Published in final edited form as:

Radiother Oncol. 2016 June ; 119(3): 480–486. doi:10.1016/j.radonc.2016.04.004.

Radiomic phenotype features predict pathological response in Non-Small Cell Lung Cancer

Thibaud P. Coroller, M.Sc.^{1,#}, Vishesh Agrawal, B.A.¹, Vivek Narayan, B.Sc.¹, Ying Hou, M.D.¹, Patrick Grossmann, M.Sc.¹, Stephanie W. Lee, B.A.¹, Raymond H. Mak, M.D.^{1,*}, and Hugo J.W.L. Aerts, Ph.D.^{1,2,*}

¹Department of Radiation Oncology, Dana-Farber Cancer Institute, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

²Department of Radiation Radiology, Dana-Farber Cancer Institute, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

Abstract

Background and purpose—Radiomics can quantify tumor phenotype characteristics non-invasively by applying advanced imaging feature algorithms. In this study we assessed if pre-treatment radiomics data are able to predict pathological response after neoadjuvant chemoradiation in patients with locally advanced non-small cell lung cancer (NSCLC).

Materials and Methods—127 NSCLC patients were included in this study. Fifteen radiomic features selected based on stability and variance were evaluated for its power to predict pathological response. Predictive power was evaluated using area under the curve (AUC). Conventional imaging features (tumor volume and diameter) were used for comparison.

Results—Seven features were predictive for pathologic gross residual disease (AUC > 0.6, p-value < 0.05), and one for pathologic complete response (AUC = 0.63, p-value = 0.01). No conventional imaging features were predictive (range AUC = 0.51–0.59, p-value > 0.05). Tumors that did not respond well to neoadjuvant chemoradiation were more likely to present rounder shape (spherical disproportionality, AUC = 0.63, p-value = 0.009) and heterogeneous texture (LoG 5 mm 3D - GLCM entropy, AUC = 0.61, p-value = 0.03).

Conclusion—We identified predictive radiomic features for pathological response, although no conventional features were significantly predictive. This study demonstrates that radiomics can provide valuable clinical information, and performed better than conventional imaging features.

[#]**Corresponding author.** Thibaud Coroller, MSc, Dana-Farber Cancer Institute, Brigham and Women's Hospital, Harvard Medical School, 450, Brookline Ave, JF518, Boston, MA 02115-5450, P – 781.690.4612, F - 617.582.6037, tcoroller@iroc.harvard.edu.

^{*}Equal contribution

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Disclaimer: None

Conflict of Interest Statement: None

Keywords

Radiomics; pathological response; NSCLC; biomarkers; quantitative imaging

Introduction

Radiomics is an emerging field of quantitative imaging that aims to describe tumors non-invasively using a large set of advanced imaging features [1–3]. These features can robustly create a unique phenotypic atlas for each tumor [4–6]. Associating clinical information to this atlas has enabled the identification of new, reproducible, image-based biomarkers which has been prognostic for clinical outcomes including overall survival [7–9] and distant metastasis [10]. Association was found with lung cancer patients of histology and stage [11] as well.

Lung cancer is the leading cause of cancer deaths worldwide [12]. Stage IIIA non-small cell lung cancer (NSCLC) can be treated using trimodality therapy that includes neoadjuvant chemoradiation followed by surgery according to NCCN guidelines [13]. However, trimodality therapy is controversial, given the observed lack of survival benefit in adding surgery compared to definitive chemoradiation alone, [14,15] which underscores the importance of identifying patients who respond completely to chemoradiation and do not require additional invasive local therapy.

Pathological response is a direct measure of tumor response to neoadjuvant chemoradiation assessed at time of surgery. It has the potential to be used as a surrogate endpoint [16] for survival/local control and has been shown to be prognostic for survival in early [17] and advanced [18] stages for NSCLC patients. Predicting pathological response at an early time point would allow modification of the treatment regimen (e.g. adding surgery versus intensifying chemoradiation) based on how the tumor is likely to respond and this adaptive approach could improve patient outcomes.

Currently, tumor response is clinically assessed using RECIST [19], which classifies changes in the sum of tumor and lymph nodes diameters on CT images before and after therapy. However, the radiographic response to chemoradiation for NSCLC tumors may be slow [20], which may limit the utility of RECIST in predicting pathological response at the end of the neoadjuvant chemoradiation shortly before surgery, and hence allow very little margin for clinicians to adapt the treatment regimen consequently.

