Clinical Target Volume Automatic Segmentation Based on Lymph Node Stations for Lung Cancer with Bulky Lump Lymph Nodes

Mingyi Di
Peking Union Medical College Hospital
Jing Shen
Peking Union Medical College Hospital
Shaobin Wang
Medmind Technology Co,Ltd.
Qi Chen
Medmind Technology Co,Ltd.
Yu Chen
Medmind Technology Co,Ltd.
Zhikai Liu
Peking Union Medical College Hospital
Xin Lian
Peking Union Medical College Hospital
Jiabin Ma
Peking Union Medical College Hospital
Yantian Pang
Peking Union Medical College Hospital
Tingting Dong
Peking Union Medical College Hospital
Bei Wang
Peking Union Medical College Hospital
Qiu Guan
Peking Union Medical College Hospital
Lei He
Peking Union Medical College Hospital
Yue Zhang
Peking Union Medical College Hospital
Hao Liang
Peking Union Medical College Hospital
Jie Shen (shenjiePUMCH@163.com)

Peking Union Medical College Hospital

Fuquan Zhang

Peking Union Medical College Hospital

Article

Keywords:

Posted Date: April 19th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1545690/v1

Abstract

The lack of standardized delineation of lymph node station in lung cancer radiotherapy leads to nonstandard clinical target volume (CTV) contouring, especially in patients with bulky lump gross target volume lymph nodes (GTVnd). Therefore, this study defines lymph node region boundaries in radiotherapy for lung cancer and automatically contours lymph node stations based on the International Association for the Study of Lung Cancer (IASLC) lymph node map. Computed tomography (CT) scans of 200 patients with small cell lung cancer are collected. The lymph node zone boundaries are defined based on the IASLC lymph node map, with adjustments to meet radiotherapy requirements. Contours of lymph node stations are confirmed by two experienced oncologists. A model (DiUNet) is constructed by incorporating the contours of GTVnd to precisely contour the boundaries. Quantitative evaluation metrics and clinical evaluations are conducted. This is the first study to propose a method that automatically contours lymph node regions station by station based on the IASLC lymph node map with bulky lump GTVnd. Delineation of lymph node stations based on the DiUNet model is a promising strategy to obtain accuracy, consistency, and efficiency for CTV delineation in lung cancer patients, especially for bulky lump GTVnd.

Introduction

Precise delineation of the clinical target volume (CTV) is essential for accurate, individualized radiotherapy (RT). This task is time-consuming and relies largely on the experience of oncologists. Intra- and inter-observer variations exist despite following the same contouring guidelines [1–3]. Metastasis in lymph nodes (LNs) is very common in lung cancers, especially in patients who are inoperable. Treatment guidelines for lung cancer recommend including the entire pathologically affected LN station in contouring the CTV [4, 5] in accordance with the Radiation Therapy Oncology Group (RTOG) regulations [6]. Thus, accurate contouring of the involved thoracic regional LN is a necessary step in the accurate definition of a CTV in RT for lung cancer; however, this accuracy is not usually achieved as clinical delineation of a CTV usually depends on the experience of the oncologist and results in great heterogeneity. Furthermore, no previous studies have defined the boundary of the mediastinal LN region in RT or the delineation range of the involved LNs. When we contour a certain lymph node region, there is no relatively reasonable evaluation system for the accuracy of its boundaries. This problem is present in both small cell lung cancer(SCLC) and non-small cell lung cancer(NSCLC).

Convolutional neural network (CNN)-based methods have been successfully applied to identify organ structures and CTVs for contouring several cancers. Deep learning-related published data for lung cancer have predominantly focused on organs at risk (OAR) [7–10], GTV of primary tumours [11], and CTV in postoperative radiation therapy (PORT) [12]. Bi et al. built a deep learning model [12] for PORT-CTV delineation; however, it could only be used in RT for non-small cell lung cancer (NSCLC) patients after complete resection, and the CTV contouring mostly depended on the experience of the oncologists. Zhang et al. proposed a modified ResNet model for auto-segmentation of the GTV of NSCLC [11], but only

the primary tumour mass was assessed for patients with LN involvement. To date, a machine learning model of automatic contouring of LN stations or CTV in patients with bulky lump gross target volume LNs (GTVnd) has not been reported. Therefore, we cannot currently accurately identify the LN region contained in a CTV and whether it was fully included in the CTV.

