CrossMark

ORIGINAL ARTICLE

Heterogeneity of CD8+ tumor-infiltrating lymphocytes in non-small-cell lung cancer: impact on patient prognostic assessments and comparison of quantification by different sampling strategies

Joseph M. Obeid[1] · Nolan A. Wages[2] · Yinin Hu[1] · Donna H. Deacon[1] · Craig L. Slingluff Jr.[1]

Abstract

Introduction Infiltration of non-small-cell lung cancer (NSCLC) by CD8+ T lymphocytes predicts improved patient survival; however, heterogeneity of intratumoral localization complicates this assessment. Strategies for tumor sampling may not accurately represent the whole tumor. We hypothesized that sampling strategies may alter the identification of tumors with high CD8 density and affect the prognostic significance.

Patients and methods Twenty-three primary NSCLC tumors were immunohistochemically stained for CD8 and were assessed using automated software with eight different sampling strategies or the whole tumor. Results of all sampling strategies were compared to the whole tumor counts (paired t tests, Pearson's r). Associations between CD8 densities and overall survival were assessed (log-rank test).

Results Counts from all eight sampling strategies significantly correlated with whole tumor counts ($p \leq 0.001$). However, the magnitude of CD8+ cell counts and categorization into high vs low infiltrate groups were affected by the sampling strategy. The most concordant values were derived from random sampling of 20 % of the tumor, a simulated core biopsy, or from sampling the tumor center. TIL infiltration was associated with survival when sampling the center ($p = 0.038$), but not the invasive margin ($p > 0.2$) or other strategies.

Conclusion Different tumor sampling strategies may yield discordant TIL density results and different stratification for risk assessment. Small biopsies may be particularly unrepresentative. Random sampling of larger tumor areas is recommended. Enumerating CD8+ T cells in the tumor center may have prognostic value.

Keywords Heterogeneity · NSCLC · Tumor-infiltrating lymphocytes · CD8 T cells · Patient survival · Tumor sampling strategies

Abbreviations
CTS   Central tumor sampling
DLA   Dense lymphoid aggregate
MAX   Largest density
MIN   Smallest density
NSCLC   Non-small-cell lung cancer
STDEV   Standard deviation
TMA   Tissue microarray
TME   Tumor microenvironment
WT   Whole tumor

✉ Craig L. Slingluff Jr.
cls8h@virginia.edu

[1] Department of Surgery, University of Virginia, Charlottesville, VA, USA

[2] Department of Public Health Sciences, University of Virginia, Charlottesville, VA, USA

Introduction

The immune system plays an important role in the host defense against malignancies. Increased CD8+ T cells in the tumor microenvironment (TME) are associated with improved survival in many cancers [1, 2] and are associated with response to chemotherapy in non-small-cell lung cancer (NSCLC) and colorectal liver metastases [3, 4]. A dominant mediator of the antitumor immune response is

believed to be tumor cell killing by cytotoxic CD8$^+$ T cells [5–7]. Successful therapy of NSCLC and other cancers with checkpoint blockade antibodies (targeting CTLA-4 and PD-1) provides direct evidence in humans that spontaneous T cell responses to human cancers can control those cancers [8]. There is growing interest in identifying patients whose tumors are infiltrated by CD8$^+$ T cells because those patients have improved prognosis and are more likely to respond to PD-1/PD-L1 blockade therapy.

CD8$^+$ tumor-infiltrating lymphocytes (TILs) can be identified in the TME by immunohistochemistry performed on paraffin-embedded tissue, and new systems for automated image analysis can enumerate CD8 antibody-stained cells with accuracy. However, the spatial distribution of TILs in the TME is heterogeneous [9]. Several classification schemes have been proposed to account for intratumoral heterogeneity of T cell infiltrates. The "immunoscore" used in colorectal cancer is based on the densities of CD8$^+$ and CD45RO$^+$ T cells in central tumor and in the invasive margin [10]. This measurement has been shown to have clinical relevance in colon cancer, and it has been proposed for inclusion in the colon cancer staging system as TNM-I (TNM-immune) [11–13]. On the other hand, we have defined "immunotypes" of metastatic melanoma based on the extent of immune cell invasion beyond the perivascular stroma and into tumor cell aggregates [2]. For NSCLC, an immunoscore concept has been proposed in which the density of CD8$^+$ T cells in the stroma is suggested as a promising prognostic marker [14, 15]. In a series of 797 patients, TILs in the stroma of the invasive margin were significantly associated with better patient outcomes [16]. However, a recent systematic review encompassing 8600 patients supports that stromal and intraepithelial CD8 counts are both associated with better overall survival [17]. Both of these studies acknowledged that the relative significance of CD8 T cell infiltration in the center of the tumor versus that of the tumor margin is not yet clear [14, 17]. In addition, associations between CD8 T cell infiltration and survival have varied across multiple studies of human NSCLC [3, 15, 18–24].

Importantly, comparing data among various studies is complicated by inconsistent strategies of sampling (stroma vs tumor nests, tissue microarray (TMA) vs submillimeter cores) and counting (small vs larger field views). Indeed, the accurate and consistent quantification of immune cells in the TME remains a challenge. In 2014, consensus guidelines attempted to standardize strategies of TIL measurement in breast cancer [25]. To date, no such guidelines exist for NSCLC. As the evidence for an association between CD8 TIL and NSCLC outcomes mounts and has greater potential for guiding patient treatment decisions, a set of clinically relevant sampling and counting guidelines should

be incorporated within the evolving NSCLC immunoscore. The purpose of this study was to assess a variety of sampling strategies to measure CD8$^+$ TIL density in NSCLC tumors. To the extent that CD8$^+$ TIL counts may be incorporated in clinical care pathways, it is important to consider whether sampling of representative areas of a whole tumor specimen may provide reliable counts in the setting of intratumoral heterogeneity. Also, because many cancers are initially evaluated with core biopsies or incisional biopsies, it is important to obtain some guidance about whether such biopsy samples can provide reliable sampling of the whole tumor. Thus, we have evaluated which of eight sampling strategies would have the closest correlation with CD8$^+$ TIL density measures of a whole tumor cross section, as well as with survival outcomes. We hypothesized that the validity of these sampling strategies would differ substantially, and that assessing CD8 counts by sampling the center of the tumor would have a better prognostic value than sampling the tumor margin.

Methods

Patients

Twenty-four non-small-cell lung tumors pathological stage I–III (stage I (13 %), II (8 %) and III (79 %)) were selected from our cancer registry from patients with ample follow-up. Surgical resections were performed from 2003 to 2012, and patients were followed through 8/2014 (mean follow-up 30 months). The study was approved by the University of Virginia institutional review board under protocol HSR # 18346. One tumor was excluded due to inadequate sampling.

Immunohistochemistry

Formalin-fixed and paraffin-embedded NSCLC samples were cut into 4 μm sections that were fixed on glass slides (Superfrost® Plus, Fisherbrand® Loughborough, England). Antigen retrieval was performed by immersing the slides in xylene for 10 min, then in decreasing concentrations of ethanol followed by PBS buffer and deionized water for 5 min each. Slides were then heated to 100 °C for 20 min in a pH 9 Tris-based solution. Mouse antihuman CD8 antibody (clone C8/144B, 1:200, DAKO, Carpinteria, CA, USA) was added for 1 h and then washed. Secondary antibody to mouse IgG was added for 30 min and washed before adding diaminobenzoic acid for 2 min to obtain ideal coloration of CD8-positive cells. Sections were counterstained with hematoxylin, dehydrated in increasing concentrations of ethanol, and immersed in xylene.

