

UC Davis

UC Davis Previously Published Works

Title

Evaluating Which Dose-Function Metrics Are Most Critical for Functional-Guided Radiation Therapy

Permalink

<https://escholarship.org/uc/item/2nd0j6gb>

Journal

International Journal of Radiation Oncology • Biology • Physics, 99(1)

ISSN

0360-3016

Authors

Faught, Austin M
Yamamoto, Tokihiro
Castillo, Richard
[et al.](#)

Publication Date

2017-09-01

DOI

10.1016/j.ijrobp.2017.03.051

Peer reviewed



Published in final edited form as:

Int J Radiat Oncol Biol Phys. 2017 September 01; 99(1): 202–209. doi:10.1016/j.ijrobp.2017.03.051.

Evaluating which dose-function metrics are most critical for functional-guided radiotherapy with CT ventilation imaging

Austin M. Faught^{1,*}, Tokihiro Yamamoto², Richard Castillo³, Edward Castillo⁴, Jingjing Zhang⁴, Moyed Miften¹, and Yevgeniy Vinogradskiy¹

¹University of Colorado School of Medicine, *Aurora, Colorado*

²University of California Davis School of Medicine, *Sacramento, CA*

³University of Texas Medical Branch of Galveston, *League City, Texas*

⁴Beaumont Health System, *Royal Oak, Michigan*

Abstract

Purpose—4DCT-ventilation imaging is increasingly being used to calculate lung ventilation and implement functional-guided radiotherapy in clinical trials. There has been little exhaustive work evaluating which dose-function metrics should be used for treatment planning and plan evaluation. The purpose of our study was to evaluate which dose-function metrics best predict for radiation pneumonitis (RP).

Methods and Materials—Seventy lung cancer patients with 4DCT imaging and pneumonitis grading were used. Pre-treatment 4DCTs of each patient were used to calculate ventilation images. We evaluated 3 types of dose function metrics that combined that patient's 4DCT-ventilation image and treatment planning dose distribution: 1) structure-based approaches 2) image-based approaches using the dose-function histogram (DFH) and 3) non-linear weighting schemes. Log-likelihood methods were used to generate normal tissue complication probability (NTCP) models predicting grade 3+ pneumonitis for all dose-function schemes. The area under the curve (AUC) was used to assess the predictive power of the models. All techniques were compared to NTCP models based on traditional, total lung dose metrics.

Results—The most predictive models were structure-based approaches that focused on the volume of functional lung receiving 20Gy (AUC=0.70). Probability of grade 3+ RP of 20% and 10% correspond to V20Gy to the functional sub-volumes of 26.8% and 9.3%, respectively. Imaging-based analysis with the DFH and non-linear weighted ventilation values yielded AUCs of 0.66 and 0.67, respectively, when evaluating the percentage of functionality receiving 20Gy. All dose-function metrics outperformed the traditional dose metrics (mean lung dose, AUC=0.55).

*Corresponding Author: University of Colorado School of Medicine, 1665 Aurora Ct. Suite 1032, Campus Mail Stop R-706, Denver, CO 80045, Tel: (720) 848-5374, Fax: (720) 848-0481, austin.faught@ucdenver.edu.

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.

Conflict of Interest: This work was partially funded by grant R01CA200817 (AF, RC, EC, YV) and 1K01-CA-181292-01 (RC)

Conclusion—A full range of dose-function metrics and functional thresholds were examined. The calculated AUC values for the most predictive functional models occupied a narrow range (0.66–0.70) and all demonstrated notable improvements over AUC from traditional lung dose metrics (0.55). Identifying the combinations most predictive of grade 3+ RP provides valuable data to inform the functional-guided radiotherapy process.

Introduction

Dose limiting normal tissue tolerances used in lung radiation therapy are based on the assumption of homogenous underlying lung function (1–4). Recent literature (5,6) suggests that functional heterogeneity is present in a large portion of stage III lung cancers. The presence of functional defects and advances in imaging techniques has led to an interest in utilizing functional imaging modalities such as single photon emission computed tomography (SPECT) (7,8), hyperpolarized Helium or Xenon MRI (8,9), and more recently 4DCT ventilation imaging (10–12), to identify areas of high functionality for the purpose of preferential sparing during the treatment planning process.

4DCT ventilation uses phase or amplitude resolved CT images to calculate pulmonary ventilation on a voxel-by-voxel basis (13–17). Studies have detailed the methodology of 4DCT ventilation imaging (15,17), the validation of the technique (15,17,18), and its potential clinical uses as a functional imaging modality (12,19–22). Because 4DCT imaging is increasingly common amongst lung cancer patients, 4DCT ventilation imaging can be implemented with little to no added dosimetric or financial cost to the patient. There are several ongoing clinical trials investigating the use of CT ventilation imaging to preferentially spare areas of higher functioning lung (NCT02528942, NCT02773238, NCT02002052, NCT02308709, NCT02843568), a technique we refer to as functional-guided radiotherapy.

Studies have shown that dose-function metrics are more predictive of radiation toxicity than dose alone(19,23,24). However, there has been little exhaustive work evaluating which dose-function metrics should be used for treatment planning and plan evaluation for functional-guided radiotherapy. The purpose of our study was to evaluate which dose-function metrics are most critical in assessing toxicity. Specifically, we retrospectively evaluated various types of dose-function metrics and determined which metrics are most critical in predicting clinical radiation pneumonitis (RP).

Methods

Patient Population

A 70 patient cohort of non-small-cell lung cancer patients with 4DCTs and demonstrating heterogeneous lung functionality was used. Evaluating the heterogeneity of functionality is an important criterion in functional-guided radiotherapy because without areas of relative ventilation defects, the planning has no area to preferentially deposit dose in an effort to avoid higher functioning lung. A previous study details the heterogeneity screening process and its validation(25).

The 70 patient population consisted of 8 with Stage 1 disease, 9 with Stage 2 disease, 49 with Stage 3 disease, and 4 with Stage 4 disease. All patients were treated with standard fraction schemes (1.8Gy–2.0Gy per fraction). The clinical endpoint for analysis was grade 3 and higher (denoted grade 3+) RP using the Common Terminology Criteria for Adverse Events (version 3.0) scoring system. The endpoint was determined using clinical presentation and radiographic findings. Among the 70 patients used, 14/70 (20%) were scored as grade 3+.