In this study we investigated the power of pre-treatment CT-based radiomic features to predict pathological response after neoadjuvant chemoradiation. We compared these results to conventional volumetric features such as tumor volume and diameter.

Methods

Patient selection

Patients with stage II–III NSCLC treated at Dana-Farber Cancer Institute between 2001 and 2013 who were treated with neoadjuvant radiotherapy and chemotherapy (chemoradiation)

prior to surgical resection were included in this study. Patients with distant metastasis at presentation or delay in surgery greater than 120 days after the completion of chemoradiation were excluded. For all patients, CT imaging at the initiation of chemoradiation and prior to surgical resection was available. No exclusion based on histology was applied. A subset of patients received adjuvant therapy and was also included in this analysis. Finally, a total of 127 patients were included for this study.

Follow-up and endpoints

The main endpoint for this study was pathological response assessed at time of surgery. The amount of residual tumor was classified based on surgical pathology reports as pathologic complete response (pCR), microscopic residual disease (MRD) or gross residual disease (GRD). Percent residual tumor in the pathological sample was not available for this study. Three other clinical endpoints were included for this study including overall survival (OS), distant metastasis (DM) and in-field local recurrence (LR). The time associated with the endpoint was defined from treatment start date to date of first event. The last date of follow-up was used for patients with no events.

Follow-up chest CT scans with contrast (unless the patient had a contraindication to contrast, e.g. renal dysfunction or allergy) were performed every three to six months after treatment for patients at our institution based on US national guidelines [13] to assess tumor progression.

CT Acquisition and Segmentation

Planning CTs were acquired according to scanning protocol at our institution using GE “lightspeed” CT scanner (GE Medical System, Milwaukee, WI, USA). Tumor segmentation was performed on radiation therapy planning CTs using Eclipse (Varian Medical System, Palo Alto, CA, USA). The primary tumor site was retrospectively contoured guided by existing treatment planning contours. Using both soft tissues and lung windows, air, vessels, normal tissue or surrounding organs were subsequently excluded from the contours (Figure 1.A). All contours were done manually (T.P.C., V.A., Y.H.), and then all individually verified by an expert radiation oncologist by (R.H.M.).

Features Extraction and Selection

Radiomic features describing tumor phenotype were extracted ($m=1603$) from the primary tumor site with an in-house Matlab 2013 (The Mathworks Inc., Natick, Massachusetts, United States) toolbox and the software 3D Slicer 4.4.0 [21] (Figure 1.B). Average voxel spacing was ($0.9\text{mm} \times 0.9\text{mm} \times 3\text{mm}$) respectively for (x,y,z) and was resampled $3 \times 3 \times 3\text{mm}^3$ prior to feature extraction to have standardized voxel spacing across the cohort. A bin width of 25 Hounsfield units (HU) was used for textural features. All features are described in the supplement of a previous study [10].

Fifteen Radiomic features were selected based on stability and variance for this study (features selection is described in Supplement I). Additionally, we defined three conventional, pre-treatment, clinically utilized, volumetric features for comparison to advanced phenotypic features prior chemoradiation. These features consisted of tumor

volume, 2D axial maximal diameter and 3D maximal diameter. 2D axial maximal diameter corresponds to the greatest diameter in the axial plane. 3D maximal diameter refers to the greatest diameter in any direction. All volume and diameter measurements were obtained from the primary tumor and did not include the sum diameters or volumes of involved lymph nodes.

Statistical Analysis

All statistical analyses were done on R software [22] version 3.1.3. Predictive performance of these remaining features were assessed using the “survcomp” package [23,24] version 1.16 from Bioconductor [25]. We computed receiver operating characteristic (ROC) area under the curve (AUC) for binary outcomes. Predictive power was reported as proportional or disproportionate to the risk of experiencing the response as the feature value is increasing.

Difference for clinical categories was assessed using chi-square or two-sided Wilcoxon-test respectively for categorical or continuous variables. Noether test was used to consider AUC significance from 0.5 (random). Survival and disease-free probability curve were computed using Kaplan-Meier analysis. A three year estimate was reported for the analysis. Log-rank test was used to assess difference in probability curves between pathological response groups. A p-value below 0.05 was considered as significant. Features with an AUC above 0.60 and a p-value below 0.05 were considered predictive.

Multivariate models were made using logistic regression for pathological response using the same subgroup for the univariate analysis to compare their performance. Three models were created with 1) Conventional (volume and axial/3D diameters), 2) Radiomics (predictive features for GRD) and 3) Combined (Conventional + radiomics) features.