Table 1 Nine strategies used to assess CD8 counts

	Sampling strategy	N	Mean area (mm^2) ± 95 % CI	Discordant classification versus whole tumor		
				Above or below median (%)	Tertiles (%)	Quartiles (%)
A	Whole tumor	23	47.8 ± 18.9	0	0	0
B	Representative areas (6 mm^2)	23	4.7 ± 0.01	17	26	35
C	CTS	23	3.0 ± 0.2	17	13	26
D	DLA	20	0.4 ± 0.1	30	40	55
E	Invasive margin	22	2.5 ± 0.5	27	36	41
F	A random 1-mm-diameter area	23	0.8 ± 0.02	26	17	70
G	4 random 1-mm-diameter areas	23	3.0 ± 0.2	17	22	48
H	Random 20 % of the tumor	23	3.1 ± 0.4	9	0	26
I	Random 10 × 1 mm sample	23	8.4 ± 0.9	0	17	26

Summary of the areas assessed (mean ± 95 % CI) and the N available in each strategy. Also shown are the rates of discordant classification of CD8 densities obtained from the different sampling strategies in comparison with those obtained by assessing the whole tumor with respect to their median, tertile and quartile cutoffs. *CTS* central tumor sampling, *DLA* dense lymphoid aggregate
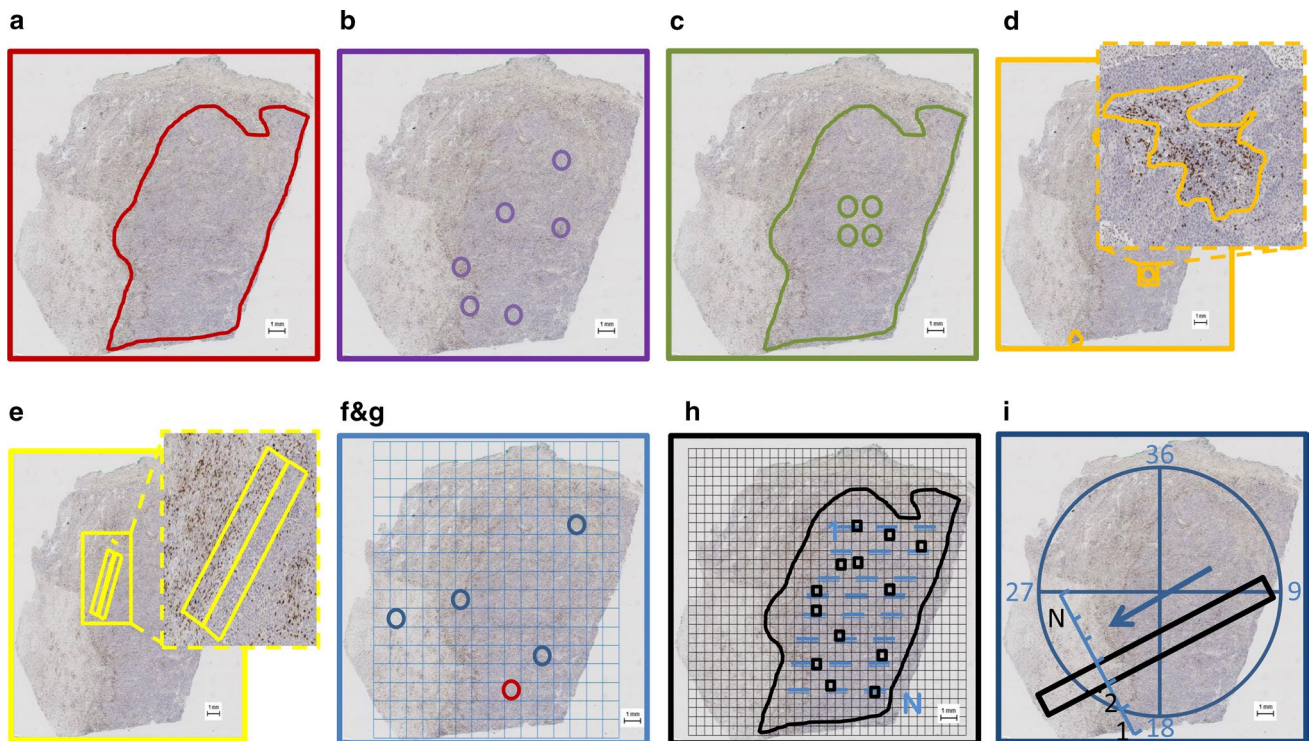
Fig. 1 **a–i** Illustrations of the sampling strategies (*scale bars*, in the lower right corner, are 1 mm in length.)

Cell counts

Slides were scanned using the Leica SCN400 ScanScope (Nussloch, Germany) and uploaded for automated counting using the Leica Digital Imaging Hub analysis software. Sampling areas were delineated on the scanned image by two investigators (Obeid and Hu). Automated cell counts are reported per millimeter square (mm^2).

Sampling strategies

Nine sampling strategies were used for each of the 23 evaluable whole tumor specimens (Table 1):

(a) Whole tumor: The tumor area was demarcated on a low magnification view of the specimen (Fig. 1a). The counts obtained by this strategy are inclusive of all the

heterogeneity within the tumor but exclude the peritumoral tissues beyond the invasive margin.

(b) Representative areas: six samples each 1 mm in diameter (equivalent to one $200 \times$ field of view) representing the three most densely infiltrated areas and the three least infiltrated areas within each tumor (Fig. 1b). The average $CD8^+$ T cell density across all six sampled areas was calculated to provide an estimate representative of the whole tumor. This sampling was based on similar strategies previously described [20, 22]. The primary advantage of this strategy is that the relatively low total sample area facilitates manual counting using an optical microscope.

(c) Central tumor sampling (CTS): Four 1 mm diameter samples of the center of the tumor were identified (Fig. 1c). In addition to its relevance in the "immunoscore," this sampling strategy also approximates the analysis of four TMA cores used in research studies [26–28]. Unless otherwise specified, TMA cores are commonly extracted from the center of a given tumor.

(d) Dense lymphoid aggregates (DLAs): Areas predominantly populated by dense lymphocytic infiltrates in the tumors were demarcated (Fig. 1d). These areas were not geometric but were drawn to encompass the irregular shapes of these aggregates.

(e) Invasive margin: 500 micron wide strips of the internal and external margin of the tumor were demarcated and counted (Fig. 1e). The average TIL density across the two samples was calculated to represent the TIL density of the tumor's invasive margin, a metric relevant to the NSCLC "immunoscore" [14, 16].

(f) Random1: After overlaying a grid of 1 mm squares on the whole tumor specimen, one square was randomly selected. A 1 mm diameter circular sample was demarcated within the random selection (Fig. 1f).

(g) Random4: Four random samples each 1 mm in diameter were collected using the method described above to simulate repeated small random biopsy samples (Fig. 1g).

(h) Random20 %: A grid of 500 μm^2 was overlaid on the tumor. Squares with the following features were excluded: Those within 500 μm of the tumor edge, squares containing false positive staining, and those containing necrotic tissue or DLAs were excluded. Squares with more than 50 % area populated by tumor cells were numbered from 1 to N. Among the numbered squares, 20 % were randomly selected for cell counting. Allowing for variations in total tumor size, a minimum of ten and a maximum of twenty squares were analyzed (with the exception of one tumor with four available squares, Fig. 1h). This strategy was designed to approximate recommendations by the 2014 Salgado et al. consensus guidelines for breast cancer [25].

