CT texture analysis as predictive factor in metastatic lung adenocarcinoma treated with tyrosine kinase inhibitors (TKIs)

Abstract
Purpose: To assess the predictive and prognostic value of pre-treatment CT texture features in lung adenocarcinoma treated with tyrosine kinase inhibitors (TKI).

Materials and Methods: Texture analysis was performed using commercially available software (TexRAD Ltd, Cambridge, UK) on pre-treatment contrast-enhanced CT studies from 50 patients with metastatic lung adenocarcinoma treated by TKI. Texture features were quantified on a 5-mm-thick central slice of the primary tumor and were correlated with progression-free and overall survival (PFS and OS) using an internally cross-validated machine learning approach then validated on a bootstrapped sample.

Results: Median PFS and OS were 10.5 and 20.7 months, respectively. A noninvasive signature based on five texture parameters predicted 6-month progression with Area Under the Curve (AUC) of 0.8 (95%CI) and 1-year progression with AUC of 0.76. A high-risk group had hazard ratios for progression of 4.63 and 5.78 when divided by median and best cut-off points, respectively. Texture signature did not correlate with OS. Available clinical variables did not correlate with PFS or with OS.

Conclusion: Texture features seem to be associated with PFS in lung adenocarcinoma treated with TKI.

Keywords: Protein Kinase Inhibitors; Receptor, Epidermal Growth Factor; Carcinoma, Non Small Cell Lung; Adenocarcinoma; Texture Analysis; Tomography, X-Ray computed; Progression-Free Survival; Prognosis.

Abbreviations
AUC, Area Under the Curve
CECT, contrast-enhanced computed tomography
CTTA, CT texture analysis
EGFR, epidermal growth factor receptor
LASSO, Least Absolute Shrinkage and Selection Operator
OS, overall survival
PFS, progression-free survival
TA, texture analysis
TKI, tyrosine kinase inhibitors

**Funding:**
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
Introduction

Lung cancer is responsible for more than 150,000 deaths every year in the United States (1). Only 15% of the patients are diagnosed with localized disease, and stage IV lung cancer has a 5-year survival rate of only 4% (2). About 20% of non-small cell lung cancer (NSCLC) presents in-frame deletions of exon 19 and 21 in the tyrosine kinase domain of the epidermal growth factor receptor (EGFR) gene (3,4). Tyrosine kinase inhibitors (TKIs) targeting the EGFR pathway, such as erlotinib or gefitinib, are the current gold standard in metastatic lung adenocarcinoma harboring EGFR mutations (5) both as first-line treatment or after failure of standard first-line chemotherapy. In these settings, TKIs can improve overall survival (OS) and have lower toxicity compared to cytotoxic drugs. However, treatment response is very heterogeneous and de novo resistance to EGFR-TKIs has been described in 20–30% of patients (6–9). This has both a clinical and an economic impact given the high cost of these drugs (10). For this reason, it would be paramount to identify predictive factors able to stratify the risk of resistance to treatment. Even if the mechanisms of de novo resistance are still not understood, some gene mutations associated with EGFR-TKI resistance have been identified (11); however, detection of these mutations in tumor samples could be biased by tumor heterogeneity (12), while liquid biopsy on circulating tumor DNA is promising but still affected by a lack of sensitivity (13).

A different (alternative or complementary) approach is provided by texture analysis (TA) of computed tomography (CT) images, which quantifies the heterogeneity of tumor density and contrast-enhancement and features/objects of different sizes and numbers. This technique is non-invasive and can be performed on routinely acquired imaging studies, such as contrast-enhanced computed tomography (CECT), which is routinely used in staging and treatment response monitoring of NSCLC. Texture features have been found to be associated with relevant phenotypic (14) and genotypic (15,16) tumor characteristics in NSCLC and are predictive of treatment response and prognosis in many oncologic settings (17–20).

The objective of this study is to assess if tumor heterogeneity as assessed by pre-treatment CECT-TA can be a predictive and/or prognostic factor in patients with advanced lung adenocarcinoma treated with EGFR-TKIs.

