HHS Public Access

Author manuscript

Int J Radiat Oncol Biol Phys. Author manuscript; available in PMC 2018 September 01.

Published in final edited form as:

Int J Radiat Oncol Biol Phys. 2017 September 01; 99(1): 80–89. doi:10.1016/j.ijrobp.2017.05.002.

Variabilities of MRI-, CT- and PET-CT-based tumor and lymph node delineations for lung cancer radiotherapy planning

Kishor Karki[1,‡], **Siddharth Saraiya**[1,4,‡], **Geoffrey D. Hugo**[1], **Nitai Mukhopadhyay**[2], **Nuzhat Jan**[1], **Jessica Schuster**[1], **Matthew Schutzer**[1], **Lester Fahrner**[3], **Robert Groves**[3], **Kathryn M. Olsen**[5], **John C. Ford**[6], and **Elisabeth Weiss**[1,†]

[1]Department of Radiation Oncology, Virginia Commonwealth University, Richmond, VA, USA

[2]Department of Biostatistics, Virginia Commonwealth University, Richmond, VA, USA

[3]Department of Radiology, Virginia Commonwealth University, Richmond, VA, USA

[4]Department of Radiation Oncology, University of Toledo, OH, USA

[5]Department of Radiology, University of Colorado, Denver, CO, USA

[6]Department of Radiation Oncology, University of Miami, Miami, FL, USA

Abstract

Purpose—Target delineation in lung cancer radiotherapy using CT and/or PET-CT is affected by large variability. MRI has excellent soft tissue visualization and better spatial resolution than PET-CT. The main purpose of this study is to analyze delineation variability for lung cancer using MRI.

Methods and materials—Seven physicians delineated the tumor volumes of ten patients for the following scenarios: (1) CT only; (2) PET-CT fusion images registered to CT ("clinical standard"); and (3) post-contrast T1-weighted MRI registered with diffusion-weighted MRI. To compute interobserver variability, the median surface was generated from all observers' contours and used as the reference surface. A physician labeled the interface types (tumor to lung, atelectasis (collapsed lung), hilum, mediastinum, or chest wall) on the median surface. Contoured volumes and bidirectional local distances (BLDs) between individual observers' contours and the reference contour were analyzed.

Results—CT- and MRI-based tumor volumes normalized relative to PET-CT-based volumes were31.62±0.76 (mean±SD) and 1.38±0.44, respectively. Volume differences between the imaging modalities were not significant. Between observers, the mean normalized volumes per patient averaged over all patients varied significantly by a factor of 1.6 (MRI) and 2.0 (CT and PET-CT) (p=4.10×10^{-5} – 3.82×10^{-9}). The tumor-atelectasis interface had a significantly higher variability

†Corresponding author: Elisabeth Weiss, 401 College Street, PO Box 980058, Richmond, VA 23298, USA. elisabeth.weiss@vcuhealth.org; Tel: (804) 828-7232; Fax: (804) 828-6042.
‡These authors contributed equally and are co-first authors.

Disclosures: See COI forms. There are no conflicts of interest.

than other interfaces for all modalities combined (p=0.0006). The interfaces with the smallest uncertainties were tumor-lung (on CT) and tumor-mediastinum (on PET-CT and MRI).

Conclusions—While MRI-based contouring showed overall larger variability than PET-CT, contouring variability depended on the interface type and was not significantly different between modalities despite of the limited observer experience with MRI. Multimodality imaging and combining different imaging characteristics might be the best approach to define the tumor volume most accurately.

I. INTRODUCTION

Target delineation in lung cancer radiotherapy has, in general, large variability [1,2]. Reliable delineation of the target volumes is required for accurate radiation dose delivery [3,4]. PET-CT has been shown to improve target definition by improving the accuracy of determining tumor volume and localization [5,6]. Therefore, PET-CT is routinely used for lung cancer target delineation for radiotherapy planning. However, even with PET-CT target delineation remains highly variable [7–9]. Compared to PET-CT, MRI has good contrast between normal and cancerous tissues with higher spatial resolution [10]. MRI has been shown to be superior to CT for identifying tumor extensions to chest wall, mediastinum, vascular and neural structures, such as brachial plexus [10–14]. In conjunction with diffusion-weighted sequences (DW-MRI), MRI has been shown to be comparable or even superior to PET-CT in primary tumor and nodal staging [15–18]. MRI has rarely been evaluated for lung cancer target definition in radiotherapy. In a recent study, DW-MRI-based gross tumor volumes (GTVs) were reported to be as reliable as PET-CT-based GTVs [19].

Several methods have been used to evaluate interobserver variability in target volume delineation [4,20]. Typical approaches include assessment of centroid shifts or volumetric overlap measures, e.g., conformity index and dice similarity coefficient [3,20,21]. To understand the association of contour uncertainty with anatomical information presented on multimodality imaging, in the present study voxel-wise contouring variability on 3D volume surfaces was analyzed. The standard deviation (SD) of the distances of the multiple observers' contour surfaces from the reference contour surface, defined as the median contour, was assessed in relation to tumor-tissue interface types [22,23] using the bidirectional local distance (BLD) measure [24]. The purpose of this study is to investigate interobserver delineation variability for GTVs of primary lung tumors (PTs) and associated pathologic lymph nodes (LNs) and for their interfaces with surrounding tissues using MRI, and to compare the results to CT alone- and PET-CT-based delineations.

II. METHODS AND MATERIALS

II.A. Patients

Multimodality imaging of ten patients with non-small cell lung cancer (NSCLC) obtained on an IRB-approved protocol (NCT xxx) was used for this study (Table 1). Patients underwent MRI, CT and 2-deoxy-2-(^{18}F)fluoro-D-glucose (FDG) PET-CT before conventionally fractionated radiation therapy. The average time between CT and PET-CT was 17 days (range: 2–39 days), between CT and MRI 33 days (9–52), between PET-CT and MRI 19

days (10–36). Nine of the ten patients had primary tumors in lung and associated lymph nodes; one patient had pathologic LN involvement only.

II.B. Imaging modalities

Diagnostic CT: Diagnostic CT image sets with intravenous contrast were used for delineation in the CT only and PET-CT sessions, slice thickness was 3.00 mm and pixel size ~0.58×0.58 mm^2.

PET-CT: Fusion images of the attenuation-corrected PET images with the corresponding CT images of PET-CT scans were used for delineation after rigid registration to diagnostic CT images. The PET images had a slice thickness of ~4.25 mm and in-plane resolution of ~3.91×3.91 mm^2.