We compare model performance with the validation AUC using the cross validation (CV). The cohort was split, using 80% for training and 20% for validation for each CV (for each 100 iterations). Patients were randomized using a conservative random split using the “caret” package [26]. Difference between the CV models performance was done using a permutation test. The outcome labels were randomly resampled ($k=1000$ times) and a new CV was computed for each random label combination. One-sided Wilcoxon test was computed for each random label models, the W_k statistic extracted and compared to W_0 (true label) to assess if a model performance was significantly greater than another.

Results

127 patients with NSCLC were included in this study. The median age was 60.5 years (range 32.7 to 77.6 years), with a majority of women (53.5%) and white (92.1%). Tumor histology was predominantly adenocarcinoma (56.6%) and AJCC [27] stage IIIA (75.6%). The median follow-up was 41.8 months (range 2.7–117.2). The distribution of pathological response was 27 (21.3%), 33 (26.0%) and 67 (52.7%) respectively for complete response, microscopic and gross residual disease.

All treatment information can be found in the Table 1. Comparison between pathological complete response (pCR) versus microscopic (MRD) and gross (GRD) residual disease,

showed no significant differences between treatment (p-value=0.49, Chi-square test) and radiation dose (p-value=0.32). Significant differences were found between overall stages (p-value= 0.02) and histology (p-value < 0.001), likely driven by the fact that the distribution is skewed for histology and overall stage.

Relationship between clinical outcomes and pathological response subgroups was investigated (Table 2). The median (range) for overall survival, distant metastasis and local recurrence was respectively 41.8 (2.7–117.2), 10.8 (2.5–73.5) and 14 (4.7–66.5) months. No significant difference was observed for survival between pathological response (p-value =0.86, Log-rank test). However, pCR patients had significantly higher probabilities at three years for distant metastasis-free (79%, p-value = 0.036) and local recurrence-free (94%, p-value = 0.013). Kaplan-Meier curves can be found in Figure S3 and concordance index for radiomics features in Figure S4 in Supplement II.

Fifteen features were selected based on stability and variance (see Supplement I) and were evaluated for performance to predict clinical outcomes. Those features included one shape, seven statistics and seven textural features. Textural features incorporated four gray-level co-occurrence matrix (GLCM) sensitive to voxel patterns and three gray-level size zone matrix (GLSZM) sensitive to flat zone (area of connecting voxel with same value). All these features are described in Table 3.

Pathological response was our primary clinical endpoint. We first determined if radiomic features could identify tumors likely to respond poorly (GRD) vs. tumors likely to respond well (pCR and MRD) to the chemoradiation (Figure 2.A). The fifteen selected advanced imaging features had an AUC of 0.53 to 0.66 for GRD (described in Table S2 in Supplement II). Seven features were significantly predictive (range AUC 0.61–0.66, p-value <0.05) for GRD. From those seven predictive features, two were risk proportionate as the feature value increases (GLCM entropy and stats root mean square) and five disproportionate (mean and skewness in voxel intensity histogram, spherical disproportionality and two GLSZM large area emphasis) to experience GRD. No conventional volumetric imaging features were significant from random or above at the threshold of AUC > 0.6 (range 0.57 to 0.59, p-value >0.05) and all were disproportionate to the risk of GRD.

We then investigated the predictive power for identifying pathologic complete response (pCR) vs. non-complete response (MRD and GRD). The AUC range of radiomic features (Figure 2.B) was 0.52–0.63 and 0.51–0.55 for conventional features (described in Table S3 in Supplement II). The best performing radiomic feature, Wavelet HLL mean, was also the only significantly predictive feature (AUC = 0.63, p-value = 0.01, Noether test) and was risk proportional. No conventional imaging features were predictive for pCR (range 0.51 to 0.55, all p-value>0.05).

Multivariate models were created for each set of features (Figure 3), including conventional (3 features), radiomics (7 predictive features for GRD) and the resulting combined set (10 features). The median AUC values of the cross validation were 0.57, 0.65 and 0.65 for GRD and 0.60, 0.61 and 0.68 for pCR respectively for conventional, radiomics and combined models. The combined and radiomics model for GRD performed significantly better on the

cross validation compare to clinical features alone. For pCR, no significant difference was found between radiomics and conventional model performance. However, the combined model significantly outperformed both radiomics and conventional features for predication of pCR.