Materials and methods
Patient selection

This retrospective study was approved by the institutional review board and was conducted in accordance with the ethical standards of the Declaration of Helsinki. Patients were selected consecutively between October 2008 and November 2016. Inclusion criteria were: ascertained histological diagnosis of stage IV lung adenocarcinoma according to the 8th AJCC edition (21), available pre-treatment CECT study, first-line treatment with TKIs targeting the EGFR pathway (i.e. erlotinib, gefitinib). Exclusion criteria were: lesion size < 20 mm, non-solid tumor. The following clinical variables were registered: age, sex, BMI and performance status (expressed as ECOG score (22)).

Chemotherapy

Erlotinib was administered at a daily dosage of 150 mg, and gefitinib was administered at a daily dosage of 250 mg. Therapy was continued until documented disease progression or unacceptable toxicity (Grade III/IV according to Common Terminology Criteria for Adverse Events v 4.0).

CT acquisition

The CT protocol included a thoracic volumetric acquisition 80 s after intravenous administration of 120 mL of iodinated contrast agent at 2 mL/s (Iomeron 350, Bracco, Milan, Italy). CT acquisition was performed on two multidetector scanners (from 2008 to 2011 Siemens Somatom Sensation 16, Forcheim, Germany; from 2011 Siemens Somatom Definition Flash, Forcheim, Germany) with similar parameters: acquisition FOV 500 mm, reconstruction FOV tailored to patient size, matrix 512×512, pixel size ranging from 0.57 to 1.02 mm, tube voltage 120 kV, reference current 150 mAs modulated with CareDose 4D. Images were reconstructed using both filtered back projection (FBP) and iterative algorithms when available; image analysis was performed on FBP images because they were the only available for all the studies.

Follow-up and survival assessment

Clinical evaluation was performed monthly. Follow-up CT was performed every 3 months or in case of clinical deterioration assessed by two experienced oncologists. Progression was defined according to RECIST 1.1 by radiologists experienced in lung cancer staging and follow-up. Progression-free survival (PFS) and OS were calculated from the date of therapy commencement until progression or death from any cause and were used as outcome indicators.

CT Texture analysis – CTTA
CT DICOM images reconstructed with 5 mm thickness (as suggested by developers of texture analysis software to minimize image noise) and soft-tissue kernel (B31 or B40) were exported from PACS to an external workstation and analysed using the proprietary texture analysis software TexRAD (TexRAD Ltd, www.texrad.com – part of Feedback Plc, Cambridge, UK). CTTA comprised a filtration-histogram technique where the initial filtration step using a band-pass Laplacian of Gaussian filter (similar to a non-orthogonal wavelet approach) extracted and enhanced objects/features of different sizes and intensity variation corresponding to the different spatial scale filters (SSF) = 0, 2, 3 and 4 mm (18), which corresponded to unfiltered/without-filtration, fine, medium and coarse texture scales, respectively (Figure 1). Quantification of textures without filtration (SSF=0) and with filtration (SSF=2, 3, 4) was performed using a number of statistical and histogram parameters such as mean gray level intensity, standard deviation (SD), entropy, mean of positive pixels, skewness, kurtosis, normalized standard deviation and normalized entropy (see Supplementary material). CTTA was performed by a single radiologist experienced in texture analysis and oncological imaging on manually drawn regions of interest (ROIs) encompassing the entire lesion avoiding cavitative components(18). Each lesion was segmented on a single slice displaying its largest cross-sectional area. Pixels with Hounsfield unit values below 0 and above 300 were excluded to remove the impact of air spaces, fat tissue and calcifications from the analysis.

**Statistical analysis**

Relevant texture features were selected using Least Absolute Shrinkage and Selection Operator (LASSO) with 10-fold internal cross-validation in order to prevent model over-fitting. For each of the selected features, LASSO calculates a coefficient, whose magnitude reflects the strength of the relationship between parameters and outcome. Positive coefficients indicate a positive correlation with the risk of progression, whereas negative coefficients suggest a protective effect. A Cox proportional hazards model was built using selected texture features and its performance was assessed by Area Under the Curve (AUC) using 200 bootstrap samples to estimate the range of AUC. Hazard ratios were calculated using the linear combination of the selected texture feature coefficients as independent variable (overall radiomic coefficient). Kaplan–Meier curves were plotted splitting the sample by the median predicted risk at 6 months. An additionary optimal cutoff was calculated by minimizing the p-value of the log-rank test between high- and low-risk groups. Predictive and prognostic value of age, sex, performance status (ECOG class) and M status according to the 8th TNM were tested separately with univariate Cox models. The significance level was fixed at 0.05. Statistical analysis was performed in R 3.3.0 (23) using the “hdnom” and “glmnet” packages,
Results

Patients’ characteristics
Fifty patients met the inclusion criteria. Median follow-up was 37 months. Patients’ characteristics are described in detail in Table 1. None of the patients discontinued TKI due to unacceptable toxicity. Thirty-one patients progressed and 24 died during follow-up. Median PFS was 10.5 months while median OS was 20.7 months.

