS as endpoints). Figure 2 presents the prognostic impact of stromal CD8$^+$ density in the total material with three different endpoints: DFS ($P < 0.001$), DSS ($P < 0.001$), and OS ($P < 0.001$). Five-year survival rates for stromal CD8$^+$ high, intermediate, and low score in the total material were as follows: DFS: 68%, 59%, and 43%; DSS: 74%, 63%, and 49%; OS: 61%, 50%, and 41%, respectively. The prognostic impact in each cohort with DFS and DSS as endpoint is shown in Fig. 3. Using OS as endpoint, the prognostic impact of stromal CD8 density in each cohort was as follows: Tromso ($P = 0.004$), Bodo ($P = 0.28$), Oslo ($P = 0.22$), and Odense ($P = 0.11$), all showing the same tendency as with DFS and DSS as endpoints.

The frequencies of recurrences, lung cancer–related deaths, and overall deaths in each cohort were as follows: Tromso: 51%, 45%, and 79%; Bodo: 42%, 41%, and 82%; Oslo: 37%, 28%, and 41%; Denmark: 52%, 51%, and 82%, respectively.

When stratified by histology and pStage, CD8$^+$ density had significant prognostic impact in all subgroups with DSS as endpoint (Table 2). Using OS as endpoint, all subgroups showed significant prognostic impact, except only showing a tendency in pStage II ($P = 0.09$), adenocarcinomas ($P = 0.09$), and large-cell carcinoma (LCC; $P = 0.06$). Using DFS as endpoint, stromal