Ventilation Image Calculation

Pre-treatment 4DCTs of each patient were used in the calculation of the ventilation images. Images were acquired at 120 kVp and 800 mAs/slice. A varied pitch based on the patient's breathing cycle was used. Segmentation of the peak inhale and peak exhale images was performed to exclude the trachea, main-stem bronchi, and pulmonary vasculature(17). Voxels from images of the two phases were linked using deformable image registration (DIR). A previous study found the accuracy of the deformation algorithm to be 1.25mm (26). All segmentation steps and the DIRs were visually inspected. A density changed based model, shown in Equation 1, was used to calculate ventilation based off a change in Hounsfield units (HU) between the two breathing phases.

$$\frac{V_{in} - V_{ex}}{V_{ex}} = 1000 \frac{HU_{in} - HU_{ex}}{HU_{ex}(1000 + HU_{in})} \quad (1)$$

In the equation, V_{in} and V_{ex} are the inhale and exhale volumes, respectively, and HU_{in} and HU_{ex} are the Hounsfield units of the individual lung voxels from the inhale and exhale phases, respectively. By assuming a linear combination of water-like material with HU value of 0 and air-like material with HU value of -1000 (27), a 3D ventilation map is calculated for each voxel within the lung (1,17).

Identification of Functional Lung

In the setting of functional-guided radiotherapy it will be important to evaluate both the dose and functional components of the treatment plan. In other words, metrics will be needed that assess the interplay of whether dose was deposited in functional lung.

To evaluate which dose-function metrics were most predictive for grade 3+ RP, we analyzed 3 different types of dose function metrics: 1) structure-based approaches 2) image-based approaches and 3) non-linear weighting schemes. The structures-based approaches create a 'functional' lung contour based on the ventilation image and calculate various dose metrics inside the functional contour. It should be underlined that there are 2 variables to consider in structure-based approaches: the threshold to use to define the functional contour and which standard dose metric (mean dose, V20Gy, etc.) to calculate inside the functional contour. Image-based approaches employ a dose-function histogram (DFH) concept(28,29) which is similar to the DVH except bins the function (defined by the 4DCT-ventilation) in each dose bin rather than binning the volume. Non-linear weighting schemes is an extension of the image-based concept where the ventilation values are weighted non-linearly and

subsequently used to calculate DFH metrics. The details of the calculation of each type of dose-function metric are provided below. Table 1 includes a summary of the dose-function metrics examined.

Structure-Based Evaluation of Dose-Function

Ventilation values from the 4DCT-ventilation image were binned into percentiles. The lung was then divided into regions of high-functioning and low-functioning lung based on a cutoff percentile value. The cutoff value of what defined functional lung was allowed to vary from 0% to 95%. When referring to a cutoff value, this means that regions of lungs with ventilation values above the cutoff value are treated as functional and regions of the lung with ventilation values below the cutoff value are treated as non-functional. As an example, a cutoff of the 90th percentile would treat the top 10% of all ventilation values as the functional region of the lung. Figure 1 displays how the functional contour changes in size when the threshold is changed from the 35th percentile to the 75th percentile.

As a second means of structure-based evaluation, areas of ventilation defect were identified as regions exhibiting a reduction in ventilation compared to a perfectly homogeneous lung. The calculation was performed by dividing the lung into superior-to-inferior thirds as reported in a previous assessment of 4DCT-ventilation imaging (6). This is a similar geometrical assessment of functionality used in nuclear medicine (30). Figure 2 shows an example of how the lungs are divided into thirds. Assuming lungs with perfectly homogenous function, each third would contain 16.7% (100% divided by six) of the total ventilation measured from the ventilation map. An area of functional defect was defined as regions in which the ventilation values deviated by a set percentage from the idealized homogenous value. The mathematical representation of the calculation is included in the Appendix for reference. NTCP models and analysis was performed for a range of values corresponding to ventilation threshold values 75% higher than the theoretical, homogeneous value to 95% lower than the homogeneous ventilation value. For both the percentile method and the geometrical-thirds method, we calculated mean dose, V5Gy, V10Gy, V20Gy, and V30Gy inside the functional contours.

Image-Based Evaluation of Dose-Function

Image based evaluation of dose-function metrics uses the dose-function histogram (DFH) for analysis. The DFH calculation tallies the dose by the physiologic function rather than the volume (as is done in a traditional DVH approach). The resulting histogram will communicate the fractional functional capacity receiving a particular dose. The DFH was used to calculate amount of functional lung receiving 5Gy (F5), 10Gy (F10), 20Gy (F20), and 30Gy (F30).

Non-Linear Weighting of Ventilation Values

Thus far the literature has focused on linear weighting of ventilation(19,28,32) and dose-function binning approaches(23). To the best of our knowledge, there has been no work done to explore whether non-linear weighting of function is a more appropriate way to assess toxicity. We explored a sigmoidal function to scale the ventilation values used for functional dose metric calculation. The idea is that function below and above certain thresholds may

not be physiologically meaningful (i.e. there are thresholds and saturation effects with regard to spatial lung function). Non-linear weighting of PET SUV has been explored as a means of dose painting to areas of tumor hypoxia (33). Raw ventilation values from the original ventilation calculation were weighted by the sigmoid function given in Equation 2.

$$V_{w,i} = \exp \left(-\exp \left(\frac{e * m * (V_i - h_c)}{V_{max}} \right) \right) \quad (2)$$

The ventilation in i th voxel is given by V_i and the maximum ventilation value in the map is V_{max} . The parameters m and h_c are sigmoid fitting parameters representing the slope and position of the sigmoid, respectively. NTCP models were calculated with values for the slope parameter, m , ranging from 0.5 to 16. The values for the position parameter, h_c , used in the calculations ranged from 0.05 to 1. By constraining the values of m and h_c , the weighting function maintained its sigmoid shape within the range of values encountered in a ventilation calculation (typically 0 to 1). Using the weighted ventilation values, a DFH was calculated for dose metric analysis.

