

CT-based quantitative evaluation of radiation-induced lung fibrosis: a study of interobserver and intraobserver variations

Jaesung Heo, MD, Oyeon Cho, MD, O Kyu Noh, MD, PhD, Young-Taek Oh, MD, PhD,
Mison Chun, MD, PhD, Mi-Hwa Kim, MS, Hae-Jin Park, PhD

Department of Radiation Oncology, Ajou University School of Medicine, Suwon, Korea

Purpose: The degree of radiation-induced lung fibrosis (RILF) can be measured quantitatively by fibrosis volume (VF) on chest computed tomography (CT) scan. The purpose of this study was to investigate the interobserver and intraobserver variability in CT-based measurement of VF.

Materials and Methods: We selected 10 non-small cell lung cancer patients developed with RILF after postoperative radiation therapy (PORT) and delineated VF on the follow-up chest CT scanned at more than 6 months after radiotherapy. Three radiation oncologists independently delineated VF to investigate the interobserver variability. Three times of delineation of VF was performed by two radiation oncologists for the analysis of intraobserver variability. We analysed the concordance index (CI) and inter/intra-class correlation coefficient (ICC).

Results: The median CI was 0.61 (range, 0.44 to 0.68) for interobserver variability and the median CIs for intraobserver variability were 0.69 (range, 0.65 to 0.79) and 0.61 (range, 0.55 to 0.65) by two observers. The ICC for interobserver variability was 0.974 ($p < 0.001$) and ICCs for intraobserver variability were 0.996 ($p < 0.001$) and 0.991 ($p < 0.001$), respectively.

Conclusion: CT-based measurement of VF with patients who received PORT was a highly consistent and reproducible quantitative method between and within observers.

Keywords: Radiation-induced lung fibrosis, Fibrosis volume, Quantitative CT, Interobserver variation, Intraobserver variation

Introduction

Radiotherapy to the tumours of thoracic region often leads to the lung injuries, such as radiation pneumonitis and radiation lung fibrosis, and these injuries may limit the sufficient dose delivery to achieve the adequate local control [1]. Many researchers have investigated about the radiation pneumonitis and its predictive factors including clinical parameters as well as

dose-volume statistics [2-5]. However, there are few studies that have investigated the dosimetric and clinical features of the radiation-induced lung fibrosis (RILF) despite its potential detrimental and sustained effect on long-term cancer survivors.

Previously, we studied RILF features related to dosimetric parameters in non-small cell lung cancer (NSCLC) patients who underwent postoperative radiotherapy [6]. In that study,

Received 23 December 2013, Revised 29 January 2014, Accepted 18 February 2014.

Correspondence: O Kyu Noh, MD, PhD, Department of Radiation Oncology, Ajou University School of Medicine, 206 World cup-ro, Yeongtong-gu, Suwon 443-721, Korea. Tel: +82-31-219-5884, Fax: +82-31-219-5894, E-mail: okyu.noh@gmail.com

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

www.e-roj.org

we implemented the fibrosis volume (VF) on follow-up chest computed tomography (CT) scan as a measurable endpoint for assessing the degree of RILF and the results showed that the VF was consistently associated with the every dose-volume parameters of the lung. The VF on CT images as a measurable endpoint initially has been proposed in analytic category of LENT-SOMA scale, which described as "assessment of lung volume and zones of fibrosis" on CT/MR imaging modalities [7,8]. However, the quantitative measurement of fibrosis zones has been less investigated as the scale for assessing the radiation-induced lung toxicities and this might be due to the lack of objective method to measure. The VF measured by the physician's contouring the extent of the fibrosis implemented in our study also may hamper the reliability of this quantitative endpoint.

Therefore, we suggested a procedure to delineate the extent of the lung fibrosis to improve the reliability of measuring the VF in this study. The purpose of this study was to investigate the inter- and intra-observer variability in the CT-based measurement of VF implementing this new procedure.

Materials and Methods

From the previous study [6], we selected randomly ten patients with RILF. All patients were diagnosed with NSCLC and received postoperative radiotherapy. The types of surgery were lobectomy in 7 and bilobectomy in 3. Nine patients were male and the median age was 70.5 years (range, 49 to 76 years).

Radiotherapy was started within 6 weeks after the operation and all patients were treated with 3-dimensional conformal radiation therapy using the mega-voltage (MV) photon beams (≥ 6 MV). Clinical target volume (CTV) was defined the area including bronchial stump, lymph nodal station with positive tumor deposit, and its next draining lymph nodal station. Planning target volume was extended from the CTV with 1.0 to 1.5 cm of margin. Total dose of 44 to 65 Gy (median, 50.4 Gy) was irradiated with conventional fractionation (1.8 to 2.0 Gy/day).