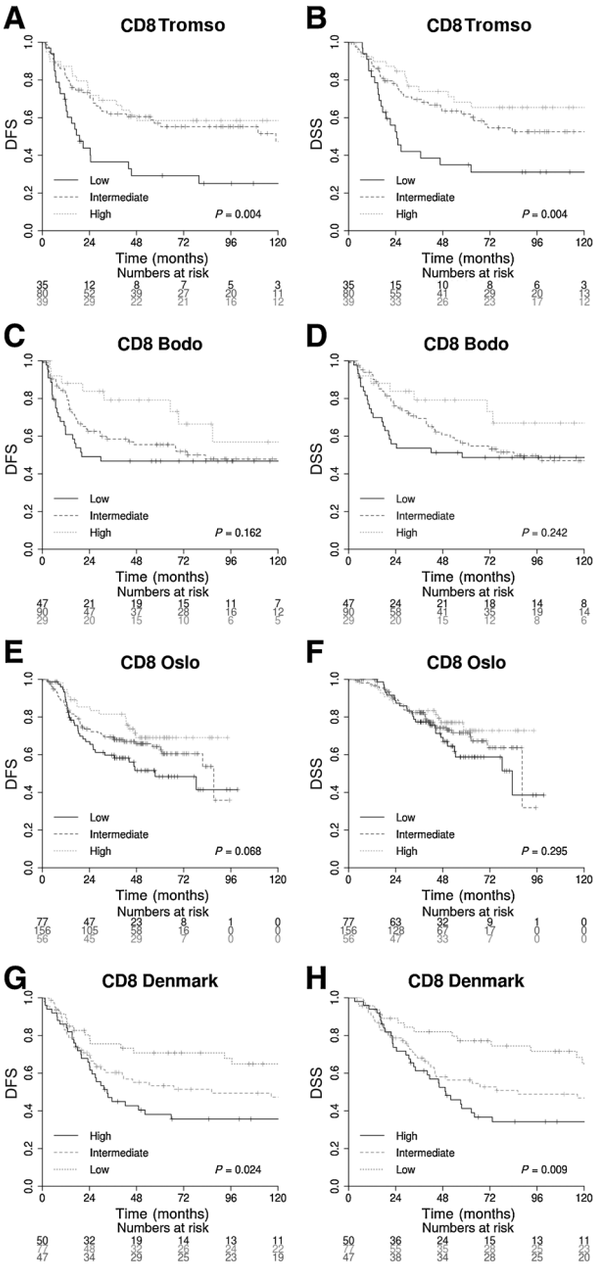

Figure 3.
DFS and DSS curves according to low, intermediate, and high stromal CD8$^+$ T-cell density in the four different cohorts; A and B, University Hospital of North Norway (Tromso, $n = 155$); C and D, Central Nordland Hospital (Bodo, $n = 169$); E and F, Oslo University Hospital (Oslo, $n = 295$); G and H, Danish cohort (Odense University Hospital, $n = 178$). The maximum score for each patient was used.

CD8$^+$ density showed significant prognostic impact in all subgroups, except within pStage I ($P = 0.07$) and LCC ($P = 0.29$).

In the Danish population, we were able to stratify the stromal CD8 density according to the central parts of the tumor versus the tumor periphery/invasive margins. As shown in Supplementary Fig. S1, the prognostic impact of stromal CD8$^+$ density with DSS as endpoint was highly significant at the tumor margins ($P = 0.008$) in contrast with the central tumor ($P = 0.67$). CD8$^+$ density in the tumor periphery/invasive margins showed 5-year survival rates of 80% (high density), 61% (intermediate density), and 42% (low density). The prognostic impact of stromal CD8 density in the invasive margin was also significant for DFS ($P = 0.021$) and tended toward significance for OS ($P = 0.09$).

Multivariate analysis

Results from the multivariate analysis are presented in Table 3. All clinicopathologic variables with $P < 0.25$ and stromal CD8$^+$ density from the univariate analyses were entered into the multivariate analysis. However, as pStage is based on a combination of T stage and N stage, multivariate analysis was done including pStage and not T stage or N stage (Table 3). Pathologic stage, differentiation, and stromal CD8$^+$ density were identified as independent prognostic factors regardless of endpoints (DSS, DFS, or OS).

Validating the Danish cohort

TMAs from the Danish cohort were reevaluated with the same IHC procedure and scored by the same pathologists as for the Oslo cohort. The interobserver reliability coefficient between the scorers was good (ICC = 0.89, $P > 0.001$). At a patient level (maximum score for each patient), the correlation coefficient, r, was 0.68, kappa value was 0.57, and $P < 0.001$ when comparing the results from the first and the second staining/scoring. The prognostic impact of stromal CD8$^+$ density in the validation of the Danish cohort, using the same predefined cutoffs, was similar to the original results: DFS, $P = 0.009$ versus $P = 0.024$; DSS, $P = 0.005$ versus $P = 0.009$; OS, $P = 0.066$ versus $P = 0.11$.

Discussion

In this large study of 797 NSCLC patients, stromal CD8$^+$ TIL density was a strong independent prognostic factor for DFS, DSS, and OS in the total material. There is a significant prognostic impact or at least the same trend in all four cohorts. Subgroup analyses revealed a significant prognostic impact within each pStage. In the Danish cohort, marker localization was further addressed and the prognostic impact of stromal CD8$^+$ TILs was related to an invasive margin/tumor periphery location.

The strength of this study lies in the number of cohorts and validated NSCLC patients. The cutoff was predefined for validation sets. Cutoff points of 25% and 50% are easy to follow and reproduce in daily practice, supported by the fact that the distribution of low, intermediate, and high score was similar in all cohorts and also independent of the core size. Finally, CD8 IHC staining has limited technical challenges, is well validated, and established as a routine marker at most pathologic departments.

Though CD8 is a routine clinical marker, we reevaluated the Danish population to show that the results were independent of the staining and scoring procedure. Some variation in the score is expected due to the fact that new slides provided are from a different depth of the tumor block. Even though the interobserver correlation coefficient was good, this may also to a limited degree affect the score. However, reassuringly, there was a highly significant correlation between the first and the second Danish staining procedure/scoring as well as a similar prognostic impact. However, as the prognostic impact of the intermediate group varies most between the different cohorts, one may consider a two-category approach in a clinical setting.