(i) 10×1 mm core biopsy simulation: To simulate tissue samples from a clinical core biopsy 1 mm in diameter, circumferential axes of approach were created at 10-degree intervals crossing at the center of each tumor specimen. A directional axis was selected randomly, and a band of tissue 1 mm wide and 10 mm long is delineated along this axis (Fig. 1i). To ensure adequate sampling and to accurately represent diagnostically useful core biopsies, we excluded bands of tissue that encompassed only the peritumoral stroma or <3 mm of tumor tissue.

Statistical methods

CD8 cell counts were obtained from the Digital Image Hub software in cells/mm^2 for each of the different sampling strategies. Cell counts derived from the whole tumor specimen are the most representative of tumor heterogeneity; therefore, cell counts from each experimental sampling strategy were compared with those obtained in the whole tumor annotation. Paired t tests and Pearson's correlation coefficients were used to compare these sampling strategies against TIL densities derived from whole tumors. Overall survival (OS) was derived from the medical record and was calculated from the date of surgery until date of death, last follow-up, or last proof of life. Survival analyses were compared for patients whose tumors had high vs low CD8 density. To estimate an optimal high/low cutoff, we used a method developed by Contal and O'Quigley [29]. Kaplan–Meier graphs were generated, and the log-rank test was used to test associations with survival. Statistical analyses were performed using R (R Core Team, Vienna, Austria) and MedCalc (MedCalc Software Inc., Mariakerke, Belgium).

Results

On histologic and immunohistologic review of the NSCLC specimens, at least four regions within the TME were identified: stroma, epithelial tumor cell nests, invasive margin, and DLA as illustrated in Fig. 2. $CD8^+$ T cells were enumerated by the nine strategies defined above, some of which specifically evaluated selected regions within the TME, and some of which were random samples. Details of the CD8 counts by each sampling strategy are provided in Supplementary Table 1. Values were obtained for all 23 tumors for each sampling strategy, except that three tumors (13 %) did not have identifiable DLAs and one (4 %) did not have an evaluable tumor margin. The three tumors without DLAs were all in tumors with low CD8 counts by the total tumor area measure (two in the lowest quartile and one in the second quartile). The tumor without evaluable

Fig. 2 Heterogeneity of CD8 T cell infiltration in the tumor microenvironment. IHC diaminobenzoic acid stain of the CD8 marker (in *brown*), nuclei are visualized in *blue* (hematoxylin counterstain). Pictures from the intraepithelial compartment, dense lymphoid aggregates, tumor margin, and perivascular stroma of the same tumor sample at ×200 magnification, obtained with the Leica Digital Image Hub (*scale bars* in the lower right corner are 100 microns in length.)

Intra-epithelial

Dense Lymphoid Aggregate

Tumor Margin

Stroma: Perivascular
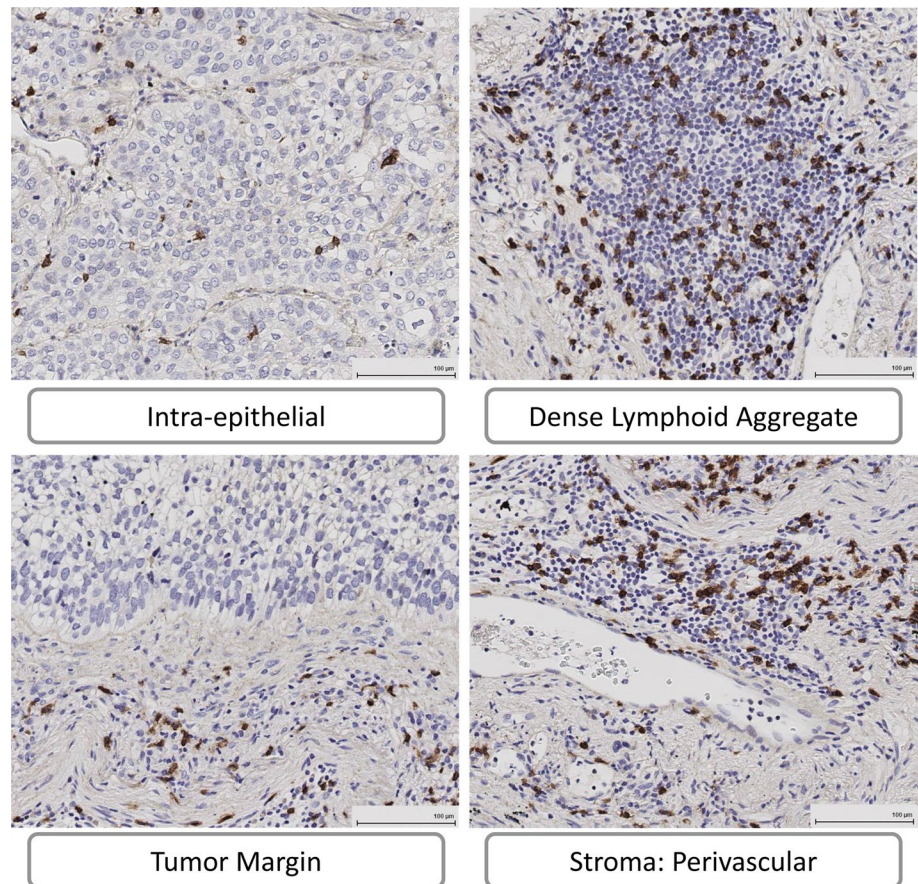
margin was in the third quartile of CD8 counts by whole tumor measure.

CD8 densities measured by each experimental sampling strategy were compared with the whole tumor CD8 density for the same tumor sections (Fig. 3). CD8 T cell counts obtained from all sampling strategies were significantly correlated with the whole tumor count ($p \leq 0.001$). Higher correlation with whole tumor counts ($r > 0.9$) was observed with the following sampling strategies: representative areas, CTS, random20 %, and the core biopsy simulation. Comparing CD8 counts between each of the 36 data pairs obtained from the different strategies demonstrates wide variability in correlation among the sampling strategies (Supplementary Figure. 1).

Prognostic biomarkers may be used to identify patients at high or low risk of cancer-associated death, based on classification of the values above and below a threshold cutoff. For a sampling strategy to be useful as a surrogate for counting the whole tumor, it should enable classification of the tumor samples into groups with high and low CD8 counts in a manner that is concordant with the same classification of the whole tumor samples. Thus, whole tumor samples were classified into high and low CD8 counts by dividing at the median CD8 density (≥ 204 vs

<204 CD8$^+$ cells/mm^2). The same process was repeated for each of the different sampling strategies. The proportions of tumors with discordant classification when compared to whole tumor are reported in Fig. 3. Highest discordancy rates were noted with the following sampling strategies: DLA (30 %), invasive margin (27 %), random1 (26 %), and random4 (17 %). Conversely, discordant classification was rare when sampling random20 % (9 %) and the axial sampling strategy replicating core biopsy (0 %). Similar calculations were also assessed across tertiles, where the lowest discordant classification rates were with random20 % (0 %), CTS (13 %), axial sampling/core biopsy simulation (17 %), and random1 (17 %). These data are included in Table 1. Thus, assessments of CD8 T cell infiltration will vary by the sampling strategy and also by the cutoffs used.