MRI: Images from a 1.5 T MRI scanner included DW-MRI and post Gd-DTPA T_1-weighted ultrafast gradient echo volume interpolated breath-hold examination (VIBE) images. Apparent diffusion coefficient (ADC) maps were obtained from the DW-MRI images of eight b-values = 0–1000 s/mm^2. VIBE images had a slice thickness of ~2.00 mm and an in-plane pixel size of ~1.56×1.56 mm^2. DW-MRI images had pixel size of ~1.98×1.98 mm^2, slice thickness/gap = 6 mm/1.2 mm. Each frame of a DW-MRI scan had 7 slices. One or more frames were used to cover the whole tumor volume.

II.C. Target delineations

Five radiation oncologists and two chest radiologists were asked to independently delineate GTV of PT (GTV_PT) and GTV of LN (GTV_LN) using MIM Software (MIM Software Inc., Cleveland, OH, USA). During contouring, sagittal and coronal views were displayed. To avoid potential confounding errors for identifying and contouring different lesions, the physicians were informed on the locations of GTV_PT and/or GTV_LN by marking the location in a representative axial slice of the images. Seven lymph nodes were contoured in four patients. Only well demarcated, clearly visualized paratracheal and hilar nodes were selected for contouring. The following delineation sessions were created (see example in Figures 1 and S1):

1. CT only: The GTV_PT contour was drawn on the lung/mediastinal window on diagnostic CT images, depending on tumor location. GTV_LN structures were contoured in the mediastinal window.
2. PET-CT: Users were instructed to first use a gradient-based auto-segmentation tool (PET-Edge, MIM) [25]. Users then manually edited the volumes on the CT images.
3. MRI: Users were provided VIBE images rigidly registered with DW-MRI images and ADC maps (Fig. 1). DW-MRI images of the highest b-value (b = 1000 s/mm^2) or, in case of very low signal to noise ratio, the second highest b-value (b = 800 s/mm^2) were used as additional information for the delineation on VIBE images. Since most of the users had little or no experience with tumor delineation on VIBE, a presentation was prepared in collaboration with a specialized chest radiologist showing contouring examples for tumor and

atelectasis on VIBE and DWI scans for various clinical scenarios including contour recommendations.

II.D. Analysis

The median contour was generated from all observers' contours for each patient and modality separately and was selected as the reference contour. The median contours were generated for each GTV following the method of Steenbakkers et al. [23]. Briefly, a voxel was labeled as part of the median GTV if at least 50% of the observers (4/7 observers) labeled this voxel as GTV. All voxels labeled as median GTV were then combined into the median GTV contour, as shown in Fig. 1. A physician labeled the interface types (GTV_PT or GTV_LN to lung, atelectasis (collapsed lung), hilum, mediastinum, airways, blood vessels and/or chest wall) on the median surface as applicable to the respective GTV_PT and GTV_LN structures (Fig. 2).

The mask images of the 3D contours including individual observers' contours, median contours and contours identifying the interfaces were processed using MATLAB. A marching cubes algorithm was used to generate a surface contour mesh from the 3D volume of the binary mask images with a threshold value of 0.5 [26]. The BLD measures for the observers' contours from the reference median contour were obtained for the different interfaces and the whole 3D surface using the methods described by Kim et al [24]. BLD was employed in the presence of complex lung tumor contours to reduce errors of the minimum distance (MD) analysis and other conventional distance-based methods by taking into account the bidirectional characteristics with forward and backward directions.

BLD has been shown to be more appropriate than other distance measures, especially around concave or folding regions of either reference or individual observer's contours or both, where MD can be too short; normal distance either cannot be established or is erroneously very long; and ComGrad (method using the combined gradient of the signed distance transforms produced from the reference and observer's contours) cannot find corresponding points (see [24] for more details). The procedure for obtaining BLD is given below:

i. The forward minimum distance, $d_{min}(P, t)$, from a point P on a reference contour surface r to an individual observer's contour surface t is determined (see Figure S2):

$$\text{Forward minimum distance} = d_{min}(P, t) = \min\{d(P, T_i)\},$$

where points T_i (i = 1, 2, …., N (total number of voxels on t)) are on contour t, and d_{min} represents the minimum distance. This step is the same as obtaining MD.

ii. All possible backward minimum distances from the points T_j to r such that $d_{min}(T_j, r) = d(T_j, P)$, are identified.

iii. BLD at P is identified as the greatest among these forward and backward minimum distances obtained in steps (i) and (ii) (in Figure S2, T_3P is the BLD measure at P). Note that $d_{min}(T_j, r) = d(T_j, P) \geq d_{min}(P, t)$.

Backward minimum distances from a number of points T_j on contour t to contour r can be obtained at P. If no backward minimum distances to r can be identified at P, then the forward minimum distance, i.e. MD from P is established as the BLD measure. This process is repeated for all voxels on the reference contour surface for each observer's contour surface. The standard deviation of the distance from the observers' contours to a single point on the reference contour was calculated, which we termed the local SD. Root mean square of the local SD values (RMS SD) was obtained as an overall SD representing the variation in distance to all points as a measure of target delineation variability.

Plots of the local SD on the surface mesh (SDSM) were made for visualization of the regional variations of local SDs. For improved visualization (in Fig. 1) 3D Gaussian smoothing of the contour surface with $\sigma = 1.0$ mm was applied, and a mean filter was then applied to the computed local SD values on the 3D smoothed surface. We tried to keep σ small to preserve the original variations of 3D GTV_PT or GTV_LN contours [24, 27]. Local SD and RMS SD were compared across the three imaging modalities and across the observers.

II. E. Statistical analysis

Comparisons of GTV_PT and GTV_LN volume differences across modalities and observers were performed using a repeated measures ANOVA model with modality and observers as main effects. To investigate the interobserver variability or systematic error made by the observers, the observers' contour volumes relative to the corresponding mean volume were compared for all modalities for both GTV_PT and GTV_LN using single factor ANOVA. Overall uncertainties for both GTV_PT and GTV_LN were computed for all interfaces and whole surfaces for all three modalities.

A random intercept clustered linear mixed effects model was used to model the regional RMSE in different imaging modalities and interface data using SAS v 9.4 and R v 3.2.1.