Table 2. The prognostic impact of stromal CD8+ density in the total material stratified by histology and pathologic stage (DSS, univariate analyses, log-rank test); the maximum score for each patient was used

	Total (n = 797, 1990–2011)			
	Number (%)	**5-year DSS (%)**	**Median (Months)**	***P***
Stromal CD8+ density, total material				**<0.001**
Low CD8+	209 (27)	49	57.6	
Intermediate CD8+	403 (51)	63	127.4	
High CD8+	171 (22)	74	191.1	
Histology				
Squamous cell carcinoma				**0.006**
Low CD8+	99 (28)	51	64.1	
Intermediate CD8+	172 (49)	71	NR	
High CD8+	78 (23)	77	NR	
Adenocarcinoma				**0.041**
Low CD8+	96 (26)	48	57.1	
Intermediate CD8+	205 (54)	57	75.4	
High CD8+	74 (20)	68	189.6	
Large cell carcinoma				**0.025**
Low CD8+	14 (24)	32	54.0	
Intermediate CD8+	26 (44)	61	98.1	
High CD8+	19 (32)	83	NR	
Pathologic stage				
Stage I				**0.028**
Low CD8+	99 (25)	65	175.2	
Intermediate CD8+	211 (53)	74	174.9	
High CD8+	88 (22)	85	191.1	
Stage II				**0.023**
Low CD8+	80 (30)	40	41.5	
Intermediate CD8+	128 (47)	58	98.1	
High CD8+	63 (23)	61	142.1	
Stage IIIA				**0.012**
Low CD8+	30 (26)	18	22.4	
Intermediate CD8+	64 (56)	40	43.1	
High CD8+	20 (18)	68	NR	

NOTE: Bold numbers are significant results.
Abbreviation: NR, not reached.

The size of the TMA cores may also be an issue, but showing the same trends regardless of core size is reassuring as we are evaluating the percentages of lymphocytes compared with the total amount of nucleated cells in the stromal compartment.

In the colorectal TNM-I, several immune cell combinations have been tried and the most promising combination so far has included CD8 and CD3 with scores from both the central parts of the tumor and the invasive margins (3–6). We have

Table 3. Results of Cox regression analysis summarizing significant independent prognostic factors in the total material (n = 797) using three different endpoints; clinicopathologic variables with $P < 0.25$ and stromal CD8+ density from univariate analyses were included, and the maximum score for each patient was used

	DSS		DFS		OS	
Prognostic factor	**HR (95% CI)**	***P***	**HR (95% CI)**	***P***	**HR (95% CI)**	***P***
Pathologic stage		**<0.001**[a]		**<0.001**[a]		**<0.001**[a]
I	1.00		1.00		1.00	
II	1.99 (1.52–2.61)	**<0.001**	1.89 (1.47–2.44)	**<0.001**	1.49 (1.22–1.82)	**<0.001**
IIIA	3.19 (2.27–4.49)	**<0.001**	2.90 (2.09–4.01)	**<0.001**	2.39 (1.83– 3.13)	**<0.001**
Differentiation		**0.001**[a]		**0.001**[a]		**0.029**[a]
Well	1.00		1.00		1.00	
Moderate	1.13 (0.73–1.74)	0.58	1.39 (0.91–2.13)	0.13	1.03 (0.76–1.38)	0.84
Poor	1.90 (1.23– 2.93)	**0.004**	2.10 (1.37–3.24)	**0.001**	1.32 (0.97–1.78)	0.14
Histology		**0.005**[a]		**0.001**[a]		NI
SCC	1.00		1.00			
AC	1.63 (1.021–2.60)	**0.041**	1.35 (0.89–2.04)	0.16		
LCC	1.10 (0.69–1.76)	0.69	0.84 (0.55–1.27)	0.39		
Age, years		NI		NS		
≤65					1.00	
>65					1.48 (1.22–1.77)	**<0.001**
Stromal CD8+ density		**<0.001**[a]		**<0.001**[a]		**<0.001**[a]
High	1.00		1.00		1.00	
Intermediate	1.48 (1.05–2.09)	**0.026**	1.37 (0.99–1.88)	0.053	1.26 (0.99–1.60)	0.065
Low	2.31 (1.61–3.31)	**<0.001**	2.27 (1.63–3.18)	**<0.001**	1.74 (1.34–2.26)	**<0.001**