Toxicity Modeling

Using the 70 patient cohort and an end-point of grade 3+ RP, normal tissue complication probability (NTCP) models were fit using log-likelihood methods. NTCP models were fit for each of the 3 types of dose-function metrics noted above. For reference, standard dose-volume metrics (MLD, V5Gy, V10Gy, V20Gy, V30Gy) were also evaluated and fit to NTCP models. The calculated NTCP models were then used in receiver operating characteristic (ROC) analysis with the corresponding area under the curve (AUC) being used to quantify the ability of each dose and dose-function metric to predict for toxicity. We report AUC values, logistic regression significance (p) values, and model fitting parameters.

While we focused on the modeling of grade 3+ RP for this study, for completeness, we also modelled grade 2+ RP. A summary of the results from the grade 2+ modelling is included in the Appendix.

Results

Structure-Based Modeling

A summary of the AUCs, p-values, and optimal threshold values for all modeling which treated functional lung as a sub-volume is presented in Table 2. For comparison, the results when modeling standard total lung metrics without regard to functionality are included. There was no difference in predictive power for calculating grade 3+ RP using structure-based analysis whether using percentile images of the ventilation map (AUC = 0.70) or changes from homogeneity (AUC = 0.70). Both methods of structure-based analysis showed superior predictive power with respect to traditional means of evaluating lung doses using mean lung dose (AUC = 0.58, p-value >0.05) and V20Gy (AUC = 0.52, p-value >0.05).

Another way to analyze structure-based approaches is to assess which threshold is optimal to define a 'functional' structure for a given DVH metric within the functional structure. A plot

of the AUC and p-value as a function of the threshold used for determining the functional lung sub-volume is shown in Figure 3. In the figure the threshold was calculated as a percentile of the ventilation value. The dose metric of analysis in this example was V20Gy. A clear trend towards higher AUC values (and reduced p-values) is observed as higher functioning lung is considered for evaluation compared to the total lung evaluated indiscriminately (0th percentile cutoff on the plot). It should also be noted that once a threshold above the 30th percentile is considered the variation in AUC is small (0.66 to 0.70). Included in the Appendix is a complete table of maximum AUCs, p-Values, and the optimum dose-function metric for each of the percentiles used as a threshold in structure-based analysis.

Presented in the results are the best performing (highest AUC) dose-function metrics for each example of structure-based modeling. In the Appendix a table of the highest AUC and associated p-Value for the other dose-function metrics (e.g. V5Gy, V10Gy, etc.) are presented.

The data presented in the current study aims to help establish guidelines for the implementation of functional-guided radiotherapy for prospective clinical trials. To that end, the optimal dose-function metric for each type of scheme is presented in Table 3 along with the dose constraints needed to achieve 10% and 20% probability of grade 3+ RP.

Image-based metrics: Dose Function Histogram and Non-linear Weighted Ventilation

The results for NTCP models generated from dose function histograms (DFH) and using nonlinear weighting of the ventilation values in the calculation of DFHs is presented in Table 2 along with the results from normal lung metric modeling for reference. Incorporation of functional status into the prediction of grade 3+ pneumonitis was superior to that of normal lung metrics. The sigmoidal weighting of the ventilation values showed a comparable predictive power for grade 3+ pneumonitis (AUC=0.67) with respect to the linear weighting used in traditional DFH analysis (AUC = 0.66). The model parameters used in the sigmoidal weighting of ventilation values that yielded the highest AUC while providing a significant fit to the data (p-Value <0.05), were found to be 0.2 for the position parameter h_c and 8 for the slope parameter m . A table of the model parameter results for grade 3+ and grade 2+ modeling are included in the Appendix. The most predictive dose-function metrics for both the DFH and non-linear weighting scheme were F20Gy. Included in the Appendix are the AUCs and associated p-values to the dose-function metrics that were not as predictive as F20Gy.

Discussion

Three salient points can be assessed from our study. The inclusion of functional information improves the prediction of radiation toxicity. Non-linear weighting of function does not demonstrate an appreciable improvement in toxicity prediction. While the dose-function metric of V20Gy/F20Gy consistently was shown to be most predictive of toxicities, the exact cutoff or delineation of functional/non-functional lung was less sensitive than broadly choosing to incorporate functional status. With regard to whether dose-function information improved toxicity, our NTCP model results for RP demonstrated improved predictive results

for dose-function metrics when compared to standard lung dose metrics. Both structure-based and image-based schemes all outperformed their standard dosimetric counterparts. As an example, the AUC results showed values of 0.58 for MLD and 0.73 for mean dose in a functional contour. Other studies have reported similar improvements in predicting RP(19,24) and radiation fibrosis(23) when incorporating functional information into predictive models. Our data along with previous work(19,23,24) demonstrates that regardless of which dose-function metric is used, incorporating functional information provides improved prediction of toxicity when compared to dose metrics alone. Retrospective demonstration of improved toxicity prediction with dose function metrics further supports the rationale for the ongoing prospective trials in thoracic functional-guided radiotherapy.

Our study was one of the first to examine a non-linear weighting scheme using a sigmoid mathematical function to weight function. The idea is that function below and above certain thresholds may not be physiologically meaningful. Our sigmoid modeling results showed a non-appreciable improvement in AUC compared to conventional DFH metrics (0.67 compared to 0.66) suggesting sigmoid-based weighting of function does not provide improved prediction of toxicity compared to structure-based or image-based metrics.