A challenge of comparing data from different studies is that different sampling strategies may result in differences of ranking and classification, as assessed above, but which may also result in different magnitudes of the cell counts. The latter would have relevance if a cutoff number of CD8$^+$ cells per mm^2 were to be employed in clinical practice. Thus, we compared the magnitudes of the CD8 counts obtained from each sampling strategy to those obtained from the whole tumor area (Fig. 4). The magnitudes were
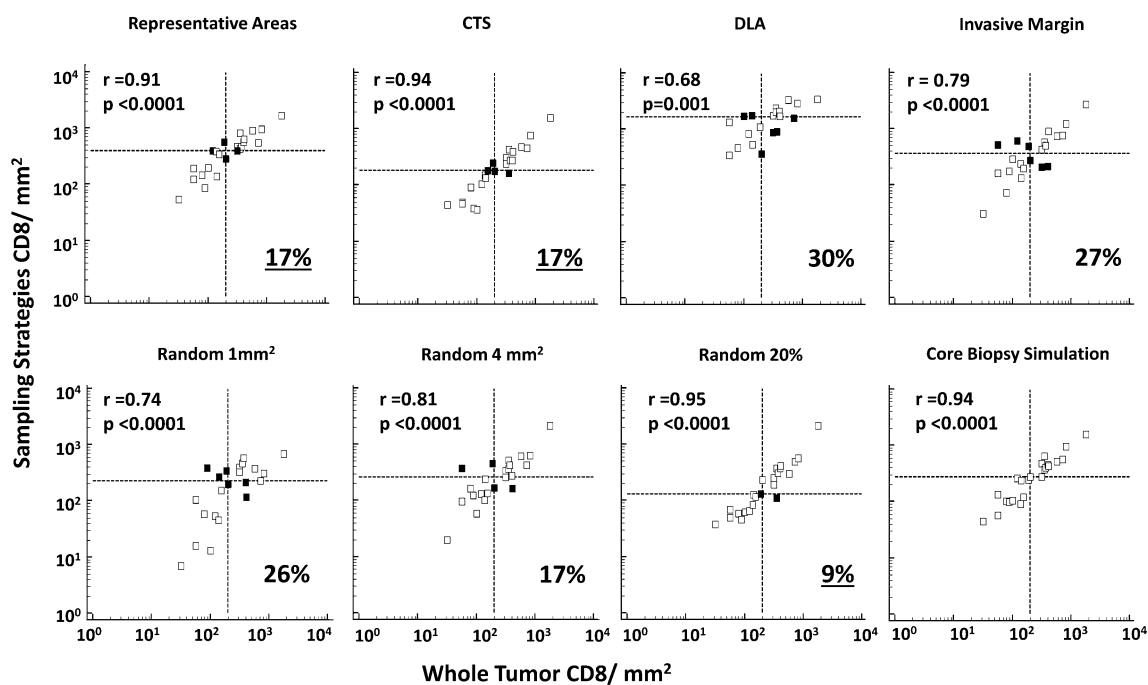
Fig. 3 Associations between CD8 densities obtained by the different sampling strategies, and whole tumor. Scatter plots represent the CD8 count from the different sampling strategies (*y*-axis) versus the whole tumor (*x*-axis) numbers; numbers are reported in log transform. Pearson correlation coefficients (*r*) and the corresponding *p* values are noted. *Dotted lines* represent the median CD8 count as obtained by each strategy. Misclassified tumors, tumors for which relationship to the median on experimental sample is different than the relationship to the median on the whole tumor, are represented in the top left and lower right quadrants of the *dotted lines* (*dark squares*); percent misclassification is noted in the right lower corner of each *graph*. *CTS* central tumor sampling, *DLA* dense lymphoid aggregate

similar to whole tumor measures for random1, random4, and simulated core biopsy sampling strategies ($p \geq 0.17$). However, there were significantly higher counts when sampling DLAs (350 % increase, $p < 0.0001$), margin areas (54 % increase, $p = 0.0025$), and representative areas (37 % increase, $p = 0.0004$). The representative areas included three regions considered to be representative of the highest CD8$^+$ T cell counts; not surprisingly, these provided mean values more than double those of the whole tumor values (Supplementary Fig. 2). CTS and random20 yielded counts that were 14 % ($p = 0.014$) and 15 % ($p = 0.046$) less than whole tumor counts, respectively (Fig. 4, Supplementary Table 1).

Overall survival was analyzed for CD8 densities obtained by each sampling strategy. The cutoff between high and low densities was determined for each sampling strategy according to the method of Contal and O'Quigley [29]. For eight of the strategies, high CD8 density was not significantly associated with patient survival ($p = 0.059$–0.399, Fig. 5). Interestingly, higher CD8$^+$ T cell infiltration in the center of the tumor (≥ 182 CD8/mm^2) was associated with better overall patient survival (median OS not reached for CD8^{high} vs 1.4 years for CD8^{low}, $p = 0.038$). No differences in CD8 densities were observed in tumors of different clinical characteristics (patient age, clinical stage, or tumor histology, Supplementary Table 2).

Discussion

Prior studies of human NSCLC have identified intratumoral heterogeneity in terms of proliferative activity [30], cellular histology [31], anaplastic lymphoma kinase rearrangements [32], and PD-L1 expression [9]. Thus, heterogeneity of CD8$^+$ T cell infiltrations is expected. Others have identified T cell infiltrations as prognostic for patients with NSCLC [19, 20, 22, 23, 33–35], and in most of these studies, the best predictor of survival was the density of the CD8$^+$ T cell subset [17, 24]. Thus, reliable quantitation of CD8$^+$ T cells in NSCLC may be valuable for assessing patient risk; furthermore, CD8$^+$ T cell infiltration has been associated with PD-L1 expression in other cancers [36] and may predict response to PD-1/PD-L1 blockade therapy. Thus, quantitation of CD8$^+$ T cells in NSCLC may have both prognostic and predictive relevance. However, we are not aware of studies that have directly compared different sampling strategies for their impact on the CD8$^+$ T cell counts. Thus, we have assessed the concordance among nine strategies of sampling tumors for CD8$^+$ T cell infiltrations in
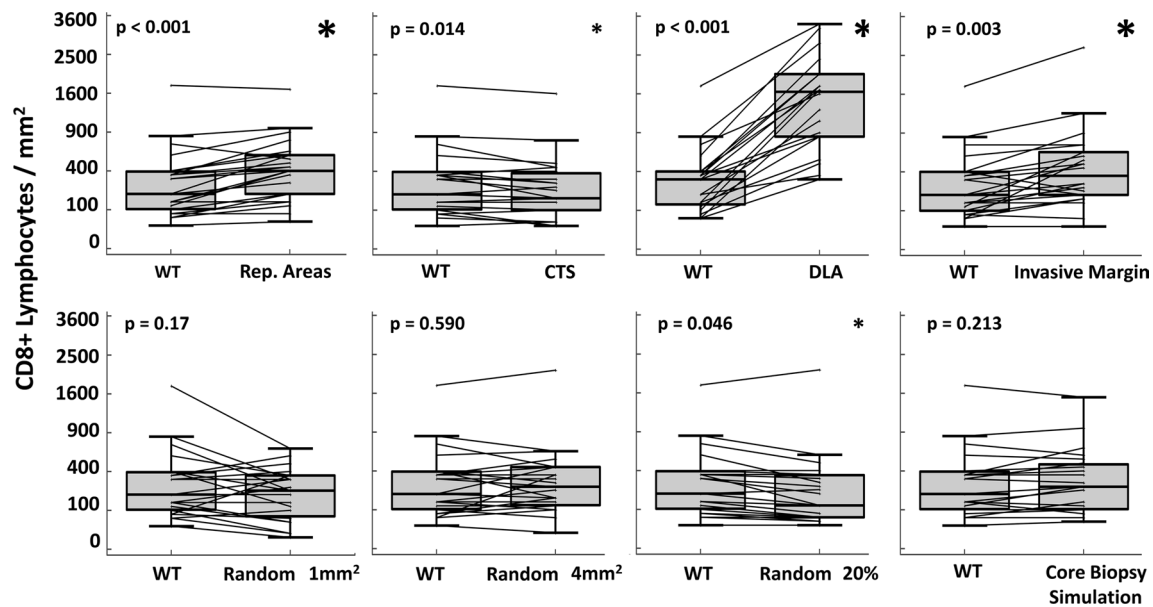
Fig. 4 Differences in magnitude of CD8 cell density by sampling strategy. The *y*-axis represents the CD8 T cell density. The *x*-axis of each *graph* is labeled with the whole tumor and another sampling strategy. Box plots show the distribution of CD8 densities from each strategy (median, quartiles). *Lines* connect the two CD8 densities relating to the same tumor. Paired t tests are used to assess the differ- ence, and *p* values are reported. A *small asterisk* denotes differences in the mean CD8 density that are smaller or equal to 15 %, whereas greater differences are noted by the large asterisk (Supplementary Table 1). *WT* whole tumor, *CTS* central tumor sampling, *DLA* dense lymphoid aggregate