NOTE: Bold numbers are significant results.
Abbreviations: AC, adenocarcinoma; CI, confidence interval; NI, not included; NS, not significant; SCC, squamous cell carcinoma.
[a]Overall significance as a prognostic factor. Age and gender were only included in OS analysis. Gender did not become significant.

previously addressed the issue of localization by examining the prognostic impact of $CD8^+$ TILs within the cancer cell nests (epithelial lymphocytes) and in the tumor-related stroma (stromal lymphocytes; ref. 11). We observed stromal $CD8^+$ density to have the most promising prognostic impact. We did not, however, distinguish between stroma centrally in the tumor versus at the tumor periphery/invasive margins. Herein, we were not able to address this issue in the three Norwegian cohorts as the stromal $CD8^+$ density localization was not registered. However, this could be examined in the Danish validation set because the localizations were recorded as central tumor versus tumor periphery/invasive margins (24). In our previous study, we found the prognostic impact of stromal $CD8^+$ density to be superior to $CD8^+$ density in the cancer nests. In this study, the Danish cohort shows that the greatest prognostic impact of stromal $CD8^+$ TIL density was observed in the tumor periphery/invasive margins. An obvious question is: would the prognostic impact of the total material would further improve if the stromal $CD8^+$ density had been scored only at the invasive margins? In our previous study, we used an average value for the cores of each patient (11). In our test set, we observed that a maximum score resulted in a slightly improved prognostic impact compared with the average score. The cores in the Norwegian cohorts are a mix from the central and the periphery of the tumors, and in most cases, we expect one core from each patient to originate from the tumor periphery. This may explain why the maximum score appears to be the best approach. Nevertheless, $CD8^+$ TIL localization in the tumor environment is an important matter that needs to be fully addressed in a prospective study.

Besides the localization of TILs, a pivotal question is which combination of immune cells should be included in an NSCLC immunoscore. The prognostic impact of other immune cells, such as B cells, natural killer cells, myeloid-derived suppressor cells, macrophages, regulatory T cells, and subsets of T-helper populations (Th2, Th17), has been shown to differ depending largely on cancer type and stage (8). In contrast, the presence of T cells, cytotoxic T cells, Th1 cells, and memory T cells has been shown to be strongly associated with beneficial clinical outcome for most cancer types (8). In our previous studies, we have explored most of these immune cells and found $CD8^+$ cells as one of the most promising candidates (11, 13,14,17, 18). In this study, we have focused on this marker to explore its robustness with the intention to include it in a subsequent prospective clinical study on NSCLC TNM-I. Based on the colorectal cancer TNM-I and previous NSCLC tumor *in situ* immunology studies, there are other promising immune markers that should also be taken into account with respect to future prospective NSCLC TNM-I studies. These are CD4 (T-helper cell), $CD45RO^+$ (memory T cells), and CD3 (pan T-cell marker; refs. 3–6, 8, 11, 17, 18, 28–30).