Texture-based model and survival analysis
LASSO selected five texture features correlated with PFS: kurtosis SSF=0, entropy SSF=3 mm, mean SSF=4 mm, skewness SSF=4 mm and kurtosis SSF=4 mm. Table 2 shows the coefficients for each variable. The overall radiomic coefficient had a median value of −0.2552 and ranged between −1.101 and 1.041. The texture-based model predicted PFS at 6 months with an AUC of 0.8 and at 1 year with an AUC of 0.76. Fluctuations of model performance over time (between 6 months and 3 years) are shown in Figure 2. The half population with an overall radiomic coefficient above the median value had a 4.63-fold (95%CI 2.08–10.34) higher risk of progression while the optimal cut-off point of −0.2949 identified a group of 28 patients with a 5.78-fold higher risk of progression (95%CI 2.32–14.42).

Kaplan–Meier curves for PFS were built using the five-feature non-invasive signature splitting the sample by the median predicted risk at 6 months (Figure 3). The log-rank p-value was <0.001. Further details on the model are described in the Supplementary material. No texture feature correlated significantly with OS. None of the clinical parameters correlated with PFS or OS (Table 3).

Discussion

Quantitative analysis of CT images, which is the basis of radiomics and includes texture analysis, produced in the last years promising results, encompassing noninvasive genomic and phenotypic
signature, prognostication and treatment response prediction in many oncological settings[]. Among these, lung cancer has probably been the most studied[].

This retrospective study investigated the potential role of CT texture analysis (CTTA) in the prognostic stratification of lung adenocarcinoma patients treated with TKIs. The filtration-histogram technique of CTTA was proposed by Miles and Ganeshan (24) and highlights relatively hyperdense (higher contrast-enhancement) areas with a predefined size-scale and suppresses relatively hypodense (lower contrast-enhancement) areas with the same size-scale within the tumor. Compared to other texture analysis methods, it produces a relatively low number of parameters which reflect quite intuitive image characteristics (25). Furthermore, image filtration makes parameters less sensitive to intrinsic image noise that largely varies according to acquisition parameters and different scanners (14).

Machine learning belongs to a field of computational approaches directed to help the radiologists and the clinicians in taking care of lung cancer patients. Many fields in which there are promising results are computer aided diagnosis and detection, genomic classification and prognosis stratification(26). Texture analysis among other techniques of unsupervised feature extraction has been showing promising results in different oncologic settings(20,27) and more specifically in assessing prognosis or therapy response in lung cancer patients(28–30). In this study, a non-invasive texture-based signature was built using a machine learning method (31). A non-invasive signature was produced composed of five texture features, which were strongly associated with PFS. In this signature, skewness SSF = 4 mm, kurtosis SSF = 4 mm and kurtosis on unfiltered images (in decreasing order of magnitude) are associated with a poorer response to TKI while entropy had a negative relationship. Even if a conclusive interpretation of this result is not feasible, a possible explanation can be deduced relying on a previous computer modeling study (25) in which the relationship between image features (related to heterogeneity) and texture parameters derived from a filtration-histogram-based technique of CTTA was assessed. Skewness is positively related to the brightness of highlighted objects and tends to zero when the number of highlighted objects is high (very heterogeneous images) due to averaging; kurtosis is inversely related to the number of highlighted objects; mean reflects the intensity of highlighted objects; entropy is a metric of tumor heterogeneity. Relying on these assumptions, we could conceive that tumors with relatively highly dense highlighted objects and relatively homogeneous texture (low number of highlighted objects) are less likely to respond to TKI. On the other hand, highly heterogeneous tumors tend to respond better and may reflect angiogenesis activation via a collateral pathway of EGFR (32), which is the target of TKIs. Therefore, high tumor heterogeneity could indirectly indicate a higher EGFR mutational content, which is associated with
better treatment response to TKI compared to tumors with a lower mutational content, as demonstrated by quantitative polymerase chain reaction (12).