Discussion

Radiomics [1] is an emerging field of quantitative imaging that aims to extract phenotypic tumor information from clinical imaging data. In this study we demonstrate the predictive power of radiomic phenotypic features for pathological response in patients with NSCLC. Pathological response is a standard endpoint, assessed at time of surgery for a direct measure of neoadjuvant chemotherapy effect. Pathologic response was significantly associated with clinical outcomes in our study (distant metastasis and local recurrence) showing the importance of predicting pathological outcome. Using data mining techniques, we selected fifteen radiomics features based on stability and variance. From those features, we identify seven features predictive of pathological gross residual disease (GRD) and one feature predictive of pathological complete responders (pCR). In contrast, no pre-treatment conventional clinically utilized features (volume or diameter) were predictive for pathological response. Lastly, the radiomics model also performed significantly better on a cross validation than the conventional model for GRD. A combined model using conventional and radiomics outperformed for pCR.

The seven significant predictive features for GRD enabled us to represent phenotypic characteristics of a lung tumor that is less likely to respond to neoadjuvant chemoradiation. Spherical disproportionality (a measure of the similarity between the tumor and a sphere with an equivalent volume) of the primary tumor site was associated with pathological response (AUC above 0.6 for both GRD and pCR). Remarkably, more complex tumors are likely to be associated with pCR while more spherical tumors are more likely to be associated with GRD. In comparison, conventional radiographic features such as volume and diameter were not predictive for pathological response (neither pCR nor GRD), and may not capture sufficient shape information as advanced shape radiomic features. It is noteworthy that larger tumor dimension (both volume and diameter) appeared to be associated with pCR. Tumor with large flat zones (area of connecting voxels of similar intensity) were associated with pCR whereas tumor with rich complex patterns (heterogeneities) were associated with GRD. Finally based on the common significant predictor for both pCR and GRD (Wavelet HLL – mean) indicated that tumor with overall lower voxel intensity (darker) were associated with GRD after Wavelet HLL filtering. Multivariate performances demonstrated that radiomics information added complementary information for pCR (significantly greater performance for the combined model) and outperformed conventional features alone for GRD.

To our knowledge this is the first study investigating pathological response to neoadjuvant chemoradiation using advanced quantitative imaging features from CT images. A previous study from Ravanelli *et al.* [28] investigated response of first line of chemotherapy (with no concurrent radiation therapy or surgery) for patients with advanced NSCLC using two textural features (uniformity and grey-level) from contrast enhanced CT images. They found

an AUC of 0.74 for their multivariate model (Leave one out cross validation). Conventional features (tumor volume and diameters) have been identified as a prognostic factor in prior studies [29–31]. However patients in those reports did not receive surgical resection as part of their therapy as did patients in our study. Conventional features may thus be more relevant as a prognostic factor in the absence of surgical intervention. Additionally, we did not find any significant association between radiomics and clinical outcomes, supporting the idea that surgical resection of the tumor undermines image-based features prediction for clinical outcomes. We believe that the underlying tumor phenotype may be more relevant for identifying pathologic response than either volume or diameter, although these may remain valuable for assessing response to therapy. Other studies [32,33] investigated the association of conventional features (tumor volume) but demonstrated lack of correlation between measured tumor volume change and pathologic response in NSCLC, supporting the fact that none of our conventional features were significant.

Limitations of this study include the cohort size (n=127). Due to a large number of available radiomic features and concern for multiple testing, dimension reduction was used with restrictive criteria to perform drastic selection, therefore excluding potential predictive features for any clinical outcomes. However despite this very restrictive data-driven approach selection (about 1% remaining after), we were able to find predictive features. Additionally, the incidence of pCR (n = 27) was much lower than GRD (n = 67), potentially explaining the differences in the predictive power for pathological complete response. Since the study period was conducted over 12 years, we recognize that there is heterogeneity in the treatments delivered to the patient cohort. However, we did not find any association between radiation dose, cycles of chemotherapy or other treatment characteristics and pathologic response. Thus, we believe the results presented remain valid despite treatment heterogeneity. A major limitation of radiomics is the lack of standardization in image acquisition that without access to raw images is difficult to control and hence standardize patient cohorts. Yet, despite the noise introduced by variation in acquisition protocols, radiomic models have shown consistent and reproducible association with outcomes in multiple independent datasets [7,8,9–11]. Furthermore, we have previously demonstrated the reproducibility [7] of the radiomics features using the RIDER [34] test / retest dataset and we resampled of imaging using a common pixel spacing (3×3×3mm3) to limit variability across patients. Finally, we acknowledge the limitations of the clinical applicability of this study. This study was initiated to find potential patterns in tumor phenotype that could predict pathological response prior to the start of therapy. Although several features were significantly predictive for pathological endpoints, we were unable to identify subgroups associated with overall survival, local recurrence or distant metastasis using radiomic features. Despite the need for further validation sets, this study provides a basis for additional research (e.g using PET-CT features) that could improve performance. We believe such radiomics based analyses can be used as a complementary method of patient stratification for NSCLC prior to the initiation of therapy as is currently being investigated in other disease sites such as breast [35,36], colorectal cancer [37], and glioblastoma [38].