multiple primary NSCLCs, in hopes of providing guidance in managing the heterogeneity integral to these cancers.

CD8$^+$ T cells may be enumerated in whole tumor cross sections, and automated imaging systems can be used to collect such data rapidly, but most hospitals do not currently have access to such systems; counting manu- ally requires sampling a subset of the tumor tissue. Also, treatment decisions may be based on core biopsies of the tumors, but it is not clear how well cell counts in the core biopsy specimens represent the whole tumors. Finally, when tumor biopsies are obtained sequentially during clini- cal trials, they often are core biopsies and sometimes are very small samples. Thus, there is a need to understand how well the CD8 counts in a whole tumor may be esti- mated by small biopsies, core biopsies, random sampling, or sampling of investigator-identified representative areas of a tumor specimen. The findings from the present study provide some information to address these questions.

This study shows significant correlations between CD8 densities obtained by all eight sampling strategies and those from the whole tumor (Fig. 3), a finding that sug- gests that the ranking of tumors should be consistent between each sampling strategy and the whole tumor. Fur- ther, it would suggest that the use of any of these sampling strategies should enable dichotomization of the population into high and low CD8 densities with results comparable to dichotomization using the whole tumor. However, we observed considerable discordance of classification into high versus low CD8 densities. The discordance increased when grouping into tertiles or quartiles (Table 1). This may explain discrepancies in the prognostic significance of CD8 counts reported from different studies.

There was greater concordance with whole tumor for multiple random samples of 20 % of the tumor or a ran- dom core biopsy measuring 10 × 1 mm than with sampling of 1–4 random areas of 1 mm diameter. The mean area counted with the 20 % random samples (3.1 mm^2 <10 % of the mean whole tumor area due to the exclusions men- tioned in the strategies section and the maximum of 20 squares analyzed) was similar to that of the four random 1-mm-diameter areas (3.0 mm^2); however, the rate of dis- cordant classification was lower for the random20 % strat- egy (Table 1). A critical difference in those two approaches is that the random20 % strategy used 10–20 smaller sam- ples (500 micron); so, it decreased the chance of skewing the total score based on one very low or very high sample. Another approach that provided high concordance was to sample a large 10 × 1 mm portion of the tumor, which mimics a clinical core biopsy. This was the largest area sampled, suggesting that greater concordance with the whole tumor may be enabled by simply increasing the size of the area sampled.

It is also informative to compare evaluation of the invasive margin versus the central tumor. The mean area
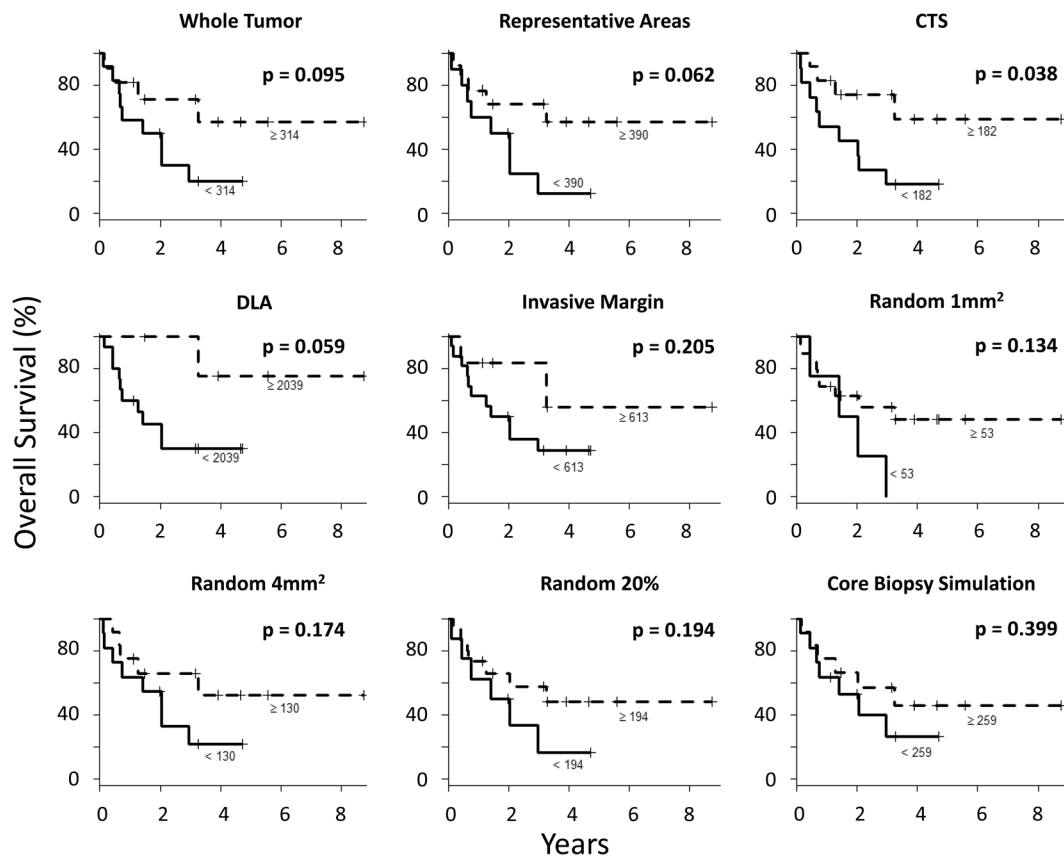
Fig. 5 Association of CD8 densities with outcomes. Kaplan–Meier curves showing the overall survival of patients whose tumors have high versus low CD8 density as obtained by each of the different sampling strategies. Cutoff values for each strategy defined using the Contal and O'Quigley method [29] are noted. Log-rank tests are used to assess the significance of the difference between the OS of the high CD8 group and that of the low CD8 group. The distribu-tion of patients into groups of high and low CD8 tumor infiltration according to each sampling strategy is as follows: 12 and 11 for the whole tumor densities, 10 and 13 representative areas, 11 and 12 central tumor sampling (CTS), 15 and 5 for dense lymphoid aggregates (DLAs), 16 and 6 invasive margin, 4 and 19 random1, 11 and 12 for random4, 8 and 15 random20 %, and 11 and 12 for core biopsy simulation

sampled was similar (2.5 vs 3.0 mm^2, respectively), but the rate of discordant classification was consistently lower with CTS than with invasive margin sampling (17 vs 27 % for 2 groups, 13 vs 36 % for tertiles, and 26 vs 41 % for quartiles, Table 1). Similarly, the correlation coefficient was more favorable for CTS counts ($r = 0.94$) than for the invasive margin ($r = 0.79$, Fig. 3). It is notable also that concordance was better for the random20 % than for the representative areas, despite the smaller area sampled for the former. This suggests that random sampling may enable more accurate counts than similar or larger areas that are selected by the investigator. Overall, these findings support assessing CD8 counts by random sampling of ten or more small areas encompassing 3 mm^2 or more. They also support use of core biopsy samples that are at least 10×1 mm in size.