III. RESULTS

III.A. Target volume variations between the modalities

The ranges of mean GTV_PT and GTV_LN volumes in PET-CT imaging modality were 3.6–284.9 and 5.2–30.4 cc, respectively. Because of a large range of volumes across the patients for each of the three modalities, mean volumes were normalized to the PET-CT volume per patient and averaged over all patients to compare volumes across the imaging modalities. CT- and MRI-based GTV_PT normalized volumes were 1.62±0.76 (mean±SD) and 1.38± 0.44 times larger than PET-CT volumes, respectively, whereas GTV_LN volumes were 1.02±0.21 and 1.30± 0.42 times larger, respectively (Figure 3). The CT-based absolute volume was larger than the PET-CT volume with an estimated difference of 77.28 cc (95% confidence interval (CI): −48.27, 202.83, p=0.21). The MRI-based volume was also larger than the PET-CT volume with an estimated difference of 35.84 cc (95% CI: −89.71, 161.4, p=0.56).

III.B. Interobserver variability for whole GTV_PT and GTV_LN contours

The individual observers' normalized volumes relative to the mean volume per patient averaged over all patients (Fig. 4) were significantly different for all modalities for GTV_PT (p=4.08×10^{-8}, 3.82×10^{-9} and 4.10×10^{-5}, single factor ANOVA for CT, PET-CT and MRI, respectively) and associated pathologic GTV_LN (p=0.00057, 0.09057 and 0.00024, respectively). Mean observers' normalized volumes per patient averaged over all patients for GTV_PT and GTV_LN varied by a factor of 1.6 and 2.2 for MRI and a factor of 2.0 and 1.7, respectively, for both CT and PET-CT, indicating better observer agreement with MRI for GTV_PT, but less for GTV_LN.

III.C. Delineation variability for interfaces

Using BLD analyses, RMS SD values of the interfaces are reported for both GTV_PT and GTV_LN (Fig. 5). The largest mean uncertainties of the whole surface for GTV_PT were observed in CT (2.96 vs 2.06 and 2.77 mm for PET-CT and MRI, respectively) and for GTV_LN in MRI (2.22 vs 1.90 and 2.01 mm for CT and PET-CT, respectively). The GTV_PT -atelectasis interface, which was present in 7/10 patients, had a significantly higher variability than other interfaces for all modalities combined (p=0.0006). Also, for the three modalities individually, the GTV_PT -atelectasis interface had the largest mean overall uncertainty (4.26, 2.53 and 4.13 mm for CT, PET-CT and MRI, respectively). The interfaces having the smallest uncertainty for CT, PET-CT and MRI were GTV_PT -lung (1.92 mm), GTV_PT -chest wall (1.61 mm) and GTV_PT -mediastinum (2.25 mm), respectively. Differences for specific interfaces were not significantly different between modalities.

IV. DISCUSSION

Contouring variability can exceed other geometric errors [28] with potentially detrimental effects on tumor control [29] and toxicity [30]. Hence, reducing contour variability has clinical implications for improving radiotherapy, particularly for highly conformal and adaptive treatment approaches. Van de Steene et al. [30] described several factors for large inter-observer variation such as methodological problems, inability to differentiate between pathologic or normal structures and tumor, as well as lack of knowledge. Our study aimed to identify tumor boundaries and differentiate between tumor and surrounding structures for various imaging modalities with particular focus on MRI.

As documented in the present study, observer variation in lung cancer delineation using CT only is a major source of geometric uncertainty [1,30–32]. Several studies [5,6,33–39] showed that use of FDG-PET has a large impact on the delineated tumor volume and localization. Bradley et al. [5] demonstrated that PET changes GTV contours in approximately two thirds of the cases compared to CT only. Steenbakkers et al's analysis [23] demonstrated that observer variation decreased with the use of PET-CT. Despite these advantages, the relatively low resolution of PET (compared to CT), improper registration, and motion blurring have been recognized as limitations for contouring. MRI provides excellent soft tissue contrast, and higher image resolution than PET, but has seen limited use in target definition for lung cancer radiotherapy [19]. Hence, radiation oncologists have generally less experience with MRI than CT for target definition. The growing use of MRI in

radiation oncology departments for treatment planning, image-guided adaptive radiotherapy, and treatment response assessment motivated us to compare the contouring variability for lung cancer on MRI to routine methods.

We selected the particular MRI sequences for several reasons. VIBE images provide anatomic details based on the different T1-relaxation values of water protons in tissues, can be acquired with sufficiently high spatial resolution, and motion artifact is minimized through breath hold acquisition. Analogous to PET-CT, incorporating functional imaging such as DW-MRI along with morphological VIBE images may assist in differentiating benign and cancerous tissue during delineation (e.g. within atelectatic lung). Variations in the apparent diffusion of water molecules due to different tissue cellularity provide contrast between tumor and the surrounding tissues in DW-MRI images and ADC maps [19].

Despite the time intervals described above for obtaining the different imaging modalities with CT acquired first, the overall tumor volume on CT only was the largest, whereas contouring on PET-CT resulted in the smallest average GTV_PT volume. Similarly, contour variability for the whole tumor surface was largest for CT only, indicating that most of the volume differences are caused by the inability to determine the tumor boundary. However, four of the seven observers on average drew larger contours on PET-CT than CT only, potentially due to the blurring effect on PET-CT, and three observers were found to have their smallest average contours on MRI.

The ability to differentiate tumor from surrounding structures depends on tumor-specific factors such as tumor shape, size, location and its surrounding structures. Our results show that the variability of all GTV_PT interfaces was smallest on PET-CT except for the interface with lung where CT contouring had the smallest variability. Functional information on PET-CT allowed good identification of collapsed lung, however, contouring variability was significantly larger on the tumor-atelectasis interface for all modalities compared to other interfaces. This is in agreement with the analysis by Steenbakkers et al [23]. MRI is currently used in routine diagnostic imaging to examine mediastinal, vascular and nerve invasion. In our study, MRI did not result in any significant reductions in delineation uncertainty, although some benefit was seen at the interfaces between GTV_PT and mediastinum and GTV_LN and vessels. Overall, MRI did not appear helpful for GTV_LN interface detection in this small cohort. However, GTV_LN volumes and interfaces were small. Even small contour differences therefore resulted in large errors relative to the median contour.

Our results indicate that some observers consistently drew larger/smaller volumes on all imaging modalities. The causes of this might be multifactorial. Observer training had been shown to reduce contouring variability [40, 41]. In this study, for each patient, graphical information on the location of the primary tumor and lymph nodes of interest was provided to minimize contouring errors due to uncertainty about tumor position and to focus the study on the detection of uncertainties arising from differences in imaging modalities only.