To establish an NSCLC TNM-I classification, the immune markers included need to add prognostic or predictive impact in a clinical setting. In this study, pStage showed 5-year survival rates of 74% (stage 1), 53% (stage II), and 39% (stage IIIA), whereas stromal $CD8^+$ density demonstrates a comparable significant prognostic impact with 5-year DSS rates of 74% (high density), 63% (intermediate density), and 49% (low density). In the Danish validation set, the stromal $CD8^+$ density in the tumor periphery/invasive margins showed 5-year survival rates of 80% (high density), 61% (intermediate density), and 42% (low density). More importantly, $CD8^+$ density appears to be a significant prognostic marker within each pStage (Table 2). For instance, among stage IIIA patients in the total material, the 5-year DSS is only 18% in low $CD8^+$ density group versus 68% in patients with high stromal $CD8^+$ density, demonstrating a large variation among patients within the same pStage. Together with pStage, histology has been the most important clinicopathologic variable in clinical NSCLC decision making. The prognostic impact of stromal $CD8^+$ density is also significant across all histologic subgroups using DSS as endpoint. The same tendency is seen with DFS and OS as endpoint.

The predictive value of $CD8^+$ has not been evaluated in this study, but as cytotoxic $CD8^+$ T cells are essential in novel therapies targeting the immune system, e.g., by blocking $CD8^+$ T cell–related ligands (PD-L1 and PD-L2) and receptors (PD1 and CTLA-4), this would be of great interest for further investigation (23, 31). As indicated by our previous study (14), the potential predictive impact of stromal $CD8^+$ density in an adjuvant NSCLC radiotherapy setting should also be examined. Further, in a recent review, it was indicated that the presence of immune cells largely reflected the underlying biology (3). Hence, immune biomarkers will be both prognostic and predictive more often than non-immuno markers as many treatment modalities affect the immune system.

Even though the stromal $CD8^+$ density seems to be reliable, reproducible, clinically relevant, and biologically meaningful, some obstacles must be considered before a prospective clinical study is initiated. TMAs as used in this study are mainly designed to reveal prognostic impact in large populations, whereas the sensitivity and specificity for each patient are crucial in a clinical setting. To better reflect the tumor localization on an individual level, there are some advantages with whole slides as for instance the localization of the cytotoxic T lymphocytes in the central tumor versus invasive margins. The scoring system should be optimized, and the utilization of digital pathology and image analysis software may be advantageous. Further, we have been studying stage I–IIIA NSCLC patients in which resected material was largely accessible. For the majority of NSCLC patients (stage IIIB/IV), only biopsy specimens will be available. In the latter group, the prognostic impact of stromal $CD8^+$ density is to our knowledge not clarified and further studies need to be done.

In colorectal cancer, a worldwide task force has started to validate TNM-I in a routine clinical setting, with a goal to implement Immunoscore in colorectal cancer staging. We propose that a similar concept may be useful for NSCLC prognostication and treatment. Stromal $CD8^+$ cytotoxic T-cell density seems to be a good candidate marker and should be included in prospective clinical trials in the effort of establishing an Immunoscore for NSCLC.

Disclosure of Potential Conflicts of Interest

H.J. Ditzel reports receiving a commercial research grant from Genmab. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: T. Donnem, S. Al-Saad, S. Andersen, R.M. Bremnes, L.-T. Busund

Development of methodology: T. Donnem, S.M. Hald, S. Al-Saad, R.M. Bremnes, L.-T. Busund

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): T. Donnem, S.M. Hald, E.-E. Paulsen, S. Al-Saad, O.T. Brustugun, A. Helland, M. Poehl, K.E. Olsen, H.J. Ditzel, O. Hansen, S. Andersen, R.M. Bremnes, L.-T. Busund

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): T. Donnem, S.M. Hald, E.-E. Paulsen, E. Richardsen, S. Al-Saad, T. Kilvaer, O.T. Brustugun, M. Lund-Iversen, M. Poehl, K.E. Olsen, S. Andersen, R.M. Bremnes

Writing, review, and/or revision of the manuscript: T. Donnem, S.M. Hald, E.-E. Paulsen, E. Richardsen, S. Al-Saad, T. Kilvaer, O.T. Brustugun, A. Helland, M. Lund-Iversen, M. Poehl, H.J. Ditzel, O. Hansen, Y. Kiselev, T.M. Sandanger, S. Andersen, F. Pezzella, R.M. Bremnes, L.-T. Busund