The magnitude of absolute difference in TIL density with different sampling strategies suggests that inconsist-encies in tumor sampling can diminish the generalizability of CD8 count cutoffs when different sampling strategies are used (Fig. 4). Contributors to higher counts include the DLAs and cells at the invasive margin. The DLAs likely include ectopic lymphoid structures or tertiary lym-phoid structures (that include the simultaneous presence of T cells, B cells, and dendritic cells), whose presence may predict better patient survival in lung and other can-cers [22, 37–39]. These lymphoid structures are associated with T cell infiltration and have been implicated mecha-nistically in supporting T cell infiltration into cancers [40]. Such aggregates are excluded in breast cancer by consen-sus [41]; it may be reasonable to consider a similar prac-tice for NSCLC. On the other hand, the apparent functional relevance of the tertiary lymphoid structures may warrant future studies to assess prognostic or predictive value in enumerating tertiary lymphoid structures specifically or in quantifying them as a proportion of the total tumor cross-sectional area.

Sampling strategies that are concordant with the whole tumor measure may or may not have the greatest prognostic

significance. In our dataset of 23 patients, those with higher CD8 counts in the center of the tumor experienced longer overall survival than those with low CD8 counts. CD8 densities obtained by all other sampling strategies were not significantly associated with survival. This study included an intensive analysis of different sampling strategies but was limited to 23 patients; however, in this population, CD8 densities and their distribution into high and low categories in the whole tumors and their centers were not affected by tumor histology, patient age, or stage of disease (Supplementary Table 2). The fact that the central tumor CD8 counts were significantly associated with survival suggests a very strong association, which is supported by other studies [24]. Other sampling strategies may also have prognostic value in larger datasets, which were beyond the scope of the present study.

Limitations of our study include the sample size and the lack of assessment of CD8 counts as predictive of the response to checkpoint blockade therapy. This study also did not specifically address differences in TIL localization in stromal versus intraepithelial compartments. A previous study including 178 NSCLC patients found that higher CD8 infiltration in the stroma of the invasive margin was associated with survival, while CD8 density in the stroma in the center of the tumors was not [16]. However, a meta-analysis representing 8600 NSCLC patients found that overall survival was associated with increased CD8$^+$ T cells in either the stroma or the tumor nests [17]. Regardless, separating the stromal and epithelial compartments is technically challenging and susceptible to investigator error, as illustrated in Supplementary Figure. 3, and may be unreliable even to assess with automated analysis [14]. Instead, we have focused on approaches that may be more amenable to routine clinical application. Another limitation of the present study is that we only evaluated CD8$^+$ T cells. A recent report suggested the value of CD45RO$^+$ cells in addition to the NSCLC immunoscore [42], and that marker has been reported to be useful for the colon cancer immunoscore [12]. We did not study CD45RO$^+$ lymphocytic infiltration within this experiment. Thus far, CD8 infiltration has been the most promising candidate in NSCLC due to its reproducibility and prognostic value across a number of studies [14, 16, 18, 43, 44]. Other immune cells, such as dendritic cells and macrophages, have important roles in NSCLC tumor progression [45]; studying their heterogeneity in the NSCLC TME will be valuable.

This study underlines the heterogeneity in the distribution of CD8$^+$ T cells in the TME of human NSCLCs. We propose that enumerating CD8$^+$ T cells in small tumor samples or small numbers of tumor samples may not reliably represent the counts throughout the whole tumor. On the other hand, sampling 10–20 small areas randomly, sampling the tumor center, or taking large core biopsies

(10 × 1 mm) may best represent the whole tumor. Small (e.g., 1 mm^2) samples, or samples selected in a non-random manner, may be particularly misleading for a large subset of patients. Since the CD8 counts from the tumor center had prognostic significance and since other studies support the prognostic significance of CD8 counts at the tumor center [17], this approach may be a good choice for future analyses. Furthermore, this sampling strategy is also practical and easily applicable to both manual and automated cell enumeration due to a limited surface area of evaluation. Among patients that only have core biopsy specimens available, TIL density evaluated through these samples is likely also valid. However, additional studies that directly measure the prognostic or predictive value of CD8 counts in core biopsy samples may be needed before relying on them to guide clinical decisions.

Compliance with ethical standards

References

1. Fridman WH, Pages F, Sautes-Fridman C, Galon J (2012) The immune contexture in human tumours: impact on clinical outcome. Nat Rev Cancer 12(4):298–306. doi:10.1038/nrc3245

2. Erdag G, Schaefer JT, Smolkin ME, Deacon DH, Shea SM, Dengel LT, Patterson JW, Slingluff CL Jr (2012) Immunotype and immunohistologic characteristics of tumor-infiltrating immune cells are associated with clinical outcome in metastatic melanoma. Cancer Res 72(5):1070–1080. doi:10.1158/0008-5472.can-11-3218

3. Liu H, Zhang T, Ye J, Li H, Huang J, Li X, Wu B, Huang X, Hou J (2012) Tumor-infiltrating lymphocytes predict response to chemotherapy in patients with advance non-small cell lung cancer. Cancer Immunol Immunother 61(10):1849–1856. doi:10.1007/s00262-012-1231-7

4. Tanis E, Julie C, Emile JF, Mauer M, Nordlinger B, Aust D, Roth A, Lutz MP, Gruenberger T, Wrba F, Sorbye H, Bechstein W, Schlag P, Fisseler A, Ruers T (2015) Prognostic impact of immune response in resectable colorectal liver metastases treated by surgery alone or surgery with perioperative FOLFOX in the randomised EORTC study 40983. Eur J Cancer 51(17):2708–2717. doi:10.1016/j.ejca.2015.08.014

5. Liu S, Lachapelle J, Leung S, Gao D, Foulkes WD, Nielsen TO (2012) CD8+ lymphocyte infiltration is an independent

favorable prognostic indicator in basal-like breast cancer. Breast Cancer Res 14(2):R48. doi:10.1186/bcr3148

6. Domagala-Kulawik J (2015) The role of the immune system in non-small cell lung carcinoma and potential for therapeutic intervention. Transl Lung Cancer Res 4(2):177–190. doi:10.3978/j.issn.2218-6751.2015.01.11

7. Motz GT, Coukos G (2013) Deciphering and reversing tumor immune suppression. Immunity 39(1):61–73. doi:10.1016/j.immuni.2013.07.005

8. Callahan MK, Postow MA, Wolchok JD (2014) CTLA-4 and PD-1 pathway blockade: combinations in the clinic. Front Oncol 4:385. doi:10.3389/fonc.2014.00385

9. McLaughlin J, Han G, Schalper KA, Carvajal-Hausdorf D, Pelekanou V, Rehman J, Velcheti V, Herbst R, LoRusso P, Rimm DL (2016) Quantitative assessment of the heterogeneity of PD-L1 expression in non-small-cell lung cancer. JAMA Oncol 2(1):46–54. doi:10.1001/jamaoncol.2015.3638