Technical factors such as slice thickness, resolution and signal-to-noise ratio (SNR) can also affect the differentiation of tumor from its surrounding structures. While VIBE offers

excellent soft tissue contrast, the low resolution of DW-MRI images and motion artifacts in the acquisition of VIBE and DW-MRI images may have potentially impacted contouring variability. Optimization of these factors may result in greater contrast-to-noise ratio (CNR) and may reduce delineation variability. Special attention should be given to artifacts/distortions such as motion, susceptibility and chemical shift artifacts associated with different types of MRI, particularly during delineation of chest tumors.

V. CONCLUSIONS

While MRI-based contouring showed overall larger variability than PET-CT, contouring variability depended on the interface type and was not significantly different between modalities despite of the limited observer experience with MRI. The GTV_PT-atelectasis interface had the largest uncertainty for all imaging modalities. With more training and experience, anatomic MRI combined with functional MRI sequences might gain clinical importance for tumor delineation close to critical interfaces, such as for differentiating cancerous processes in atelectatic lung from inflammation where PET-CT has known limitations. Although MRI showed less variability for the GTV_LN-vessel interface, larger contour variations and volumes for GTV_LN in general indicate that MRI can currently not be recommended for GTV_LN contouring with the MRI sequences used in this study. Multimodality imaging and combining different imaging characteristics might be the best approach to define the GTV most accurately.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by a VCU Massey Cancer Center pilot project and NIH grant P30CA016059.

References

1. Giraud P, Elles S, Helfre S, De Rycke Y, Servois V, Carette MF, Alzieu C, Bondiau PY, Dubray B, Touboul E, et al. Conformal radiotherapy for lung cancer: different delineation of the gross tumor volume (GTV) by radiologists and radiation oncologists. Radiother Oncol. 2002; 62(1):27–36. [PubMed: 11830310]
2. Vorwerk H, Beckmann G, Bremer M, Degen M, Dietl B, Fietkau R, Gsanger T, Hermann RM, Alfred Herrmann MK, Holler U, et al. The delineation of target volumes for radiotherapy of lung cancer patients. Radiother Oncol. 2009; 91(3):455–460. [PubMed: 19339069]
3. Peulen H, Belderbos J, Guckenberger M, Hope A, Grills I, van Herk M, Sonke JJ. Target delineation variability and corresponding margins of peripheral early stage NSCLC treated with stereotactic body radiotherapy. Radiother Oncol. 2015; 114(3):361–366. [PubMed: 25770872]
4. Van der Put RW, Raaymakers BW, Kerkhof EM, van Vulpen M, Lagendijk JJ. A novel method for comparing 3D target volume delineations in radiotherapy. Phys Med Biol. 2008; 53(8):2149–2159. [PubMed: 18379021]
5. Bradley J, Thorstad WL, Mutic S, Miller TR, Dehdashti F, Siegel BA, Bosch W, Bertrand RJ. Impact of FDG-PET on radiation therapy volume delineation in non-small-cell lung cancer. Int J Radiat Oncol Biol Phys. 2004; 59(1):78–86. [PubMed: 15093902]
6. Nestle U, Hellwig D, Schmidt S, Licht N, Walter K, Ukena D, Rube C, Baumann M, Kirsch CM. 2-Deoxy-2-[18F]fluoro-D-glucose positron emission tomography in target volume definition for

radiotherapy of patients with non-small-cell lung cancer. Mol Imaging Biol. 2002; 4(3):257–263. [PubMed: 14537131]

7. Apostolova I, Wiemker R, Paulus T, Kabus S, Dreilich T, van den Hoff J, Plotkin M, Mester J, Brenner W, Buchert R, et al. Combined correction of recovery effect and motion blur for SUV quantification of solitary pulmonary nodules in FDG PET/CT. Eur Radiol. 2010; 20(8):1868–1877. [PubMed: 20306084]
8. Erdi YE, Nehmeh SA, Pan T, Pevsner A, Rosenzweig KE, Mageras G, Yorke ED, Schoder H, Hsiao W, Squire OD, et al. The CT motion quantitation of lung lesions and its impact on PET-measured SUVs. J Nucl Med. 2004; 45(8):1287–1292. [PubMed: 15299050]
9. Goerres GW, Kamel E, Seifert B, Burger C, Buck A, Hany TF, Von Schulthess GK. Accuracy of image coregistration of pulmonary lesions in patients with non-small cell lung cancer using an integrated PET/CT system. J Nucl Med. 2002; 43(11):1469–1475. [PubMed: 12411550]
10. Khalil A, Bouhela T, Carette MF. Contribution of MRI in lung cancer staging. Jbr-Btr. 2013; 96(3): 132–141. [PubMed: 23971169]
11. Beale R, Slater R, Hennington M, Keagy B. Pancoast tumor: use of MRI for tumor staging. South Med J. 1992; 85(12):1260–1263. [PubMed: 1470977]
12. Heelan RT, Demas BE, Caravelli JF, Martini N, Bains MS, McCormack PM, Burt M, Panicek DM, Mitzner A. Superior sulcus tumors: CT and MR imaging. Radiology. 1989; 170(3 Pt 1):637–641. [PubMed: 2916014]
13. Rapoport S, Blair DN, McCarthy SM, Desser TS, Hammers LW, Sostman HD. Brachial plexus: correlation of MR imaging with CT and pathologic findings. Radiology. 1988; 167(1):161–165. [PubMed: 3347719]
14. Takasugi JE, Rapoport S, Shaw C. Superior sulcus tumors: the role of imaging. J Thorac Imaging. 1989; 4(1):41–48. [PubMed: 2643714]
15. Henzler T, Schmid-Bindert G, Schoenberg SO, Fink C. Diffusion and perfusion MRI of the lung and mediastinum. Eur J Radiol. 2010; 76(3):329–336. [PubMed: 20627435]
16. Metcalfe P, Liney GP, Holloway L, Walker A, Barton M, Delaney GP, Vinod S, Tome W. The potential for an enhanced role for MRI in radiation-therapy treatment planning. Technol Cancer Res Treat. 2013; 12(5):429–446. [PubMed: 23617289]
17. Mori T, Nomori H, Ikeda K, Kawanaka K, Shiraishi S, Katahira K, Yamashita Y. Diffusion-weighted magnetic resonance imaging for diagnosing malignant pulmonary nodules/masses: comparison with positron emission tomography. J Thorac Oncol. 2008; 3(4):358–364. [PubMed: 18379353]
18. Usuda K, Zhao XT, Sagawa M, Matoba M, Kuginuki Y, Taniguchi M, Ueda Y, Sakuma T. Diffusion-weighted imaging is superior to positron emission tomography in the detection and nodal assessment of lung cancers. Ann Thorac Surg. 2011; 91(6):1689–1695. [PubMed: 21619964]
19. Fleckenstein J, Jelden M, Kremp S, Jagoda P, Stroeder J, Khrelsh F, Ezziddin S, Buecker A, Rübe C, Schneider G. The Impact of Diffusion-Weighted MRI on the Definition of Gross Tumor Volume in Radiotherapy of Non-Small-Cell Lung Cancer. PLOS ONE. Sep 9.2016
20. Fotina I, Lutgendorf-Caucig C, Stock M, Potter R, Georg D. Critical discussion of evaluation parameters for inter-observer variability in target definition for radiation therapy. Strahlenther Onkol. 2012; 188(2):160–167. [PubMed: 22281878]
21. Hanna GG, Hounsell AR, O'Sullivan JM. Geometrical analysis of radiotherapy target volume delineation: a systematic review of reported comparison methods. Clin Oncol (R Coll Radiol). 2010; 22(7):515–525. [PubMed: 20554168]
22. Fitton I, Steenbakkers RJ, Gilhuijs K, Duppen JC, Nowak PJ, van Herk M, Rasch CR. Impact of anatomical location on value of CT-PET co-registration for delineation of lung tumors. Int J Radiat Oncol Biol Phys. 2008; 70(5):1403–1407. [PubMed: 17980511]
23. Steenbakkers RJ, Duppen JC, Fitton I, Deurloo KE, Zijp LJ, Comans EF, Uitterhoeve AL, Rodrigus PT, Kramer GW, Bussink J, et al. Reduction of observer variation using matched CT-PET for lung cancer delineation: a three-dimensional analysis. Int J Radiat Oncol Biol Phys. 2006; 64(2):435–448. [PubMed: 16198064]