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): T. Donnem, S.M. Hald, E. Richardsen, S. Al-Saad, T. Kilvaer, O.T. Brustugun, A. Helland, K.E. Olsen, H.J. Ditzel, Y. Kiselev, S. Andersen, R.M. Bremnes, L.-T. Busund

Study supervision: T. Donnem, R.M. Bremnes, L.-T. Busund

Grant Support

This study was financially supported by the Norwegian Cancer Society and Northern Norway Health Region Authority.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received July 25, 2014; revised January 15, 2015; accepted February 1, 2015; published OnlineFirst February 13, 2015.

References

1. Detterbeck FC, Postmus PE, Tanoue LT. The stage classification of lung cancer: diagnosis and management of lung cancer, 3rd ed: American College Of Chest Physicians evidence-based clinical practice guidelines. Chest 2013;143:e191S–e210S.
2. Sobin LH, Compton CC. TNM seventh edition: what's new, what's changed: communication from the international union against cancer and the American Joint Committee On Cancer. Cancer 2010;116:5336–9.
3. Angell H, Galon J. From the immune contexture to the immunoscore: the role of prognostic and predictive immune markers in cancer. Curr Opin Immunol 2013;25:261–7.
4. Galon J, Pages F, Marincola FM, Angell HK, Thurin M, Lugli A, et al. Cancer classification using the immunoscore: a worldwide task force. J Transl Med 2012;10:205.
5. Galon J, Angell HK, Bedognetti D, Marincola FM. The continuum of cancer immunosurveillance: prognostic, predictive, and mechanistic signatures. Immunity 2013;39:11–26.
6. Galon J, Mlecnik B, Bindea G, Angell HK, Berger A, Lagorce C, et al. Towards the introduction of the `Immunoscore' in the classification of malignant tumours. J Pathol 2014;232:199–209.
7. Barry M, Bleackley RC. Cytotoxic T lymphocytes: all roads lead to death. Nat Rev Immunol 2002;2:401–9.
8. Fridman WH, Pages F, Sautes-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. Nat Rev Cancer 2012;12:298–306.
9. Klebanoff CA, Gattinoni L, Restifo NP. CD8+ T-cell memory in tumor immunology and immunotherapy. Immunol Rev 2006;211:214–24.
10. Restifo NP, Dudley ME, Rosenberg SA. Adoptive immunotherapy for cancer: harnessing the T cell response. Nat Rev Immunol 2012;12:269–81.
11. Al-Shibli KI, Donnem T, Al-Saad S, Persson M, Bremnes RM, Busund LT. Prognostic effect of epithelial and stromal lymphocyte infiltration in non-small cell lung cancer. Clin Cancer Res 2008;14:5220–7.
12. Dieu-Nosjean MC, Antoine M, Danel C, Heudes D, Wislez M, Poulot V, et al. Long-term survival for patients with non-small-cell lung cancer with intratumoral lymphoid structures. J Clin Oncol 2008;26:4410–7.
13. Donnem T, Al-Shibli K, Andersen S, Al-Saad S, Busund LT, Bremnes RM. Combination of low vascular endothelial growth factor A (VEGF-A)/VEGF receptor 2 expression and high lymphocyte infiltration is a strong and independent favorable prognostic factor in patients with nonsmall cell lung cancer. Cancer 2010;116:4318–25.
14. Hald SM, Bremnes RM, Al-Shibli K, Al-Saad S, Andersen S, Stenvold H, et al. CD4/CD8 co-expression shows independent prognostic impact in resected non-small cell lung cancer patients treated with adjuvant radiotherapy. Lung Cancer 2013;80:209–15.
15. Hiraoka K, Miyamoto M, Cho Y, Suzuoki M, Oshikiri T, Nakakubo Y, et al. Concurrent infiltration by CD8+ T cells and CD4+ T cells is a favourable prognostic factor in non-small-cell lung carcinoma. Br J Cancer 2006;94:275–80.