Interestingly, texture features were not associated with OS. All patients underwent TKI as first-line chemotherapy and after progression most of them were submitted to further lines of therapy with variable responses. For this reason, the mismatch of predictive power for PFS and OS suggests that texture characteristics are not associated to overall tumor aggressiveness (in this case they would have been predictive of both PFS and OS) but more specifically of tumor responsiveness to TKI. If confirmed on larger series, this could help oncologists in decision making. For example, TKI could be administered beyond minimal radiological progression in patients with a good texture-based probability of response; on the other hand, patients with a low texture-based probability of response might be addressed or switched earlier to different treatment schedules, including standard chemotherapy, new target drugs or combined therapies which will be available in the future (33).

Recently, Song et al. (34) published a study on a large multicenter series of 314 patients with metastatic EGFR-mutant NSCLC treated with TKI: in these patients, a radiomic signature of 12 quantitative CT features showed a good predictive accuracy and allowed a group of non-responsive patients to be identified whose PFS was similar to that of a cohort of EGFR-mutant patients treated with standard chemotherapy. Our study, which would be the second one published on this topic, offers an additional clue that quantitative CT parameters (in our case, specifically texture parameters) could have the very intriguing potential to predict response to TKI in terms of PFS; on the other hand, there are some differences between these two studies which are worth mentioning. Certainly, the very large number of patients enrolled in Song’s study is a decisive strength compared to our study. The comprehensive radiomic approach used by Song et al. led to more than 1,000 parameters being computed per patient, while the filtration-histogram method we adopted is less complex and produces a small number of parameters (30 primary parameters and 12 size-normalized parameters); the non-invasive signature proposed by Song et al. included 12 parameters, whereas ours included 5.

Theoretically, when model complexity (due to a high number of parameters) exceeds a threshold, bias on the training set decreases but variance (variability of prediction among different datasets) increases (35,36); for this reason, between two models with equal performance, the simpler one should be more confidently chosen.

In the future, it may be hypothesized that advancements in the field of machine learning and deep learning will dramatically improve performance and robustness of imaging-based predictive models and consequently their clinical impact. The progresses in this field might belong to three distinct areas: automated lesion detection and volumetric segmentation through effective recognition
algorithms; unsupervised detection of complex lesion patterns, for example using artificial neural networks with multiple hidden layers, which could enhance model predictive power; finally, deep learning could help to overcome biases related to the variability of texture features due to the use of different scanners and acquisition parameters by building advanced correction models for image standardization or (especially when using high throughput methods) identifying subsets of robust features. At the moment, the most important limiting factor to the diffusion of these revolutionary approaches seems to be the difficulty in collecting a sufficient number of patients (probably in the order of thousands) with sufficiently homogeneous clinical characteristics in order to build robust and clinically useful predictive models. In this really complex scenario, the simplicity of our method should not be considered itself as a limitation even if it does not overcome all the aforementioned issues.

This study has several limitations. First of all, the retrospective design and low sample size hamper the robustness of the results. In particular, low sample size did not allow the model to be tested on an external validation group while just an internal cross-validation for model building and a bootstrap approach for quantifying the uncertainty of the predictive model were feasible. Future studies will be addressed to externally validate these preliminary results. Second, texture analysis was performed only on primary tumors, which might have different phenotypic and biologic characteristics compared to metastases; on the other hand, metastases were often too small to be analysed or not suitable for an accurate segmentation (for example, bone metastases). Third, texture analysis was performed on a single slice which might not be representative of the entire tumor volume, even if a significant advantage of multislice over single slice texture analysis is controversial (37,38). Fourth, lack of pre-processing on native images with slightly different spatial resolution could theoretically influence the results of SSF=0 parameters, while this effect should be significantly mitigated on images after LoG filtering. Finally, a test–retest analysis was not performed to check the repeatability of texture parameters(39,40).

**Conclusions**

The results of this pilot study suggest that non-invasive tumor characterization by CT texture features could help to stratify patients with metastatic lung adenocarcinoma treated with TKIs. Unfortunately, sample size is too small to draw definitive conclusions. If these very preliminary results will be confirmed on larger external databases, filtration-histogram CTTA could be used as a predictive imaging biomarker of responsiveness to TKI in lung adenocarcinoma.
References


