Radiotherapy and Oncology xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original article

Radiomic phenotype features predict pathological response in non-small cell lung cancer

Thibaud P. Coroller [a,*], Vishesh Agrawal [a], Vivek Narayan [a], Ying Hou [a], Patrick Grossmann [a], Stephanie W. Lee [a], Raymond H. Mak [a,1], Hugo J.W.L. Aerts [a,b,1]

[a] Department of Radiation Oncology; and [b] Department of Radiology, Dana-Farber Cancer Institute, Brigham and Women's Hospital, Harvard Medical School, Boston, USA

ARTICLE INFO

Article history:
Received 9 September 2015
Received in revised form 3 March 2016
Accepted 3 April 2016
Available online xxxx

Keywords:
Radiomics
Pathological response
NSCLC
Biomarkers
Quantitative imaging

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 a 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.

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 node 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

* Corresponding author at: Dana-Farber Cancer Institute, Brigham and Women's Hospital, Harvard Medical School, 450 Brookline Ave, JF518, Boston, MA 02115-5450, USA.
E-mail address: tcoroller@lroc.harvard.edu (T.P. Coroller).
[1] Equal contribution.

http://dx.doi.org/10.1016/j.radonc.2016.04.004
0167-8140/

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 (Fig. 1A). All contours were done manually (T.P.C., V.A., Y.H.), and then all individually verified by an expert radiation oncologist (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] (Fig. 1B). Average voxel spacing was 0.9mm×0.9mm×3mm respectively for (x, y, z) and was resampled $3\times3\times3mm^3$ 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 to 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 of 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 32.7–77.6 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

Please cite this article in press as: Coroller TP et al. Radiomic phenotype features predict pathological response in non-small cell lung cancer. Radiother Oncol (2016), http://dx.doi.org/10.1016/j.radonc.2016.04.004