16. Kawai O, Ishii G, Kubota K, Murata Y, Naito Y, Mizuno T, et al. Predominant infiltration of macrophages and CD8(+) T Cells in cancer nests is a significant predictor of survival in stage IV nonsmall cell lung cancer. Cancer 2008;113:1387–95.
17. Al-Shibli K, Al-Saad S, Donnem T, Persson M, Bremnes RM, Busund LT. The prognostic value of intraepithelial and stromal innate immune system cells in non-small cell lung carcinoma. Histopathology 2009;55:301–12.
18. Al-Shibli K, Al-Saad S, Andersen S, Donnem T, Bremnes RM, Busund LT. The prognostic value of intraepithelial and stromal CD3-, CD117- and CD138-positive cells in non-small cell lung carcinoma. APMIS 2010;118:371–82.
19. Donnem T, Al-Saad S, Al-Shibli K, Delghandi MP, Persson M, Nilsen MN, et al. Inverse prognostic impact of angiogenic marker expression in tumor cells versus stromal cells in non small cell lung cancer. Clin Cancer Res 2007;13:6649–57.
20. Donnem T, Al-Saad S, Al-Shibli K, Andersen S, Busund LT, Bremnes RM. Prognostic impact of platelet-derived growth factors in non-small cell lung cancer tumor and stromal cells. J Thorac Oncol 2008;3:963–70.
21. Donnem T, Al-Saad S, Al-Shibli K, Busund LT, Bremnes RM. Co-expression of PDGF-B and VEGFR-3 strongly correlates with lymph node metastasis and poor survival in non-small-cell lung cancer. Ann Oncol 2010;21:223–31.
22. Al-Saad S, Al-Shibli K, Donnem T, Persson M, Bremnes RM, Busund LT. The prognostic impact of NF-kappaB p105, vimentin, E-cadherin and Par6 expression in epithelial and stromal compartment in non-small-cell lung cancer. Br J Cancer 2008;99:1476–83.
23. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer 2012;12:252–64.
24. Pohl M, Olsen KE, Holst R, Ditzel HJ, Hansen O. Tissue microarrays in non-small-cell lung cancer: reliability of immunohistochemically-determined biomarkers. Clin Lung Cancer 2014;15:222–30.
25. Rami-Porta R, Crowley JJ, Goldstraw P. The revised TNM staging system for lung cancer. Ann Thorac Cardiovasc Surg 2009;15:4–9.
26. Travis WD, Brambilla E, Riely GJ. New pathologic classification of lung cancer: relevance for clinical practice and clinical trials. J Clin Oncol 2013;31:992–1001.
27. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. Reporting recommendations for tumor marker prognostic studies (REMARK). J Natl Cancer Inst 2005;97:1180–4.
28. Bremnes RM, Al-Shibli K, Donnem T, Sirera R, Al-Saad S, Andersen S, et al. The role of tumor-infiltrating immune cells and chronic inflammation at the tumor site on cancer development, progression, and prognosis: emphasis on non-small cell lung cancer. J Thorac Oncol 2011;6:824–33.
29. Fridman WH, Dieu-Nosjean MC, Pages F, Cremer I, Damotte D, Sautes-Fridman C, et al. The immune microenvironment of human tumors: general significance and clinical impact. Cancer Microenviron 2013;6:117–22.
30. Prado-Garcia H, Aguilar-Cazares D, Flores-Vergara H, Mandoki JJ, Lopez-Gonzalez JS. Effector, memory and naive CD8+ T cells in peripheral blood and pleural effusion from lung adenocarcinoma patients. Lung Cancer 2005;47:361–71.
31. Sznol M, Chen L. Antagonist antibodies to PD-1 and B7-H1 (PD-L1) in the treatment of advanced human cancer. Clin Cancer Res 2013;19:1021–34.