24. Kim HS, Park SB, Lo SS, Monroe JI, Sohn JW. Bidirectional local distance measure for comparing segmentations. Med Phys. 2012; 39(11):6779–6790. [PubMed: 23127072]
25. Werner-Wasik M, Nelson AD, Choi W, Arai Y, Faulhaber PF, Kang P, Almeida FD, Xiao Y, Ohri N, Brockway KD, et al. What is the best way to contour lung tumors on PET scans? Multiobserver validation of a gradient-based method using a NSCLC digital PET phantom. Int J Radiat Oncol Biol Phys. 2012; 82(3):1164–1171. [PubMed: 21531085]
26. Lorensen, WE., Cline, HE. Marching cubes: A high resolution 3D surface construction algorithm. 14th Annual Conference on Computer Graphics and Interactive Techniques; 1987; Anaheim: ACM, New York; 1987. p. 163-169.
27. xxx
28. Gwynne S, Spezi E, Sebag-Montefiore D, Mukherjee S, Miles E, Conibear J, Staffurth J. Improving radiotherapy quality assurance in clinical trials: assessment of target volume delineation of the pre-accrual benchmark case. Br J Radiol. 2013; 86(1024):20120398. [PubMed: 23392188]
29. Hamilton CS, Ebert MA. Volumetric uncertainty in radiotherapy. Clin Oncol (R Coll Radiol). 2005; 17(6):456–464. [PubMed: 16149290]
30. Van de Steene J, Linthout N, de Mey J, Vinh-Hung V, Claassens C, Noppen M, Bel A, Storme G. Definition of gross tumor volume in lung cancer: inter-observer variability. Radiother Oncol. 2002; 62(1):37–49. [PubMed: 11830311]
31. Bowden P, Fisher R, Mac Manus M, Wirth A, Duchesne G, Millward M, McKenzie A, Andrews J, Ball D. Measurement of lung tumor volumes using three-dimensional computer planning software. Int J Radiat Oncol Biol Phys. 2002; 53(3):566–573. [PubMed: 12062598]
32. Senan S, van Sornsen de Koste J, Samson M, Tankink H, Jansen P, Nowak PJ, Krol AD, Schmitz P, Lagerwaard FJ. Evaluation of a target contouring protocol for 3D conformal radiotherapy in non-small cell lung cancer. Radiother Oncol. 1999; 53(3):247–255. [PubMed: 10660205]
33. Ciernik IF, Dizendorf E, Baumert BG, Reiner B, Burger C, Davis JB, Lutolf UM, Steinert HC, Von Schulthess GK. Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study. Int J Radiat Oncol Biol Phys. 2003; 57(3):853–863. [PubMed: 14529793]
34. Dizendorf EV, Baumert BG, von Schulthess GK, Lutolf UM, Steinert HC. Impact of whole-body 18F-FDG PET on staging and managing patients for radiation therapy. J Nucl Med. 2003; 44(1): 24–29. [PubMed: 12515872]
35. Erdi YE, Rosenzweig K, Erdi AK, Macapinlac HA, Hu YC, Braban LE, Humm JL, Squire OD, Chui CS, Larson SM, et al. Radiotherapy treatment planning for patients with non-small cell lung cancer using positron emission tomography (PET). Radiother Oncol. 2002; 62(1):51–60. [PubMed: 11830312]
36. Giraud P, Grahek D, Montravers F, Carette MF, Deniaud-Alexandre E, Julia F, Rosenwald JC, Cosset JM, Talbot JN, Housset M, et al. CT and (18)F-deoxyglucose (FDG) image fusion for optimization of conformal radiotherapy of lung cancers. Int J Radiat Oncol Biol Phys. 2001; 49(5): 1249–1257. [PubMed: 11286831]
37. Mah K, Caldwell CB, Ung YC, Danjoux CE, Balogh JM, Ganguli SN, Ehrlich LE, Tirona R. The impact of (18)FDG-PET on target and critical organs in CT-based treatment planning of patients with poorly defined non-small-cell lung carcinoma: a prospective study. Int J Radiat Oncol Biol Phys. 2002; 52(2):339–350. [PubMed: 11872279]
38. van Der Wel A, Nijsten S, Hochstenbag M, Lamers R, Boersma L, Wanders R, Lutgens L, Zimny M, Bentzen SM, Wouters B, et al. Increased therapeutic ratio by 18FDG-PET CT planning in patients with clinical CT stage N2-N3M0 non-small-cell lung cancer: a modeling study. Int J Radiat Oncol Biol Phys. 2005; 61(3):649–655. [PubMed: 15708242]
39. Vanuytsel LJ, Vansteenkiste JF, Stroobants SG, De Leyn PR, De Wever W, Verbeken EK, Gatti GG, Huyskens DP, Kutcher GJ. The impact of (18)F-fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET) lymph node staging on the radiation treatment volumes in patients with non-small cell lung cancer. Radiother Oncol. 2000; 55(3):317–324. [PubMed: 10869746]
40. Dewas S, Bibault JE, Blanchard P, Vautravers-Dewas C, Pointreau Y, Denis F, Brauner M, Giraud P. Delineation in thoracic oncology: a prospective study of the effect of training on contour variability and dosimetric consequences. Radiat Oncol. 2011; 6:118. [PubMed: 21929770]