10. Galon J, Pages F, Marincola FM, Angell HK, Thurin M, Lugli A, Zlobec I, Berger A, Bifulco C, Botti G, Tatangelo F, Britten CM, Kreiter S, Chouchane L, Delrio P, Arndt H, Asslaber M, Maio M, Masucci GV, Mihm M, Vidal-Vanaclocha F, Allison JP, Gnjatic S, Hakansson L, Huber C, Singh-Jasuja H, Ottensmeier C, Zwierzina H, Laghi L, Grizzi F, Ohashi PS, Shaw PA, Clarke BA, Wouters BG, Kawakami Y, Hazama S, Okuno K, Wang E, O'Donnell-Tormey J, Lagorce C, Pawelec G, Nishimura MI, Hawkins R, Lapointe R, Lundqvist A, Khleif SN, Ogino S, Gibbs P, Waring P, Sato N, Torigoe T, Itoh K, Patel PS, Shukla SN, Palmqvist R, Nagtegaal ID, Wang Y, D'Arrigo C, Kopetz S, Sinicrope FA, Trinchieri G, Gajewski TF, Ascierto PA, Fox BA (2012) Cancer classification using the immunoscore: a worldwide task force. J Transl Med 10:205. doi:10.1186/1479-5876-10-205

11. Galon J, Mlecnik B, Bindea G, Angell HK, Berger A, Lagorce C, Lugli A, Zlobec I, Hartmann A, Bifulco C, Nagtegaal ID, Palmqvist R, Masucci GV, Botti G, Tatangelo F, Delrio P, Maio M, Laghi L, Grizzi F, Asslaber M, D'Arrigo C, Vidal-Vanaclocha F, Zavadova E, Chouchane L, Ohashi PS, Hafezi-Bakhtiari S, Wouters BG, Roehrl M, Nguyen L, Kawakami Y, Hazama S, Okuno K, Ogino S, Gibbs P, Waring P, Sato N, Torigoe T, Itoh K, Patel PS, Shukla SN, Wang Y, Kopetz S, Sinicrope FA, Scripcariu V, Ascierto PA, Marincola FM, Fox BA, Pages F (2014) Towards the introduction of the 'Immunoscore' in the classification of malignant tumours. J Pathol 232(2):199–209. doi:10.1002/path.4287

12. Pages F, Kirilovsky A, Mlecnik B, Asslaber M, Tosolini M, Bindea G, Lagorce C, Wind P, Marliot F, Bruneval P, Zatloukal K, Trajanoski Z, Berger A, Fridman WH, Galon J (2009) In situ cytotoxic and memory T cells predict outcome in patients with early-stage colorectal cancer. J Clin Oncol 27(35):5944–5951. doi:10.1200/JCO.2008.19.6147

13. Mlecnik B, Tosolini M, Kirilovsky A, Berger A, Bindea G, Meatchi T, Bruneval P, Trajanoski Z, Fridman WH, Pages F, Galon J (2011) Histopathologic-based prognostic factors of colorectal cancers are associated with the state of the local immune reaction. J Clin Oncol 29(6):610–618. doi:10.1200/JCO.2010.30.5425

14. Donnem T, Kilvaer TK, Andersen S, Richardsen E, Paulsen E, Hald SM, Al-Saad S, Brustugun OT, Helland A, Lund-Iversen M, Solberg S, Gronberg BH, Wahl SG, Helgeland L, Flotten O, Pohl M, Al-Shibli K, Sandanger TM, Pezzella F, Busund L, Bremnes RM (2015) Strategies for clinical implementation of TNM-immunoscore in resected non-small cell lung cancer. Ann Oncol 27(2):225–232. doi:10.1093/annonc/mdv560

15. Tian C, Lu S, Fan Q, Zhang W, Jiao S, Zhao X, Wu Z, Sun L, Wang L (2015) Prognostic significance of tumor-infiltrating CD8(+) or CD3(+) T lymphocytes and interleukin-2 expression in radically resected non-small cell lung cancer. Chin Med J 128(1):105–110. doi:10.4103/0366-6999.147828

16. Donnem T, Hald SM, Paulsen EE, Richardsen E, Al-Saad S, Kilvaer TK, Brustugun OT, Helland A, Lund-Iversen M, Poehl M, Olsen KE, Ditzel HJ, Hansen O, Al-Shibli K, Kiselev Y, Sandanger TM, Andersen S, Pezzella F, Bremnes RM, Busund LT (2015) Stromal CD8+ T-cell density—a promising supplement to TNM staging in non-small cell lung cancer. Clin Cancer Res 21(11):2635–2643. doi:10.1158/1078-0432.CCR-14-1905

17. Geng Y, Shao Y, He W, Hu W, Xu Y, Chen J, Wu C, Jiang J (2015) Prognostic role of tumor-infiltrating lymphocytes in lung cancer: a meta-analysis. Cell Physiol Biochem 37(4):1560–1571. doi:10.1159/000438523

18. Al-Shibli K, Al-Saad S, Andersen S, Donnem T, Bremnes RM, Busund LT (2010) The prognostic value of intraepithelial and stromal CD3-, CD117- and CD138-positive cells in non-small cell lung carcinoma. APMIS 118(5):371–382. doi:10.1111/j.1600-0463.2010.02609

19. Dai F, Liu L, Che G, Yu N, Pu Q, Zhang S, Ma J, Ma L, You Z (2010) The number and microlocalization of tumor-associated immune cells are associated with patient's survival time in non-small cell lung cancer. BMC Cancer 10:220. doi:10.1186/1471-2407-10-220

20. Djenidi F, Adam J, Goubar A, Durgeau A, Meurice G, de Montpreville V, Validire P, Besse B, Mami-Chouaib F (2015) CD8+ CD103+ tumor-infiltrating lymphocytes are tumor-specific tissue-resident memory T cells and a prognostic factor for survival in lung cancer patients. J Immunol 194(7):3475–3486. doi:10.4049/jimmunol.1402711

21. Mori M, Ohtani H, Naito Y, Sagawa M, Sato M, Fujimura S, Nagura H (2000) Infiltration of CD8+ T cells in non-small cell lung cancer is associated with dedifferentiation of cancer cells, but not with prognosis. Tohoku J Exp Med 191(2):113–118

22. Schalper KA, Brown J, Carvajal-Hausdorf D, McLaughlin J, Velcheti V, Syrigos KN, Herbst RS, Rimm DL (2015) Objective measurement and clinical significance of TILs in non-small cell lung cancer. J Natl Cancer Inst. doi:10.1093/jnci/dju435

23. Wakabayashi O, Yamazaki K, Oizumi S, Hommura F, Kinoshita I, Ogura S, Dosaka-Akita H, Nishimura M (2003) CD4+ T cells in cancer stroma, not CD8+ T cells in cancer cell nests, are associated with favorable prognosis in human non-small cell lung cancers. Cancer Sci 94(11):1003–1009

24. Zeng DQ, Yu YF, Ou QY, Li XY, Zhong RZ, Xie CM, Hu QG (2016) Prognostic and predictive value of tumor-infiltrating lymphocytes for clinical therapeutic research in patients with non-small cell lung cancer. Oncotarget 7(12):13765–13781. doi:10.18632/oncotarget.7282

25. Salgado R, Denkert C, Demaria S, Sirtaine N, Klauschen F, Pruneri G, Wienert S, Van den Eynden G, Baehner FL, Penault-Llorca F, Perez EA, Thompson EA, Symmans WF, Richardson AL, Brock J, Criscitiello C, Bailey H, Ignatiadis M, Floris G, Sparano J, Kos Z, Nielsen T, Rimm DL, Allison KH, Reis-Filho JS, Loibl S, Sotiriou C, Viale G, Badve S, Adams S, Willard-Gallo K, Loi S (2015) The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs working group 2014. Ann Oncol 26(2):259–271. doi:10.1093/annonc/mdu450