The VF was delineated on the follow-up chest CT taken at more than 6 months after radiotherapy according to the predefined procedure as described below (Fig. 1). Firstly, we

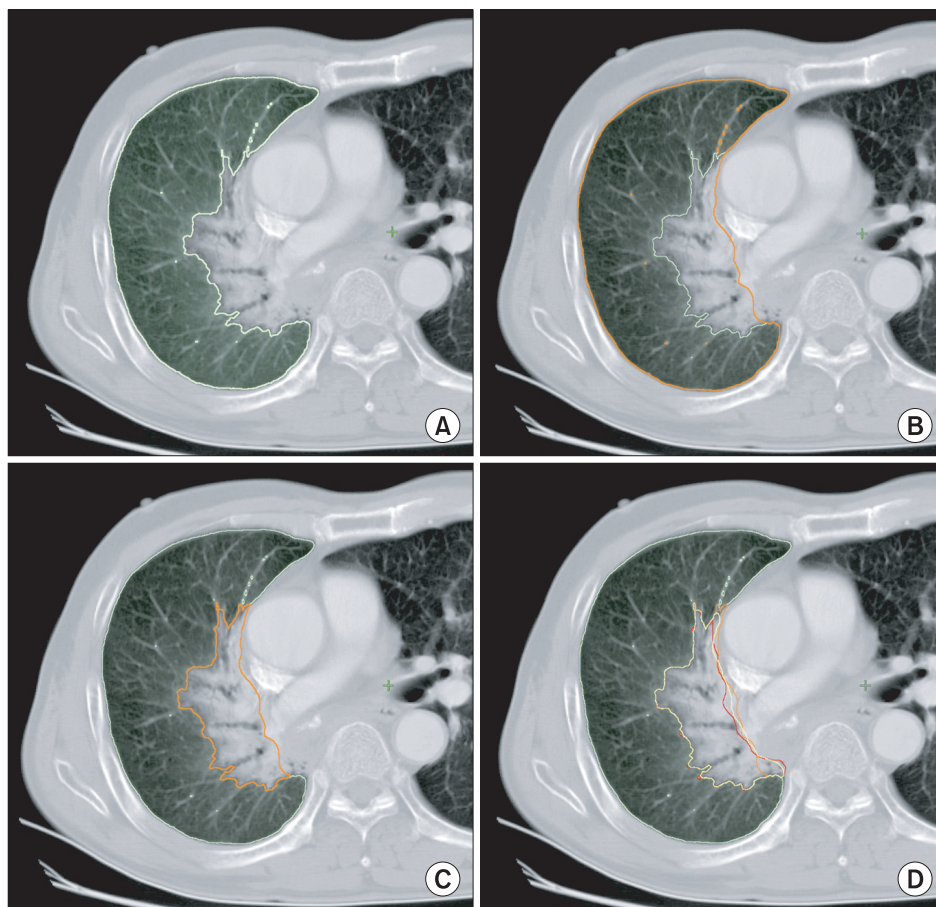


Fig. 1. Procedure of measuring fibrosis volume (VF) on follow-up chest computed tomography images. (A) segmentation of normal lung parenchyma volume (V_{lung}), (B) auxiliary volume ($V_{lung+fibrosis}$) by adding the pulmonary fibrosis area on the normal lung parenchyma volume, (C) measuring VF by subtracting the normal lung volume from the auxiliary volume ($VF = V_{lung+fibrosis} - V_{lung}$), and (D) variable delineated VF among observers.

contoured the volume of normal lung parenchyma (V_{lung}). V_{lung} was segmented automatically based on the CT density range from -500 HU to -1900 HU (Fig. 1A). Then, we made an auxiliary volume ($V_{lung+fibrosis}$) by adding the area of the RILF delineated manually on the normal lung volume (Fig. 1B). RILF was defined the newly developed fibrotic dense area in radiation field excluding the pleural fluid, major pulmonary vessels and air-space just as blebs and main bronchial space. We included the bronchial space distal to second carina level and pulmonary vessels intermingled with fibrotic area. Lastly, we assessed the VF by subtracting the normal lung volume from the auxiliary volume ($VF = V_{lung+fibrosis} - V_{lung}$) (Fig. 1C). All procedure of CT-based FV measurement was done using the Varian Eclipse External Planning System, ver. 7.1 (Varian Medical System, Palo Alto, CA, USA).