41. xxx

Summary

Target delineation in lung cancer radiotherapy is affected by large variability. MRI has excellent soft tissue visualization and was therefore compared to CT only and PET-CT for lung cancer target definition in a multiple observer delineation study. Contouring variability on MRI was usually less than CT, but greater than PET-CT without significant differences. Contouring variability was also observed dependent on the interface type with the largest variability identified for the tumor-atelectasis interface.

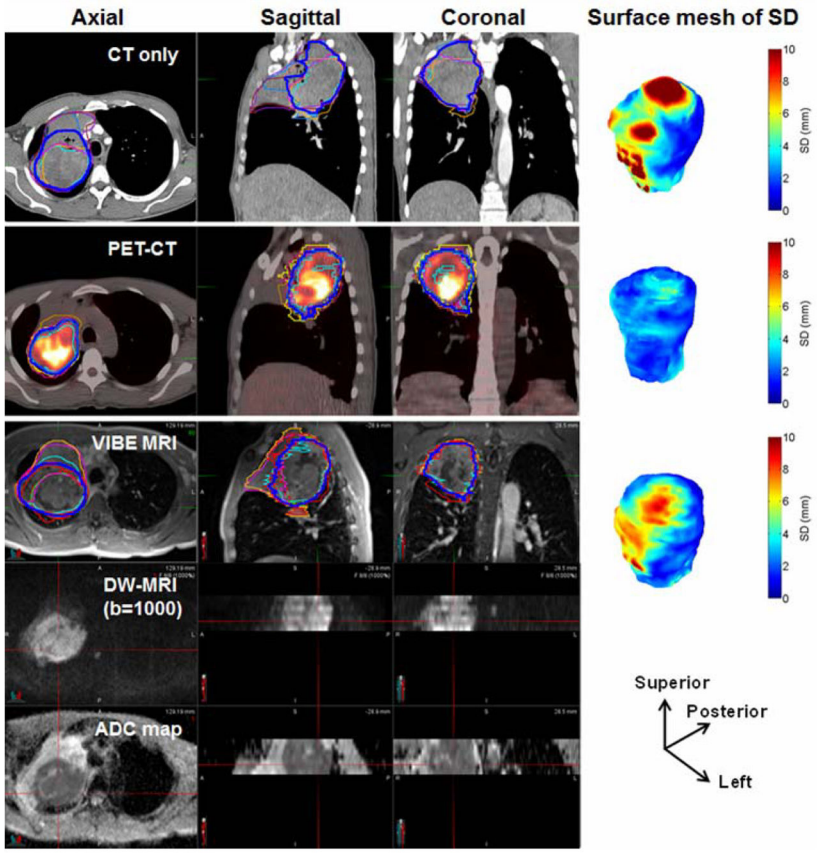

Figure 1.

An example of a large primary tumor GTV delineated by seven observers (different thin colored lines) and median contour (dark blue colored line) in different imaging modalities (for reference see also Figure S1 without contours). DW-MRI images and the corresponding ADC maps of a frame covered ~5.0 cm thickness. The surface mesh of local SD (SDSM) for the three modalities in an oblique view with the directions is shown. In total, three frames of DW-MRI and ADC maps were used to cover this tumor. The contoured volumes were 346, 282 and 328 cm^3 on CT, PET-CT and MRI, respectively.