The central aim of this work was to provide data to help identify which metrics should be used for treatment planning and treatment evaluation in functional-guided radiotherapy. We sampled a wide dose-function sample space including 3 different types of dose function metrics (structure-based, image-based, and sigmoid weighting based), various thresholds to identify functional versus non-functional lung in structure-based approaches, and various dose metrics within each type of dose-function scheme. Our results revealed that for structure-based approaches the optimal metric was V20Gy (AUC=0.7) using the 84th percentile as the threshold to define functional lung for the percentile method and -35% reduction from homogenous for the percent deviation model. Similarly, the optimal metric was F20Gy for the DFH method and the sigmoid weighting method (AUC = 0.66 and AUC=0.67, respectively). The AUC values of 0.70 for structure-based methods slightly outperformed the AUC value of 0.67 and 0.66 for image-based methods using sigmoidal weighting and traditional DFH analysis, respectively. One important observation in evaluating a full suite of dose-function metrics is that the sensitivity between the treatment of functional lung (e.g. structure based analysis versus the DFH approach) is small. Even within a single means of analysis, the exact cutoff and metric is less important than the broad use of functional status. For example, in Figure 3 in the cutoff range between the 30th percentile and 95th percentile the AUC only varies from a 0.66 to 0.69. This difference is much smaller than the difference compared to normal lung dose metrics where the AUC was 0.52. We interpret these numbers as assigning value to performing functional-guided radiotherapy as a process much more than the small differences in the choice of functional/nonfunctional delineations.

Our study advances functional-guided radiotherapy in several critical ways. We offer the results of modeling a clinically pertinent toxicity end-point (RP) while other studies have focused on imaging-based changes(23). We evaluated a full suite of dose-function metrics including varying types of dose-function assessment, thresholds to define functional lung,

weighting schemes, and dose metrics. Previous studies have typically analyzed a few chosen dose-function metrics(32,34) but have not assessed them in a systematic or continuous way as is done in the current study. We also present an examination of non-linear weighting schemes that evaluated whether physiological function has upper and lower thresholds beyond which changes in ventilation values may not have physiological significance. The data presented in the current study aims to help establish guidelines for the implementation of functional-guided radiotherapy for prospective clinical trials. The results presented in Table 3 provide dosimetric goals for the reduction of toxicity probability from 20% to 10%. With the goal of functional-guided radiotherapy being the reduction of toxicities based on physiological information, the results in Table 3 are valuable, quantitative guidelines for this exciting new way to treat lung cancer patients.

In addressing the limitations to the study it is important to note that 4DCT ventilation imaging is a new technology under continual improvement. Specific efforts are being made to reduce the impact of imaging artifacts, uncertainties in the deformable image registration(35), means of normalizing the ventilation values(36), and the actual calculation of the ventilation values(37). This study also examines the initial ventilation maps calculated from pre-treatment 4DCT scans. There is evidence to suggest that as treatment progresses, changes in the treatment volume may result in changes in the ventilation maps corresponding to re-ventilation of previously occluded regions of the lung(38). Although radiation pneumonitis is a clinically pertinent end-point, the CTCAE grading system of pneumonitis has been noted to be subjective to both the patient and the clinician. Due to the subjectivity of the grading, we have chosen to also include toxicity modeling results for grade 2+ RP in the Appendix. Most of our optimal dose-function metrics were related to V20Gy. Because it is challenging to fully isolate the dose and function components, it is possible that our dose-function results were related to the most predictive dose metrics within our study cohort. Data set where other dose metrics (MLD for example(39)) are most significant in predicting toxicity may find other dose-function metrics to be relevant. Final treatment planning objectives will be based on the ongoing clinical trials in functional-guided radiotherapy; however, the retrospective data presented here provides a seminal baseline for the choice of dose-function metrics used in functional planning.

Conclusion

4DCT-ventilation imaging presents exciting opportunities in radiation therapy to account for functional lung status in creating individualized treatment plans for patients. With acquisition of 4DCT scans increasingly common in the current clinical care routine, ventilation images can be calculated at no extra cost to the patient and without delivering additional imaging dose. We used structure-based and image-based concepts of evaluating lung function and calculated NTCP models for a full suite of dose-function metrics to predict for radiation toxicity. The calculated AUC values for the most predictive functional models occupied a narrow range in values, 0.66–0.70, and all demonstrated notable improvements over traditional total lung dose metrics. The work presented in this paper shows a benefit towards using functional lung status to predict grade 3+ radiation pneumonitis and provides the metrics with highest predictive power for those toxicities. With clinical trials underway to prospectively apply functional-guided radiotherapy using

4DCT-ventilation imaging, this work provides valuable data for the evaluation and optimization of functional avoidance plans for lung cancer patients.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Appendix

Calculation of Percent Deviation from Homogeneity

Areas of functional defect were identified as regions with ventilation values below a calculated cutoff. The calculation of the cutoff value is based on a percent deviation from an ideal, homogeneous functional distribution within the lung. The theoretical ventilation value that corresponds to perfectly homogenous functionality is a mean of the average ventilation value in six lung regions, divided superior to inferior into thirds for both the left and right lung. The cutoff at which defects were identified is given in Equation A. 1.

$$V_T = \left(\frac{100}{6}\right) \left(\frac{100 - x}{100}\right) \sum_{i=1}^6 V_{ave,i} \quad (\text{A.1})$$

Where x is the percent reduction from homogeneous function, and $V_{ave,i}$ is the average ventilation value of the i th third. All voxels with ventilation values greater than V_T were classified as functional lung and included in the functional lung contour used to calculate functional dose metrics.

Mathematically the values for x can span +100% to -500%. Realistically the range is much more limited as it is unlikely that all of a patient's ventilation occurs within a highly localized region ($x = -500\%$) and in the limit that x approaches +100%, metrics would be derived for the entire lung and be equivalent to standard lung planning dose metrics. As an example, an input value of -50% would produce a ventilation threshold equal to 1.5 times theoretically homogeneous value. Likewise, an input of +50% means that the threshold value for identifying functional lung would be one half of the theoretically homogeneous value. In other words, in the latter case, all areas of the lung with ventilation values greater than 0.5 times the homogeneous ventilation value would be binned to the functional contour.