In conclusion, we identified CT-based radiomic features predictive for pathological response in patient with locally advanced NSCLC who previously received trimodality therapy. This study demonstrates that radiomics can provide additional phenotypic information that may

reflect underlying tumor sensitivity to chemoradiation. Predicting pathological response prior to initiation of neoadjuvant chemoradiation has significant potential clinical applications such as developing adaptive therapy based on pre-treatment tumor phenotype. These radiomic features which can be captured from clinically-available imaging modalities performed better than conventionally reported metrics.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: Authors acknowledge financial support from the National Institute of Health (NIH-USA U24CA194354, and NIH-USA U01CA190234). This project was partially funded by the Kaye Scholar Award and the Brigham and Women's Hospital Department of Radiation Oncology Clinical Translational Grant.

References

1. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur. J. Cancer*. 2012; 48:441–446. [PubMed: 22257792]
2. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. *Magn. Reson. Imaging*. 2012; 30:1234–1248. [PubMed: 22898692]
3. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2015:151169.
4. Parmar, C.; Rios Velazquez, E.; Leijenaar, R.; Jermoumi, M.; Carvalho, S.; Mak, RH., et al. Robust Radiomics Feature Quantification Using Semiautomatic Volumetric Segmentation. In: Woloschak, GE., editor. *PLoS ONE*. Vol. 9. 2014. p. e102107
5. Rios Velazquez E, Aerts HJWL, Gu Y, Goldgof DB, De Ruyscher D, Dekker A, et al. A semiautomatic CT-based ensemble segmentation of lung tumors: Comparison with oncologists' delineations and with the surgical specimen. *Radiother. Oncol*. 2012; 105:167–173. [PubMed: 23157978]
6. Leijenaar RTH, Carvalho S, Velazquez ER, van Elmpt WJC, Parmar C, Hoekstra OS, et al. Stability of FDG-PET Radiomics features: An integrated analysis of test-retest and interobserver variability. *Acta Oncol*. 2013; 52:1391–1397. [PubMed: 24047337]
7. Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Cavalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat. Commun*. 2014:5.
8. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine Learning methods for Quantitative Radiomic Biomarkers. *Sci. Rep*. 2015; 5:13087. [PubMed: 26278466]
9. Leijenaar RT, Carvalho S, Hoebbers FJ, Aerts HJ, van Elmpt WJ, Huang SH, et al. External validation of a prognostic CT-based radiomic signature in oropharyngeal squamous cell carcinoma. *Acta Oncol*. 2015:1–7.
10. Coroller TP, Grossmann P, Hou Y, Rios Velazquez E, Leijenaar RTH, Hermann G, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. *Radiother. Oncol*. 2015; 114:345–350. [PubMed: 25746350]
11. Parmar C, Leijenaar RTH, Grossmann P, Rios Velazquez E, Bussink J, Rietveld D, et al. Radiomic feature clusters and Prognostic Signatures specific for Lung and Head & Neck cancer. *Sci. Rep*. 2015; 5:11044. [PubMed: 26251068]
12. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014: Cancer Statistics, 2014. *CA. Cancer J. Clin*. 2014; 64:9–29. [PubMed: 24399786]
13. National Comprehensive Cancer Network. NCCN guidelines : Non-Small Cell Lung Cancer (Version 7.2015). 2015 http://www.nccn.org/professionals/physician_gls/pdf/nscl.pdf.