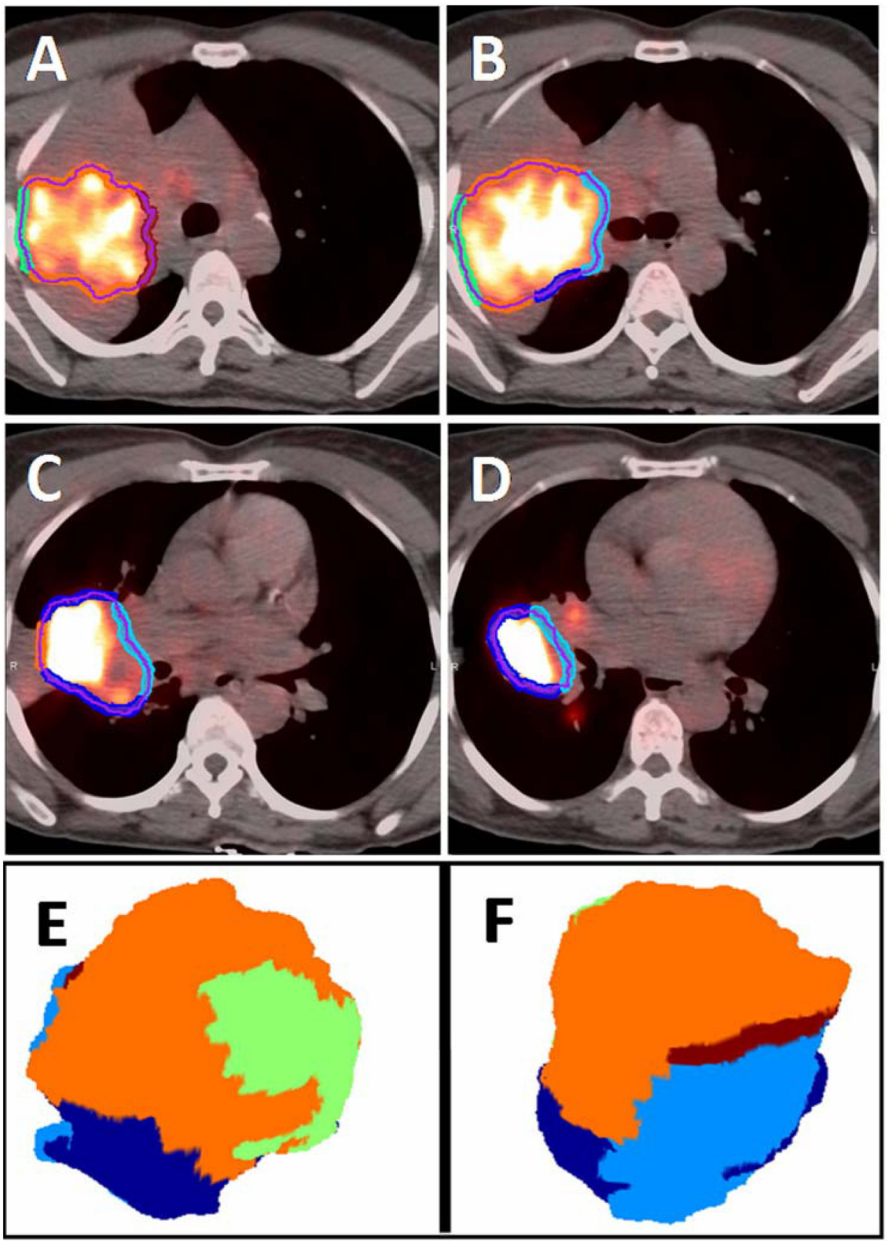

Figure 2.

The interfaces of primary tumor GTV with lung (dark-blue), hilum (light blue), atelectasis (orange), chest-wall (green) and mediastinum (dark red) on median contour surfaces of PET-CT of an example patient in in four representative axial slices (A–D), and in a 3D oblique (E) and its flipped views (F).

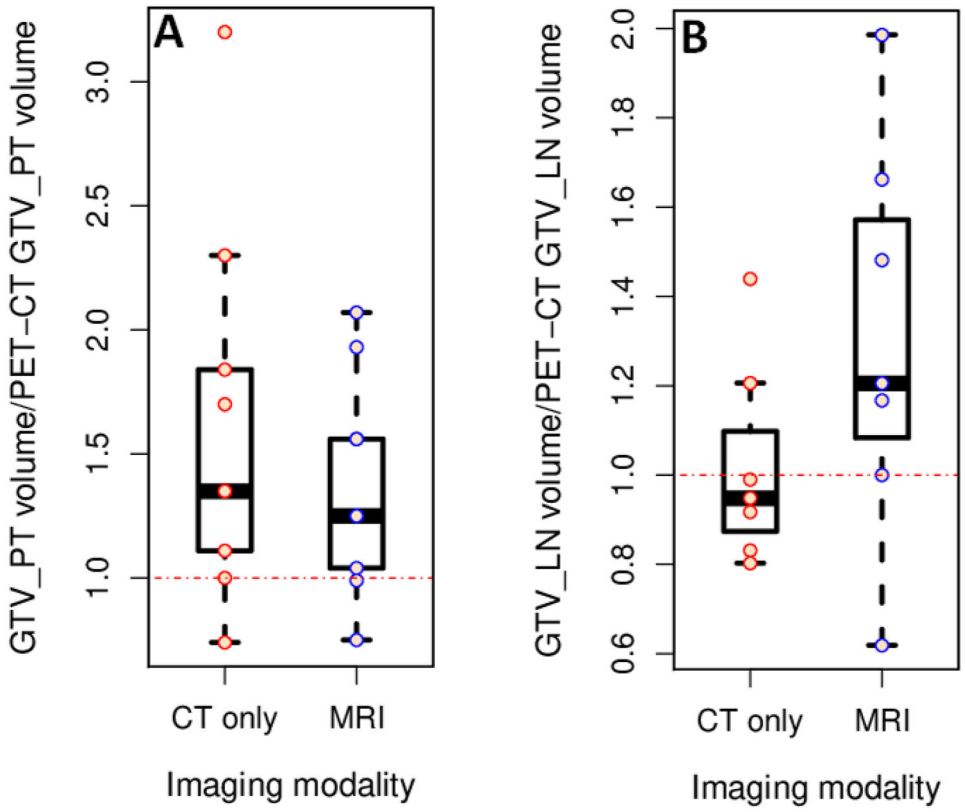

Figure 3.

CT and MRI based primary tumor (A) and lymph node (B) gross tumor volumes (GTVs) normalized by dividing average volume per patient by the corresponding PET-CT based volume. Boxes represent the inter-quartile range (IQR) extending from first quartile (Q1) to third quartile (Q3), respectively. Whiskers extend from Q1 and Q3 to end points that are typically defined as the most extreme data points within Q1-1.5*IQR and Q3 +1.5*IQR, respectively. The small open circles represent the individual data points. Fig. 3(A) shows that the normalized CT-based GTV_PT volume was generally larger than the MRI-based volume, both volumes were larger than tumor volumes on PET-CT.