Modeling Results for Grade 2+ Radiation Pneumonitis

Table A.1

The predictive power (AUC) and significance of fit (p-Values) for NTCP models using different functional sub-volumes is presented for Grade 2+ radiation pneumonitis. For comparison, the normal means of evaluating lung dose from the total lung is reported for the examined patient cohort.

| Grade 2+ Pneumonitis | | | | |
|--------------------------|---------|-------------|---------|-----------------------------|
| Model | Metrics | Maximum AUC | p-Value | Threshold |
| Percentile | V20Gy | 0.73 | <0.01 | 86 th Percentile |
| | Mean | 0.73 | <0.01 | 69 th Percentile |
| Percent Deviation | V20Gy | 0.73 | <0.01 | -50 |
| | Mean | 0.73 | <0.01 | -15 |
| Total Lung | V20Gy | 0.55 | 0.36 | N/A |
| | Mean | 0.58 | 0.12 | N/A |

Table A.2

The predictive power (AUC) and significance of fit (p-Values) for NTCP models using different dose function histogram calculations is presented for Grade 2+ radiation pneumonitis. For comparison, the normal means of evaluating lung dose from the total lung is reported for the examined patient cohort.

| Grade 2+ Pneumonitis | | | |
|----------------------------|--------|------|---------|
| Model | Metric | AUC | p-Value |
| DFH | F10Gy | 0.70 | 0.02 |
| Sigmoidal Weighting | F20Gy | 0.74 | <0.01 |
| Total Lung | V20Gy | 0.55 | 0.36 |
| | Mean | 0.58 | 0.12 |

Calculating the Dose-Function of Interest for Different Structure-Based Thresholds

Table A.3

The dose metric providing the highest predictive power (AUC) for structure-based modeling is reported for different functional cutoff thresholds. The functional cutoff thresholds used were based on a percentile image of the ventilation values. For example, the 90th percentile would treat the top 10% of all ventilation values as the functional sub-volume used for analysis.

| Grade 3+ Pneumonitis | | | |
|-----------------------------|------------|----------------|---------------|
| Threshold | AUC | p-Value | Metric |
| Total Lung | 0.55 | 0.290 | Mean |
| 5th | 0.58 | 0.160 | V30Gy |
| 10th | 0.61 | 0.111 | V30Gy |
| 15th | 0.63 | 0.076 | V30Gy |
| 20th | 0.64 | 0.055 | V30Gy |
| 25th | 0.65 | 0.044 | V20Gy |
| 30th | 0.66 | 0.036 | V20Gy |
| 35th | 0.66 | 0.032 | V20Gy |
| 40th | 0.67 | 0.029 | V20Gy |
| 45th | 0.67 | 0.027 | V20Gy |
| 50th | 0.68 | 0.035 | V30Gy |
| 55th | 0.68 | 0.024 | V20Gy |
| 60th | 0.68 | 0.023 | V20Gy |
| 65th | 0.69 | 0.021 | V20Gy |
| 70th | 0.69 | 0.020 | V20Gy |
| 75th | 0.68 | 0.021 | V20Gy |
| 80th | 0.69 | 0.024 | V20Gy |
| 85th | 0.70 | 0.033 | V20Gy |
| 90th | 0.68 | 0.062 | V20Gy |
| 95th | 0.66 | 0.110 | V20Gy |

Model Parameters for Non-Linear Weighting

Table A.4

The model parameters for the sigmoidal weighting of ventilation values prior to DFH calculation are presented. These parameters provided the highest predictive power (as determined by AUC) while maintaining significance (p -Value < 0.05) with respect to predicting Grade 2+ and Grade 3+ radiation pneumonitis. Also presented are the dose-function metrics which provided the highest predictive power for the model. The parameter h_c corresponds to the relative position of the sigmoid and the parameter m represents the slope of the central region.

| Pneumonitis Grade | Dose-Function Metric | h_c | m | AUC | p-Value |
|-------------------|----------------------|-------|-----|------|---------|
| Grade 2+ | F20Gy | 0.3 | 8 | 0.74 | <0.01 |
| Grade 3+ | F20Gy | 0.2 | 8 | 0.67 | 0.03 |

Table A.5

The highest AUCs and associated p-values are presented for all the structure-based dose-function metrics examined in the study.

| Dose-Function Metric | | | | | | | | | | |
|----------------------|------|---------|-------|---------|-------|---------|-------|---------|-----------|---------|
| | V5Gy | | V10Gy | | V20Gy | | V30Gy | | Mean Dose | |
| Method | AUC | p-Value | AUC | p-Value | AUC | p-Value | AUC | p-Value | AUC | p-Value |
| % Deviation | 0.55 | 0.78 | 0.60 | 0.25 | 0.70 | 0.02 | 0.68 | 0.02 | 0.66 | 0.04 |
| Percentile | 0.55 | 0.61 | 0.59 | 0.35 | 0.70 | 0.03 | 0.68 | 0.03 | 0.66 | 0.06 |

Table A.6

The highest AUCs and associated p-values are presented for all the image-based dose-function metrics examined in the study.

| Dose-Function Metric | | | | | | | | | |
|----------------------|------|---------|-------|---------|-------|---------|-------|---------|--|
| | F5Gy | | F10Gy | | F20Gy | | F30Gy | | |
| Method | AUC | p-Value | AUC | p-Value | AUC | p-Value | AUC | p-Value | |
| DFH | 0.48 | 0.92 | 0.54 | 0.55 | 0.66 | 0.04 | 0.66 | 0.04 | |
| Non-Linear Weighting | 0.53 | 0.75 | 0.57 | 0.39 | 0.67 | 0.03 | 0.65 | 0.04 | |

References

1. Yaremko BP, Guerrero TM, Noyola-Martinez J, et al. Reduction of normal lung irradiation in locally advanced non-small-cell lung cancer patients, using ventilation images for functional avoidance. *International Journal of Radiation Oncology Biology Physics*. 2007; 68:562–571.
2. Martel MK, Ten Haken RK, Hazuka MB, et al. Estimation of tumor control probability model parameters from 3-d dose distributions of non-small cell lung cancer patients. *Lung Cancer*. 1999; 24:31–37. [PubMed: 10403692]
3. Graham MV, Purdy JA, Emami B, et al. Clinical dose-volume histogram analysis for pneumonitis after 3d treatment for non-small cell lung cancer (nsccl). *International Journal of Radiation Oncology Biology Physics*. 1999; 45:323–329.