26. Al-Shibli KI, Donnem T, Al-Saad S, Persson M, Bremnes RM, Busund LT (2008) Prognostic effect of epithelial and stromal lymphocyte infiltration in non-small cell lung cancer. Clin Cancer Res 14(16):5220–5227. doi:10.1158/1078-0432.CCR-08-0133

27. Ilie M, Hofman V, Ortholan C, Bonnetaud C, Coelle C, Mouroux J, Hofman P (2012) Predictive clinical outcome of the intratumoral CD66b-positive neutrophil-to-CD8-positive T-cell ratio in patients with resectable nonsmall cell lung cancer. Cancer 118(6):1726–1737. doi:10.1002/cncr.26456

28. Kayser G, Schulte-Uentrop L, Sienel W, Werner M, Fisch P, Passlick B, Zur Hausen A, Stremmel C (2012) Stromal CD4/CD25 positive T-cells are a strong and independent prognostic factor in non-small cell lung cancer patients, especially with adenocarcinomas. Lung Cancer 76(3):445–451. doi:10.1016/j.lungcan.2012.01.004

29. Contal C, O'Quigley J (1999) An application of change point methods in studying the effect of age on survival in breast cancer. Comput Stat Data Anal 30(3):253–270. doi:10.1016/S0167-9473(98)00096-6

30. Del Gobbo A, Pellegrinelli A, Gaudioso G, Castellani M, Zito Marino F, Franco R, Palleschi A, Nosotti M, Bosari S, Vaira V, Ferrero S (2016) Analysis of NSCLC tumour heterogeneity, proliferative and 18F-FDG PET indices reveals Ki67 prognostic role in adenocarcinomas. Histopathology 68(5):746–751. doi:10.1111/his.12808

31. Cadioli A, Rossi G, Costantini M, Cavazza A, Migaldi M, Colby TV (2014) Lung cancer histologic and immunohistochemical heterogeneity in the era of molecular therapies: analysis of 172 consecutive surgically resected, entirely sampled pulmonary carcinomas. Am J Surg Pathol 38(4):502–509. doi:10.1097/pas.0000000000000154

32. Abe H, Kawahara A, Azuma K, Taira T, Takase Y, Fukumitsu C, Murata K, Yamaguchi T, Akiba J, Ishii H, Okamoto I, Hoshino T, Takamori S, Kage M (2015) Heterogeneity of anaplastic lymphoma kinase gene rearrangement in non-small-cell lung carcinomas: a comparative study between small biopsy and excision samples. J Thorac Oncol 10(5):800–805. doi:10.1097/jto.0000000000000507

33. Zhuang X, Xia X, Wang C, Gao F, Shan N, Zhang L, Zhang L (2010) A high number of CD8 + T cells infiltrated in NSCLC tissues is associated with a favorable prognosis. Appl Immunohistochem Mol Morphol 18(1):24–28. doi:10.1097/PAI.0b013e3181b6a741

34. Hiraoka K, Miyamoto M, Cho Y, Suzuoki M, Oshikiri T, Nakakubo Y, Itoh T, Ohbuchi T, Kondo S, Katoh H (2006) Concurrent infiltration by CD8+ T cells and CD4+ T cells is a favourable prognostic factor in non-small-cell lung carcinoma. Br J Cancer 94(2):275–280. doi:10.1038/sj.bjc.6602934

35. Eerola AK, Soini Y, Paakko P (2000) A high number of tumor-infiltrating lymphocytes are associated with a small tumor size, low tumor stage, and a favorable prognosis in operated small cell lung carcinoma. Clin Cancer Res 6(5):1875–1881

36. Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJ, Robert L, Chmielowski B, Spasic M, Henry G, Ciobanu V, West AN, Carmona M, Kivork C, Seja E, Cherry G, Gutierrez AJ, Grogan TR, Mateus C, Tomasic G, Glaspy JA, Emerson RO, Robins H, Pierce RH, Elashoff DA, Robert C, Ribas A (2014) PD-1 blockade induces responses by inhibiting adaptive immune resistance. Nature 515(7528):568–571. doi:10.1038/nature13954

37. Dieu-Nosjean MC, Antoine M, Danel C, Heudes D, Wislez M, Poulot V, Rabbe N, Laurans L, Tartour E, De Chaisemartin L, Lebecque S, Fridman WH, Cadranel J (2008) Long-term survival for patients with non-small-cell lung cancer with intratumoral lymphoid structures. J Clin Oncol 26(27):4410–4417

38. Messina JL, Fenstermacher DA, Eschrich S, Qu X, Berglund AE, Lloyd MC, Schell MJ, Sondak VK, Weber JS, Mule JJ (2012) 12-Chemokine gene signature identifies lymph node-like structures in melanoma: potential for patient selection for immunotherapy? Sci Rep 2:765. doi:10.1038/srep00765

39. Nawaz S, Heindl A, Koelble K, Yuan Y (2015) Beyond immune density: critical role of spatial heterogeneity in estrogen receptor-negative breast cancer. Mod Pathol 28(6):766–777. doi:10.1038/modpathol.2015.37

40. Peske JD, Thompson ED, Gemta L, Baylis RA, Fu YX, Engelhard VH (2015) Effector lymphocyte-induced lymph node-like vasculature enables naive T-cell entry into tumours and enhanced anti-tumour immunity. Nat Commun 6:7114. doi:10.1038/ncomms8114

41. Salgado R, Denkert C, Campbell C, Savas P, Nucifero P, Aura C, de Azambuja E, Eidtmann H, Ellis CE, Baselga J, Piccart-Gebhart MJ, Michiels S, Bradbury I, Sotiriou C, Loi S (2015) Tumor-infiltrating lymphocytes and associations with pathological complete response and event-free survival in HER2-positive early-stage breast cancer treated with lapatinib and trastuzumab: a secondary analysis of the NeoALTTO trial. JAMA Oncol 1(4):448–454. doi:10.1001/jamaoncol.2015.0830

42. Paulsen EE, Kilvaer T, Khanehkenari MR, Maurseth RJ, Al-Saad S, Hald SM, Al-Shibli K, Andersen S, Richardsen E, Busund LT, Bremnes R, Donnem T (2015) CD45RO(+) memory T lymphocytes—a candidate marker for TNM-immunoscore in squamous non-small cell lung cancer. Neoplasia 17(11):839–848. doi:10.1016/j.neo.2015.11.004

43. Horne ZD, Jack R, Gray ZT, Siegfried JM, Wilson DO, Yousem SA, Nason KS, Landreneau RJ, Luketich JD, Schuchert MJ (2011) Increased levels of tumor-infiltrating lymphocytes are associated with improved recurrence-free survival in stage 1A non-small-cell lung cancer. J Surg Res 171(1):1–5. doi:10.1016/j.jss.2011.03.068

44. Kilic A, Landreneau RJ, Luketich JD, Pennathur A, Schuchert MJ (2011) Density of tumor-infiltrating lymphocytes correlates with disease recurrence and survival in patients with large non-small-cell lung cancer tumors. J Surg Res 167(2):207–210. doi:10.1016/j.jss.2009.08.029

45. Inoshima N, Nakanishi Y, Minami T, Izumi M, Takayama K, Yoshino I, Hara N (2002) The influence of dendritic cell infiltration and vascular endothelial growth factor expression on the prognosis of non-small cell lung cancer. Clin Cancer Res 8(11):3480–3486