14. Albain KS, Swann RS, Rusch VW, Turrisi AT, Shepherd FA, Smith C, et al. Radiotherapy plus chemotherapy with or without surgical resection for stage III non-small-cell lung cancer: a phase III randomised controlled trial. *The Lancet*. 2009; 374:379–386.
15. van Meerbeeck JP, Kramer GWPM, Van Schil PEY, Legrand C, Smit EF, Schramel F, et al. Randomized Controlled Trial of Resection Versus Radiotherapy After Induction Chemotherapy in Stage IIIA-N2 Non-Small-Cell Lung Cancer. *JNCI J. Natl. Cancer Inst.* 2007; 99:442–450. [PubMed: 17374834]
16. Hellmann MD, Chaft JE, William WN, Rusch V, Pisters KM, Kalhor N, et al. Pathological response after neoadjuvant chemotherapy in resectable non-small-cell lung cancers: proposal for the use of major pathological response as a surrogate endpoint. *Lancet Oncol.* 2014; 15:e42–e50. [PubMed: 24384493]
17. Mouillet G, Monnet E, Milleron B, Puyraveau M, Quoix E, David P, et al. Pathologic complete response to preoperative chemotherapy predicts cure in early-stage non-small-cell lung cancer: combined analysis of two IFCT randomized trials. *J. Thorac. Oncol. Off. Publ. Int. Assoc. Study Lung Cancer.* 2012; 7:841–849.
18. Isobe K, Hata Y, Sakaguchi S, Sato F, Takahashi S, Sato K, et al. Pathological response and prognosis of stage III non-small cell lung cancer patients treated with induction chemoradiation. *Asia Pac. J. Clin. Oncol.* 2012; 8:260–266. [PubMed: 22897895]
19. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur. J. Cancer.* 2009; 45:228–247. [PubMed: 19097774]
20. Werner-Wasik M, Xiao Y, Pequignot E, Curran WJ, Hauck W. Assessment of lung cancer response after nonoperative therapy: tumor diameter, bidimensional product, and volume. A serial CT scan-based study. *Int. J. Radiat. Oncol. Biol. Phys.* 2001; 51:56–61. [PubMed: 11516851]
21. Pieper S, Halle M, Kikinis R. 3D Slicer. *IEEE Int. Symp. Biomed. Imaging Nano Macro* 2004. 2004; 1:632–635.
22. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013. ISBN 3-900051-07-0, URL <http://www.R-project.org/>
23. Haibe-Kains B, Desmedt C, Sotiriou C, Bontempi G. A comparative study of survival models for breast cancer prognostication based on microarray data: does a single gene beat them all? *Bioinformatics.* 2008; 24:2200–2208. [PubMed: 18635567]
24. Schröder MS, Culhane AC, Quackenbush J, Haibe-Kains B. survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. *Bioinformatics.* 2011; 27:3206–3208. [PubMed: 21903630]
25. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 2004; 5:R80. [PubMed: 15461798]
26. Kuhn M. Building predictive models in R using the caret package. *J. Stat. Softw.* 2008; 28:1–26.
27. American Joint Committee on Cancer. *AJCC Cancer Staging Manual*. 7th. New York: Springer; 2010. Available from: American Joint Committee on Cancer. *AJCC Cancer Staging Manual*, 7th ed. New York: Springer 2010
28. Ravanelli M, Farina D, Morassi M, Roca E, Cavalleri G, Tassi G, et al. Texture analysis of advanced non-small cell lung cancer (NSCLC) on contrast-enhanced computed tomography: prediction of the response to the first-line chemotherapy. *Eur. Radiol.* 2013; 23:3450–3455. [PubMed: 23835926]
29. Bradley JD, Ieumwananonthachai N, Purdy JA, Wasserman TH, Lockett MA, Graham MV, et al. Gross tumor volume, critical prognostic factor in patients treated with three-dimensional conformal radiation therapy for non-small-cell lung carcinoma. *Int. J. Radiat. Oncol. Biol. Phys.* 2002; 52:49–57. [PubMed: 11777621]
30. Alexander BM, Othus M, Caglar HB, Allen AM. Tumor Volume Is a Prognostic Factor in Non-Small-Cell Lung Cancer Treated With Chemoradiotherapy. *Int. J. Radiat. Oncol.* 2011; 79:1381–1387.