4. Kwa SLS, Lebesque JV, Theuvs JCM, et al. Radiation pneumonitis as a function of mean lung dose: An analysis of pooled data of 540 patients. *International Journal of Radiation Oncology* Biology* Physics*. 1998; 42:1–9.
5. Redacted
6. Redacted
7. Lavrenkov K, Christian JA, Partridge M, et al. A potential to reduce pulmonary toxicity: The use of perfusion spect with imrt for functional lung avoidance in radiotherapy of non-small cell lung cancer. *Radiother Oncol*. 2007; 83:156–62. [PubMed: 17493699]
8. Bates EL, Bragg CM, Wild JM, et al. Functional image-based radiotherapy planning for non-small cell lung cancer: A simulation study. *Radiother Oncol*. 2009; 93:32–6. [PubMed: 19552978]
9. Hodge CW, Tomé WA, Fain SB, et al. On the use of hyperpolarized helium mri for conformal avoidance lung radiotherapy. *Med Dosim*. 2010; 35:297–303. [PubMed: 19944585]
10. Yamamoto T, Kabus S, Bal M, et al. The first patient treatment of computed tomography ventilation functional image-guided radiotherapy for lung cancer. *Radiotherapy and Oncology*. 2015
11. Yamamoto T, Kabus S, von Berg J, et al. Impact of four-dimensional ct-derived pulmonary ventilation images on radiotherapy treatment planning for lung cancer. *International Journal of Radiation Oncology Biology Physics*. 2009; 75:S443–S443.
12. Yamamoto T, Kabus S, von Berg J, et al. Impact of four-dimensional computed tomography pulmonary ventilation imaging-based functional avoidance for lung cancer radiotherapy. *International Journal of Radiation Oncology Biology Physics*. 2011; 79:279–288.
13. Kabus S, von Berg J, Yamamoto T, et al. Lung ventilation estimation based on 4d-ct imaging. :73–83.
14. Guerrero T, Sanders K, Castillo E, et al. Dynamic ventilation imaging from four-dimensional computed tomography. *Phys Med Biol*. 2006; 51:777–91. [PubMed: 16467578]
15. Reinhardt JM, Ding K, Cao K, et al. Registration-based estimates of local lung tissue expansion compared to xenon ct measures of specific ventilation. *Medical Image Analysis*. 2008; 12:752–763. [PubMed: 18501665]
16. Christensen GE, Song JH, Lu W, et al. Tracking lung tissue motion and expansion/compression with inverse consistent image registration and spirometry. *Medical physics*. 2007; 34:2155–2163. [PubMed: 17654918]
17. Castillo R, Castillo E, Martinez J, et al. Ventilation from four-dimensional computed tomography: Density versus jacobian methods. *Physics in Medicine and Biology*. 2010; 55:4661–4685. [PubMed: 20671351]
18. Du KF, Bayouth JE, Cao KL, et al. Reproducibility of registration-based measures of lung tissue expansion. *Medical Physics*. 2012; 39:1595–1608. [PubMed: 22380392]
19. Redacted
20. Yamamoto T, Kabus S, Klinder T, et al. Investigation of four-dimensional computed tomography-based pulmonary ventilation imaging in patients with emphysematous lung regions. *Physics in Medicine and Biology*. 2011; 56:2279–2298. [PubMed: 21411868]
21. Bayouth J, Du K, Christensen G, et al. Establishing a relationship between radiosensitivity of lung tissue and ventilation. *International Journal of Radiation Oncology* Biology* Physics*. 2012; 84:S31–S32.
22. Ding K, Bayouth JE, Buatti JM, et al. 4dct-based measurement of changes in pulmonary function following a course of radiation therapy. *Medical Physics*. 2010; 37:1261–1272. [PubMed: 20384264]
23. Lan F, Jeudy J, Senan S, et al. Should regional ventilation function be considered during radiation treatment planning to prevent radiation-induced complications? *Medical Physics*. 2016; 43:5072. [PubMed: 27587037]
24. Seppenwoolde Y, De Jaeger K, Boersma LJ, et al. Regional differences in lung radiosensitivity after radiotherapy for non-small-cell lung cancer. *International Journal of Radiation Oncology Biology P hysics*. 2004; 60:748–758.
25. Redacted

26. Castillo E, Castillo R, Martinez J, et al. Four-dimensional deformable image registration using trajectory modeling. *Physics in Medicine and Biology*. 2010; 55:305–327. [PubMed: 20009196]
27. Simon BA. Non-invasive imaging of regional lung function using x-ray computed tomography. *Journal of Clinical Monitoring and Computing*. 2000; 16:433–442. [PubMed: 12580227]
28. Marks LB, Sherouse GW, Munley MT, et al. Incorporation of functional status into dose-volume analysis. *Medical Physics*. 1999; 26:196–199. [PubMed: 10076973]
29. Miften MM, Das SK, Su M, et al. Incorporation of functional imaging data in the evaluation of dose distributions using the generalized concept of equivalent uniform dose. *Physics in Medicine and Biology*. 2004; 49:1711–1721. [PubMed: 15152926]
30. Parker JA, Coleman RE, Grady E, et al. Snm practice guideline for lung scintigraphy 4.0. *Journal of nuclear medicine technology*. 2012; 40:57–65. [PubMed: 22282651]
31. Marks LB, Spencer DP, Bentel GC, et al. The utility of spect lung perfusion scans in minimizing and assessing the physiological consequences of thoracic irradiation. *International Journal of Radiation Oncology Biology Physics*. 1993; 26:659–668.
32. Seppenwoolde Y, Engelsman M, De Jaeger K, et al. Optimizing radiation treatment plans for lung cancer using lung perfusion information. *Radiotherapy and oncology*. 2002; 63:165–177. [PubMed: 12063006]
33. Bowen SR, Flynn RT, Bentzen SM, et al. On the sensitivity of imrt dose optimization to the mathematical form of a biological imaging-based prescription function. *Physics in medicine and biology*. 2009; 54:1483. [PubMed: 19218733]
34. Bradley JD, Paulus R, Komaki R, et al. A randomized phase iii comparison of standard-dose (60 gy) versus high-dose (74 gy) conformal chemoradiotherapy with or without cetuximab for stage iii non-small cell lung cancer: Results on radiation dose in rtog 0617. *J Clin Oncol*. 2013; 31:7501.
35. Pioped I. Value of the ventilation/perfusion scan in acute pulmonary embolism. Results of the prospective investigation of pulmonary embolism diagnosis (pioped). *Jama*. 1990; 263:2753. [PubMed: 2332918]
36. Du K, Reinhardt JM, Christensen GE, et al. Respiratory effort correction strategies to improve the reproducibility of lung expansion measurements. *Medical physics*. 2013; 40:123504. [PubMed: 24320544]
37. Palma DA, de Koste JvS, Verbakel WFAR, et al. Lung density changes after stereotactic radiotherapy: A quantitative analysis in 50 patients. *International Journal of Radiation Oncology Biology Physics*. 2011; 81
38. Osoba D, Rodrigues G, Myles J, et al. Interpreting the significance of changes in health-related quality-of-life scores. *Journal of Clinical Oncology*. 1998; 16:139–144. [PubMed: 9440735]
39. Martel MK, Ten Haken RK, Hazuka MB, et al. Dose-volume histogram and 3-d treatment planning evaluation of patients with pneumonitis. *International Journal of Radiation Oncology* Biology* Physics*. 1994; 28:575–581.