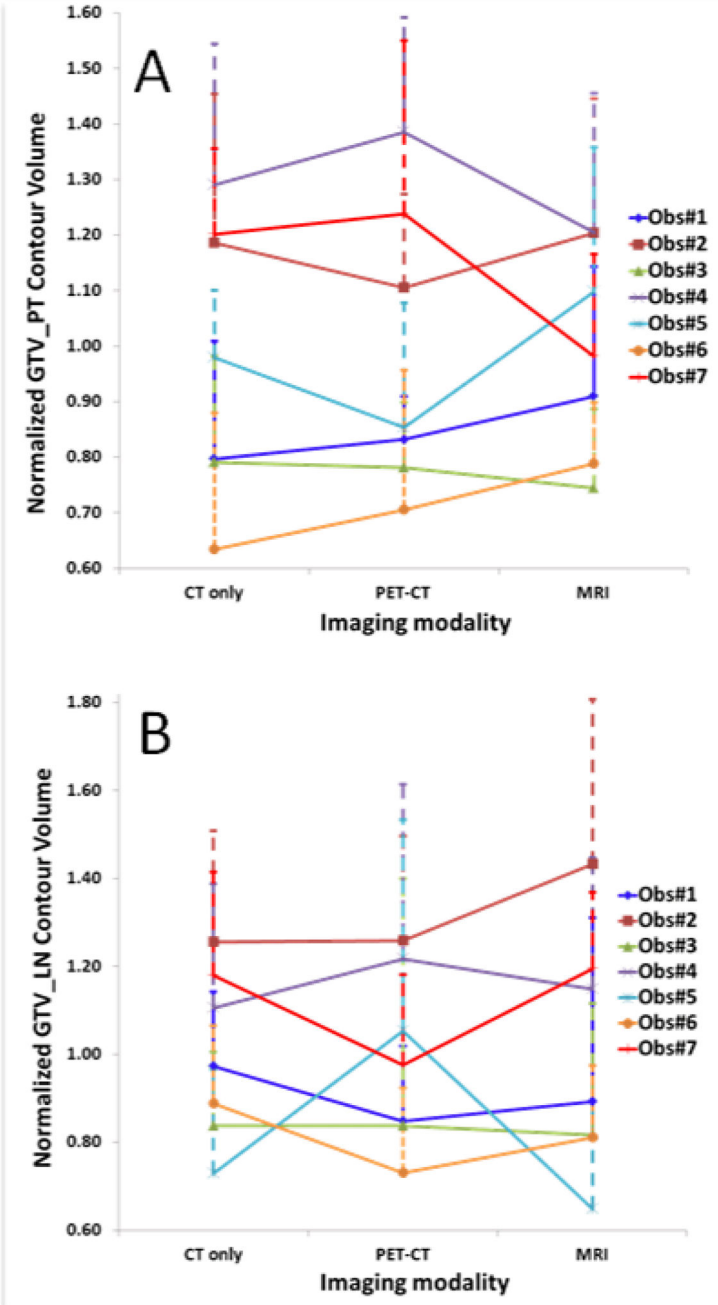

Figure 4.

Individual observers' contour volumes (mean±SD) normalized relative to the corresponding mean volume per patient and averaged over all patients for the gross tumor volumes (GTV) of primary tumor (PT) (A) and lymph node (LN) (B) per imaging modality. Some observers consistently delineated comparatively large or small volumes. Volumes delineated by radiologists (Observer#3 and Observer#5) were relatively small compared to radiation oncologists.

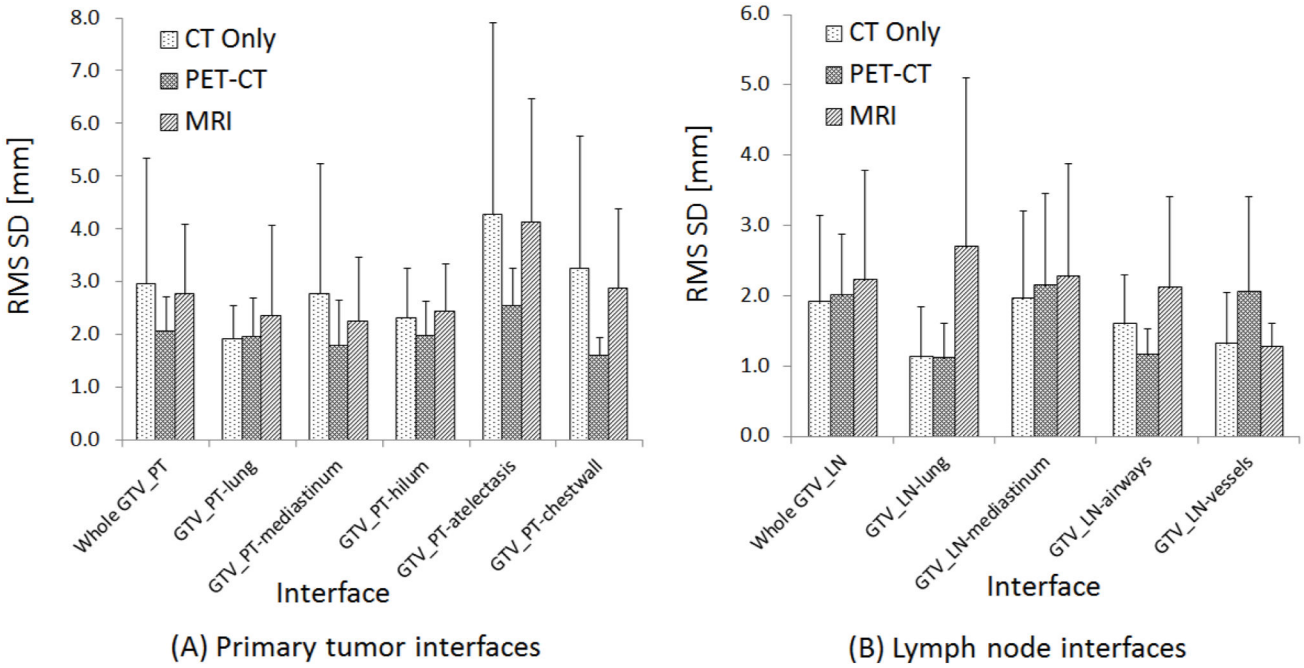

Figure 5.

Overall uncertainty (RMS SD) of BLD values (mean±SD) for the gross tumor volumes (GTVs) of primary tumor (PT) interfaces (A) and lymph node (LN) interfaces (B). Uncertainties were largest for the PT- atelectasis interface for all modalities.

Table 1

Patient Characteristics

Patient	Age (y)	Race	Sex	Histology	Stage	Tumor Location	Mean GTV_PT Volume on PET-CT (in ml)	SUV_{max} of GTV_PT
A	57	AA	F	SCC	IIIB	RUL	231.4	28.8
B	61	W	M	SCC	IIIB	LUL	50.4	24.3
C	62	W	F	Adeno	IIIA	RUL	95.7	11.8
D	57	W	M	SCC	IIIA	RUL	3.6	18.1
E	60	W	F	Adeno	IIIB	LUL	117.4	32.9
F	50	W	M	SCC	IIIA	RUL	107.8	11.1
G	56	AA	M	SCC	IIIB	RUL	284.9	63
H	51	AA	M	SCC	IIB	RUL/RML	58.7	66
I	64	AA	M	Adeno	IIIB	LLL	78.2	17.1
J	53	AA	M	Undiff. Ca	IIIA	LN only	-	41

AA, African-American; Adeno, adenocarcinoma; Ca, carcinoma; F, female; GTV_PT, gross target volume of primary tumor; LLL, left-lower lobe; LN, lymph nodes; LUL, left-upper lobe; M, male; RML, right-middle lobe; RUL, right-upper lobe; SCC, squamous cell carcinoma; Undiff., undifferentiated; W, white. All patients except C, D and J had atelectasis or collapsed lung.

Int J Radiat Oncol Biol Phys. Author manuscript; available in PMC 2018 September 01.