Three radiation oncologists independently contoured VF to evaluate the interobserver variability. Of three radiation oncologists, two radiation oncologists contoured the VFs of same patients twice more (total three times) with 1 or 2 weeks blinding interval for the analysis of intraobserver variability (Fig. 1D). We calculated the concordance index (CI) and inter/intra-class correlation coefficient (ICC) for the interobserver and intraobserver variability. The CI is defined as the overlap VF of the three observations to the total delineated VF.

$$Concordance\ index\ (CI) = \frac{VF1 \cap VF2 \cap VF3}{VF1 \cup VF2 \cup VF3}$$

VF1, VF2, and VF3 are three VFs contoured by the three observers or three VFs contoured three times by two observers. Also, we calculated ICC to assess the reproducibility of the result of the test. An ICC is measured on a scale of 0 to 1, 1

presents perfect reliability, whereas 0 indicates no reliability. All statistical analyses were performed with IBM SPSS statistics software ver. 19.0 (IBM, New York, NY, USA).

Results

The time interval between the completion of radiotherapy and the measurement of VF on follow-up chest CT ranged from 6 to 19 months (median, 10 months). The values of mean VF and CI were summarized in Table 1. The median CI was 0.61 (range, 0.44 to 0.68) for interobserver variability. The median CIs of intraobserver variability were 0.69 (range, 0.65 to 0.79) for observer 2 and 0.61 (range, 0.55 to 0.65) for observer 3. The ICC for interobserver variability was 0.974 ($p < 0.001$) and ICCs for intraobserver variability were 0.996 ($p < 0.001$) and 0.991 ($p < 0.001$), respectively.

Discussion and Conclusion

In this study, we investigated the interobserver and intraobserver variability of the CT-based fibrosis measurements and the results showed that this method could be used as one of the reliable and consistent evaluation tools for assessing RILF. The ICC values of more than 0.9 suggests that this evaluation method has very high level of reliability regardless of the repetitive delineations by same or different observer. However, the results showed moderate values of CI ranging from 0.44 to 0.69 for the interobserver and from 0.55 to 0.79 for the intraobserver variability. These moderate CI values are due to the discordances of defining the fibrosis area abutting to pleura of the mediastinum or chest wall (Fig. 1D). The

Table 1. Fibrosis volumes and interobserver and intraobserver concordance indices

Patient	Interobserver		Intraobserver (observer 2)		Intraobserver (observer 3)	
	Mean VF ± SD (mL)	CI	Mean VF ± SD (mL)	CI	Mean VF ± SD (mL)	CI
1	49.3 ± 5.3	0.62	49.3 ± 3.3	0.73	37.9 ± 4.7	0.57
2	103.4 ± 21.4	0.55	102.4 ± 10.2	0.67	84.9 ± 1.5	0.57
3	48.1 ± 2.3	0.59	46.8 ± 0.5	0.70	51.5 ± 2.4	0.61
4	67.3 ± 19.8	0.50	66.4 ± 0.7	0.65	90.6 ± 3.2	0.64
5	41.1 ± 11.1	0.44	34.2 ± 3.6	0.67	41.5 ± 10.1	0.55
6	84.1 ± 6.9	0.57	72.5 ± 4.4	0.65	87.1 ± 2.5	0.62
7	102.4 ± 9.0	0.64	98.4 ± 5.5	0.77	107.3 ± 3.9	0.65
8	90.0 ± 6.1	0.66	91.6 ± 3.0	0.77	88.3 ± 7.7	0.62
9	73.6 ± 1.6	0.65	70.1 ± 4.0	0.68	76.2 ± 5.4	0.59
10	207.1 ± 24.6	0.68	201.7 ± 3.7	0.79	212.5 ± 21.3	0.60

VF, fibrosis volume; SD, standard deviation; CI, concordance index.

pulmonary fibrotic density may not be clearly differentiated from the density of soft tissue or pleural effusion beside the mediastinum or chest wall. The degree of this discordance is similar within or between observers (median CI, 0.61 for interobserver, 0.69 and 0.61 for intraobserver). These interobserver and intraobserver spatial discrepancies seemed to be traded off over the next contouring. As a result, the variability of the defining the border of fibrosis did not affect the reproducible VF measurement. These patterns of high level of ICC with moderate level of CI suggest that the variability occurred randomly and do not hamper the reliability of the CT-based measurement.