31. Stinchcombe TE, Morris DE, Moore DT, Bechtel JH, Halle JS, Mears A, et al. Post-chemotherapy gross tumor volume is predictive of survival in patients with stage III non-small cell lung cancer treated with combined modality therapy. *Lung Cancer Amst. Neth.* 2006; 52:67–74.
32. Cerfolio RJ, Bryant AS, Winokur TS, Ohja B, Bartolucci AA. Repeat FDG-PET After Neoadjuvant Therapy is a Predictor of Pathologic Response in Patients With Non-Small Cell Lung Cancer. *Ann. Thorac. Surg.* 2004; 78:1903–1909. [PubMed: 15560998]
33. Poettgen C, Theegarten D, Eberhardt W, Levegruen S, Gauler T, Krbek T, et al. Correlation of PET/CT Findings and Histopathology after Neoadjuvant Therapy in Non-Small Cell Lung Cancer. *Oncology.* 2007; 73:316–323. [PubMed: 18497503]
34. Zhao B, James LP, Moskowitz CS, Guo P, Ginsberg MS, Lefkowitz RA, et al. Evaluating Variability in Tumor Measurements from Same-day Repeat CT Scans of Patients with Non- Small Cell Lung Cancer 1. *Radiology.* 2009; 252:263–272. [PubMed: 19561260]
35. Pickles MD, Lowry M, Gibbs P. Pretreatment Prognostic Value of Dynamic Contrast- Enhanced Magnetic Resonance Imaging Vascular, Texture, Shape, and Size Parameters Compared With Traditional Survival Indicators Obtained From Locally Advanced Breast Cancer Patients: *Invest. Radiol.* 2016; 51:177–185.
36. Fox MJ, Gibbs P, Pickles MD. Minkowski functionals: An MRI texture analysis tool for determination of the aggressiveness of breast cancer. *J. Magn. Reson. Imaging JMRI.* 2015
37. Lubner MG, Stabo N, Lubner SJ, Del Rio AM, Song C, Halberg RB, et al. CT textural analysis of hepatic metastatic colorectal cancer: pre-treatment tumor heterogeneity correlates with pathology and clinical outcomes. *Abdom. Imaging.* 2015; 40:2331–2337. [PubMed: 25968046]
38. Lee J, Jain R, Khalil K, Griffith B, Bosca R, Rao G, et al. Texture Feature Ratios from Relative CBV Maps of Perfusion MRI Are Associated with Patient Survival in Glioblastoma. *AJNR Am. J. Neuroradiol.* 2016; 37:37–43. [PubMed: 26471746]

Highlights

- Early prediction of response to neoadjuvant chemoradiation is crucial for improving overall treatment and patient outcomes.
- This study demonstrated an association between Radiomic features and pathological response for lung cancer patients.
- A Radiomics-Conventional combined model shown better performance for pathological response.

Figure 1.

Radiomic analysis workflow description: A) Lung primary tumor site was manually contoured from treatment planning images (shown in green on CT image on the left and the 3D mask on the right). B) All images were subsequently resampled and from those contours the radiomic features describing tumor phenotype were extracted using three feature groups: Shape, Statistic and Textural features, with and without Wavelet and Laplacian of Gaussian filtering. C) Finally, association between radiomic features and the clinical outcomes were investigated for image-based biomarkers discovery.

Figure 2.

AUC of Radiomic features and conventional volumetric imaging features for A) poor responders (gross residual disease) vs. good responders (pathologic complete response and microscopic residual disease) and B) pathological complete responders vs. non-complete responders (microscopic and gross residual disease). Predicting power was reported as proportional or disproportionate to the risk of experiencing the response as the feature value is increasing. Features reported with a “*” are significant from random (Noether test, p-value < 0.05). Legend colors indicates feature group.

Figure 3.

Comparison of multivariate models for A) Gross residual disease and B) complete pathological response. AUC from the validation is reported from the Cross-Validation (100 iterations, 80% training and 20% validation) for each model. Comparison between models were done using a permutation test. A “*” is reported if the model performance is significantly greater than the other, else “ns” for non-significant.

Table 1

Patient and treatment characteristics. Median (range) is reported for continuous and counts (percentage) for categorical variables. Statistical difference between complete pathological responders vs. non-complete responders was computed using Chi-Square or Wilcoxon-test respectively for categorical and continuous variables.