Summary

We performed a retrospective analysis of 70 lung cancer patients with 4DCT images to assess which dose-function metrics are most predictive of clinical radiation pneumonitis and should be used in functional-guided radiotherapy. Normal tissue complication probability models were developed based on standard lung dose metrics and four additional functional schemes that considered lung function as determined from 4DCT ventilation imaging. Results provide valuable data in the guidance of prospective clinical trials in functional-guided radiotherapy.

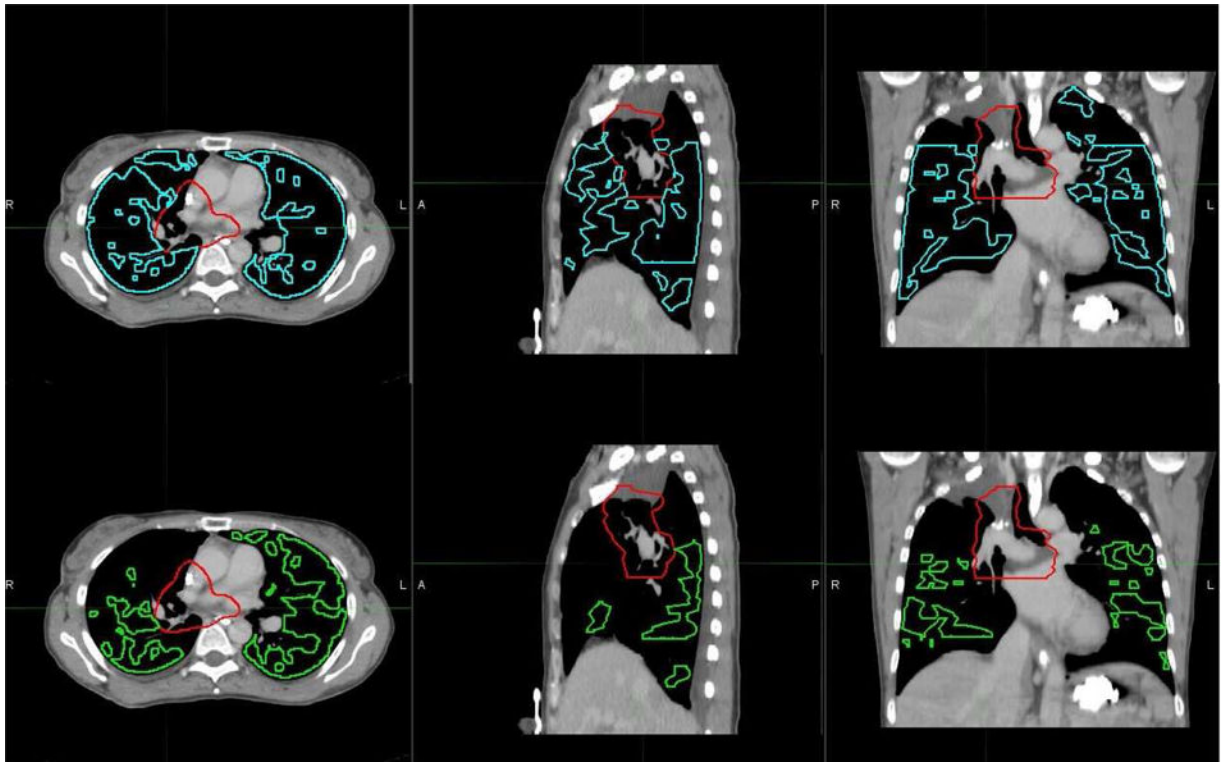


Figure 1.

Axial (left), sagittal (center), and coronal (right) planes of the functional sub-volume contour are presented for threshold values of 35th percentile (top) and 75th percentile (bottom). The red contour represents the target and blue and green represent the respective functional contours. The choice of threshold in delineating between functional and non-functional lung affects the size of the functional contour used during treatment planning optimization.