Defining the fibrosis zone from normal lung parenchyma is difficult because of the fibrosis border can varies according to the window setting as well as various threshold of each observer. Therefore, we implemented indirect semi-objective contouring method by adding the step of auto-segmentation tool based on the normal lung parenchyma CT density (from -500 HU to -1900 HU) instead of direct delineating the VF (Fig. 1). We considered the area of CT density more than -500 HU as the fibrosis zone of more than RGOG grade 1 [9]. This semi-objective method served identical fibrosis contours closed to normal lung parenchyma to observers in every measurement procedure and it helped to measure the reproducible VF. Although the value of CT density was used in defining the line between the fibrosis zone and normal lung parenchyma, we did not use the value of CT density in delineating the fibrosis zone itself. Despite the use of CT density range seems to be objective method in defining the fibrosis zone, there are several difficulties to implement it. The fibrosis zone developed already prior to radiotherapy is hard to discriminate automatically from the newly developed zone induced by radiotherapy. In a case of the pre-existing fibrosis zone is located in the radiation fields, the exact differentiation become nearly impossible. The radiation changes on CT scans can be showed in various patterns [10] and it may interfere the automatic assessment of the VF. The vessel density also provides one of the ambiguities for measuring the VF. Therefore, we defined the zone of RILF manually by the experienced radiation oncologists just like the target volumes are defined manually by radiation oncologists, who can integrate various information of several diagnostic images and published data, clinical experiences considering many uncertain factors. The values of ICC for interobserver variations suggest that this manual assessment of VF is highly reproducible between observers and can be applied as a reliable endpoint for RILF.

Although our results showed that assessment of the VF could be one of the reliable quantitative endpoints for RILF, the clinical validation should be followed in diverse clinical settings applying it. Our previous study, showing the association between the VF and dose-volume histogram parameters of lung, can be considered one of the validation data of this method [6]. However, this validation was limited to the population with NSCLC patients undergoing postoperative radiation therapy, in which the effect of primary mass was removed by resection. Because the fibrosis zone can be intermingled with the primary tumour, the VF measurement may be inappropriate in the patients with definitive radiotherapy. However, in cases of thoracic irradiation without pulmonary tumour, such as radiotherapy for breast, oesophageal, and mediastinal lymphoma, measurement of the VF can be useful to investigate the radiation-induced lung damages. The extent of fibrosis zone can be influenced by the radiation change itself. The shrinkage or expansion of VF can be occurred by the inflammatory responses to the radiation. The extent of fibrosis can be changed over time and may be affected by the combined adjuvant chemotherapy. Therefore, these factors should be considered and controlled to apply and interpret the measurement of VF.

The clinical significance of the VF can limit the usefulness of this quantitative method for the RILF. Although it has been postulated that the radiation fibrosis could affect negatively to the clinical outcome for the patients treated with thoracic radiotherapy, there is no definite evidence about its detrimental effects until now. The further studies investigating the clinical impacts on pulmonary function, survival rate as well as quality of life for long-term survivors. The quantitative method suggested in this study may help to analyse the outcomes in the further studies.

In conclusion, CT-based measurement of VF with patients who received postoperative radiation therapy was a highly consistent and reproducible quantitative method between/within observers. The clinical significance and usefulness of this method should be validated in the further investigations.

Conflicts of Interest

No potential conflict of interest relevant to this article was reported.

References

1. Mehta V. Radiation pneumonitis and pulmonary fibrosis in non-small-cell lung cancer: pulmonary function, prediction, and prevention. *Int J Radiat Oncol Biol Phys* 2005;63:5-24.
2. Hernando ML, Marks LB, Bentel GC, et al. Radiation-induced pulmonary toxicity: a dose-volume histogram analysis in 201 patients with lung cancer. *Int J Radiat Oncol Biol Phys* 2001; 51:650-9.
3. Claude L, Perol D, Ginestet C, et al. A prospective study on radiation pneumonitis following conformal radiation therapy in non-small-cell lung cancer: clinical and dosimetric factors analysis. *Radiother Oncol* 2004;71:175-81.
4. Rancati T, Ceresoli GL, Gagliardi G, Schipani S, Cattaneo GM. Factors predicting radiation pneumonitis in lung cancer patients: a retrospective study. *Radiother Oncol* 2003;67:275-83.
5. Moiseenko V, Craig T, Bezjak A, Van Dyk J. Dose-volume analysis of lung complications in the radiation treatment of malignant thymoma: a retrospective review. *Radiother Oncol* 2003;67:265-74.
6. Oh YT, Noh OK, Jang H, et al. The features of radiation induced lung fibrosis related with dosimetric parameters. *Radiother Oncol* 2012;102:343-6.
7. Pavy JJ, Denekamp J, Letschert J, et al. Late effects toxicity scoring: the SOMA scale. *Radiother Oncol* 1995;35:11-5.
8. LENT SOMA scales for all anatomic sites. *Int J Radiat Oncol Biol Phys* 1995;31:1049-91.
9. Cox JD, Stetz J, Pajak TF. Toxicity criteria of the Radiation Therapy Oncology Group (RTOG) and the European Organization for Research and Treatment of Cancer (EORTC). *Int J Radiat Oncol Biol Phys* 1995;31:1341-6.
10. Libshitz HI, Shuman LS. Radiation-induced pulmonary change: CT findings. *J Comput Assist Tomogr* 1984;8:15-9.