Variable	Group	Median (Range) / Count (%)	pCR (n=27)	MRD & GRD (n=100)	P-value
Age [years]		60.5 (32.7 – 77.6)	61.5 (32.7 – 75.2)	60.4 (33.1 – 77.6)	0.93
Performance Status	0	60 (47.2%)	10 (37.0%)	50 (50%)	0.33
	1	59 (46.5%)	14 (51.9%)	45 (45%)	
	2–3	8 (6.3%)	3 (11.1%)	5 (5%)	
Gender	Female	68 (53.5%)	11 (40.7%)	57 (57%)	0.19
	Male	59 (46.5%)	16 (59.3%)	43 (43%)	
Ethnicity	White	117 (92.1%)	25 (92.6%)	92 (92%)	0.61
	Black	5 (3.9%)	1 (3.7%)	4 (4%)	
	Hispanic	3 (2.4%)	0 (0%)	3 (3%)	
	Asian	2 (1.6%)	1 (3.7%)	1 (1%)	
Histology	Adenocarcinoma	72 (56.6%)	5 (18.5%)	67 (67%)	<0.001
	Squamous cell Carcinoma	32 (25.3%)	14 (51.9%)	18 (18%)	
	NSCLC	18 (14.2%)	6 (22.2%)	12 (12%)	
	Others ¹	5 (3.9%)	2 (7.4%)	3 (3%)	
Overall stages	IIA	2 (1.5%)	1 (3.7%)	1 (1%)	0.02
	IIB	8 (6.3%)	5 (18.5%)	3 (3%)	
	IIIA	96 (75.6%)	18 (66.7%)	78 (78%)	
	IIIB	21 (16.6%)	3 (11.1%)	18 (18%)	
Treatment sequence	Concurrent	111 (87.4%)	24 (88.8%)	87 (87%)	0.49
	Induction	14 (11.0%)	2 (7.4%)	12 (12%)	
	Others ²	2 (1.6%)	1 (3.8%)	1 (1%)	

Variable	Group	Median (Range) / Count (%)	pCR (n=27)	MRD & GRD (n=100)	p-value
Radiation Dose [Gy]		54 (45–70)	54 (46–70)	54 (45–70)	0.32
Pathological response	pCR	27 (21.3%)	27 (100%)	0 (0%)	-
	MRD	33 (26.0%)	0 (0%)	33 (33%)	
	GRD	67 (52.7%)	0 (0%)	67 (67%)	

Label: pCR = pathologic complete response, MRD=microscopic residual disease, GRD = gross residual disease.

¹ Large Cell neuroendocrine carcinoma, Mixed NSCLC and SCLC, Adenoid cystic carcinoma, Sarcomatoid.
² RT Only, Sequential.

Table 2

Three years estimate from Kaplan-Meier survival curve for each pathological response subgroup. Difference between groups was assessed with Log-Rank test.

Three years estimate probability	pCR (n=27)	MRD (n=33)	GRD (n=67)	<i>p-value</i>
Overall Survival	72%	53%	52%	<i>0.86</i>
Distant Metastasis Free	79%	59%	50%	<i>0.036</i>
Local Recurrence Free	94%	75%	62%	<i>0.013</i>

Label: pCR = pathologic complete response, MRD=microscopic residual disease, GRD = gross residual disease.

Table 3

Description of radiomic features with associated feature group and filter.

Selected Radiomic feature	Radiomic group	Filter associated	Description
Sphere Disproportionality	Shape	None	Ratio between tumor area and a sphere with the same volume as the tumor
Root Mean Square	Statistics	Wavelet HLL	Root mean square of the voxels intensity value
Range	Statistics	Wavelet LLH Wavelet LHH	The range of voxels intensity values
Energy	Statistics	Wavelet HLL	Describe the energy of the image
Mean	Statistics	Wavelet HLL	The mean voxel intensity value
Kurtosis	Statistics	LoG 3D - 5mm	Describe the shape of a probability distribution of the voxel intensity histogram
Skewness	Statistics	LoG 3D - 4mm	Describe the shape of a probability distribution of the voxel intensity histogram
Correlation	GLCM	Wavelet LHH	Correlation of the GLCM matrix
Entropy	GLCM	LoG 3D - 5mm	Complexity of the GLCM matrix (sensitive to the number of unique voxel patterns in the tumor)
Homogeneity 2	GLCM	LoG 2D - 4mm	Homogeneity of voxels patterns (similar patterns across the whole tumor)
Cluster Prominence	GLCM	LoG 3D - 3mm	Sensitive to flat zone (area of connecting voxel with same value)
Low Intensity Large Area Emphasis	GLSZM	Wavelet LHH	Sensitive to flat zone with low intensity voxel (e.g. necrotic area)
Large Area Emphasis	GLSZM	LoG 3D - 5mm	Sensitive to flat zone
High Intensity Large Area Emphasis	GLSZM	LoG 3D - 5mm	Sensitive to flat zone with high intensity voxel (e.g. calcifications, blood vessels)

Label: GLCM = Gray-level Co-occurrence Matrix, GLSZM = Gray-Level Size Zone Matrix, LoG=Laplacian of Gaussian, L=Low, H=High.