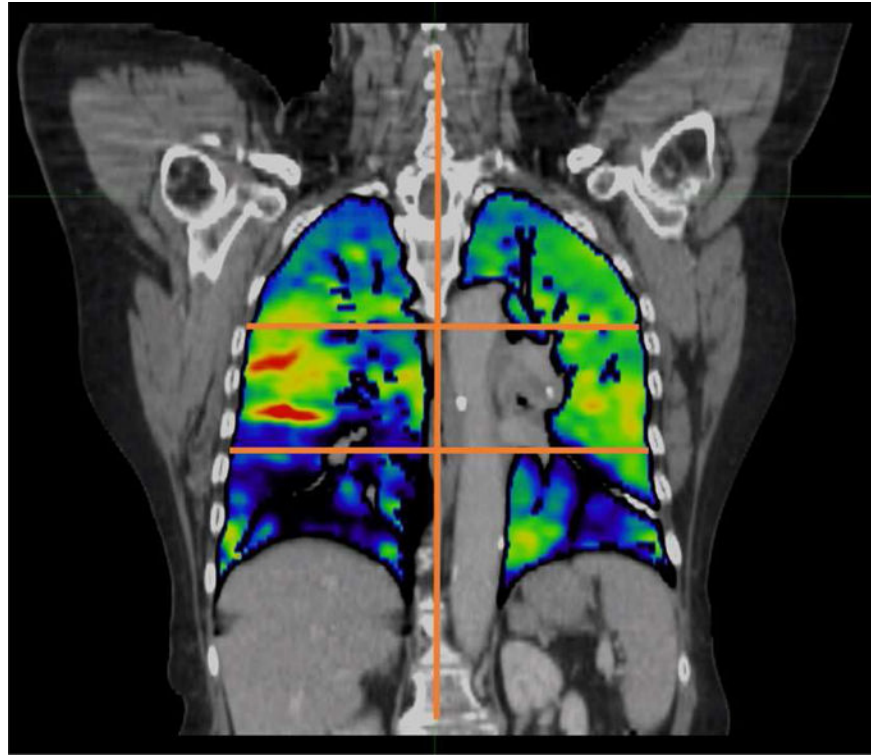


Figure 2.

A coronal image of the ventilation calculation registered to the CT is presented. The right and left lung are divided into thirds as shown by the overlaid lines on the image. Brighter colors (reds, yellows, and greens) correspond to areas of higher ventilation while the darker colors (blue and purple) correspond to regions of lower ventilation. A clear ventilation defect is present in the right lower lobe of this patient.

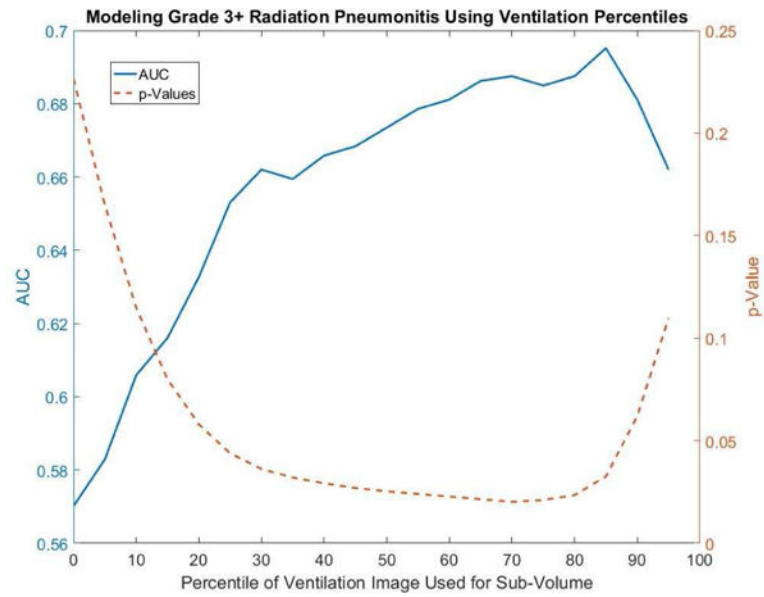


Figure 3.

A plot of the area under the curve (solid blue) and the p-Value (dashed orange) as a function of the cutoff in determining functional lung sub-volumes demonstrates the trend toward higher predictive power in modeling grade 3+ pneumonitis when evaluating areas of higher functioning lung. The metric of evaluation is the percent of functional lung receiving 20Gy.

A summary of the dose-function metrics examined for each method investigated. VXGy indicates the percent volume receiving XGy and FXGy indicates the percent functionality receiving XGy.

Table 1

| Method | Dose-Function Metrics | | | | | |
|-----------------------------|-----------------------|-------|-------|-------|-------|-----------|
| Structure-Based | V5Gy | V10Gy | V20Gy | V30Gy | V30Gy | Mean Dose |
| Imaging-Based | F5Gy | F10Gy | F20Gy | F20Gy | F30Gy | |
| Non-Linear Weighting | F5Gy | F10Gy | F20Gy | F20Gy | F30Gy | |

Table 2

The predictive power (AUC) and significance of fit (p-Values) for NTCP models using different structure-based and imaging-based analysis is presented for grade 3+ radiation pneumonitis. For comparison, the normal means of evaluating lung dose from the total lung is reported for the examined patient cohort.

| Structure-Based Modeling of Grade 3+ Pneumonitis | | | | |
|---|----------------|--------------------|----------------|-----------------------------|
| Model | Metrics | Maximum AUC | p-Value | Threshold |
| Structure-Based Percentile | V20Gy | 0.70 | 0.03 | 84 th Percentile |
| Structure-Based Percent Deviation | V20Gy | 0.70 | 0.02 | -35 |
| Imaging-Based DFH | F20Gy | 0.66 | 0.04 | N/A |
| Imaging-Based Sigmoidal Weighting | F20Gy | 0.67 | 0.03 | N/A |
| Total Lung | V20Gy | 0.52 | 0.49 | N/A |
| | Mean | 0.55 | 0.29 | N/A |

Table 3

The dose reductions required to reduce the normal tissue complication probability of Grade 3+ radiation pneumonitis is presented for the two functional sub-volume metrics shown to be most predictive as well as for the normal, total lung dose metrics.

| Model | Sub-Volume Cutoff | Metric | NTCP = 10% | NTCP = 20% |
|----------------------------|-----------------------------|--------|------------|------------|
| % Deviation | -35% | V20Gy | 13.3% | 27.3% |
| Percentile | 84 th Percentile | V20Gy | 9.3% | 26.8% |
| DFH | N/A | F20Gy | 20.4% | 31.0% |
| Sigmoidal Weighting | N/A | F20Gy | 17.2% | 30.2% |
| Total Lung | N/A | V20Gy | 4.5% | 34.0% |
| | N/A | Mean | 10.2 Gy | 20.7 Gy |