Since the radiographic imaging of the lymph node partition is based on the anatomical structure, which also associated with the site of lymph node metastasis in lung cancer. LN region boundaries in RT for lung cancer were redefined in this study for accuracy, using classifications from the International Association for the Study of Lung Cancer (IASLC). Based on this, a previously proposed CNN model, whicth successfully used in cervical cancer in our center [13], was employed to automatically contour 1–11 LN stations in 200 patients. The GTVnd was added to the CT slice as a second input to improve the accuracy of LN region delineation for bulky lump GTVnd cases. This is the first study to define LN region boundaries for lung cancer in RT and propose automatic segmentation of the LN stations, which can assist in consistently and automatically contouring the CTVs of lung cancer with involved LN stations. This work also allowed different doctors to uniformly and accurately regulate the inclusion range of lymph nodes in CTV delineation.

Methods

Data and preprocessing

CT data from 200 patients with stage III-IV small cell lung cancer (SCLC) were collected from November 2010 to January 2021. The total slice number was 16676, data were 512 × 512 p, pixel spacing ranged from 0.89 × 0.89 to 1.32 × 1.32 mm, and a 5 mm thickness was acquired with a Brilliance CT Big Bore (Philips Healthcare, Best, the Netherlands).

GTVnd, defined as enlarged regional LNs (> 1 cm in the short axis), were delineated according to the International Commission on Radiation Units and Measurements Report 50 guidelines. GTVnd were outlined on mediastinal and lung windows with no expansion for potential microscopic disease. A CTV was defined according to the RTOG guidelines [14, 15]. The LN stations, GTVnd, and CTV manually delineated by trained radiation oncologists were used as the segmentation man contouring (MC). The LN stations were contoured based on the widely accepted anatomic definitions of the nodal stations in the IASLC [16] (Supplementary Table S1 and Fig. S1); examples of the contouring of each LN station are shown in Fig. 1.

During the contouring process, corrections were made for the boundaries of the LN regions to ensure CTV coherence. For example, the blood vessels were not avoided in the contouring of the supraclavicular zone (Station 1R/L), as required by the guidelines in the IASLC. Another example is the perivascular LNs (Station 3A), where the posterior border in the IASLC was located in front of the anterior border of the superior vena cava and left carotid artery. Here, blood vessels were included in Stations 2 and 3A. Involved LNs were defined as metabolically avid in positron emission tomography or abnormally enlarged

regional LNs measuring over 1.0 cm along their short axis in CT scans. Additionally, the CTV contained the entire LN station defined by the IASLC LN map, according to the RTOG regulations. Moreover, the CTV was restrained to avoid extending beyond anatomic boundaries, such as the chest wall, vertebral body, great vessels, heart, and oesophagus, except for verifiable sufficient evidence of the invasion.

The boundary was extended as far as possible in contouring because bulky lump GTVnd in LN stations caused boundary deformation. The GTVnd crossed the 10R and 7 Stations (contoured independently, Fig. 2(a,b)) and spread beyond the boundary of Station 4R (Fig. 2(c,d)). This pattern is common in centrally located lung cancers and greatly increased the heterogeneity of the training sets. The LN stations contained all the GTVnd to confirm the anatomical boundaries defined by the IASCL.

All delineations were reviewed and modified by 2 oncologists with more than 10 years of experience in RT for lung tumours to ensure quality.

Network architectures of DiUnet

It was difficult for existing AI algorithms to accurately define the LN stations because their boundaries varied widely as the characteristics of the GTVnd changed. The previous machine learning model for GTV aimed at NSCLC patients' primary tumours and failed to define accurate boundaries of GTVnd. To incorporate the GTV boundary into the LN stations, prior expert GTV knowledge was encoded into the model by adding a GTV mask as an additional input. A U-Net backbone architecture was used in the detailed CNN architecture, consisting of two encoding paths and a decoding path (Supplementary Fig. S2). The encoding part aggregated semantic information from the GTV masks and original CT slices by reducing the spatial information to learn features from part to whole. The extracted feature maps were summed and input to the decoding part, which received semantic information from the bottom. The decoding path was combined with the two encoding paths using skip connections, which concatenated features at multiple levels and recovered the image details passed from the convolutional layers in the encoding paths. The dual-path network architecture was used in the encoders to improve feature extraction. Four micro-blocks that combined residual [17] and dense blocks [18] into a dual path architecture were embedded into the decoder to replace the standard convolution operations. As a 2.5D architecture, the input to the model was three adjacent 2D CT slices, and the output was the corresponding segmentation result of the middle CT slice. The number of channel outputs was the same as the desired number of structures.

This model combined the information of the CT slices and GTV masks. The output mask boundaries were constrained by the spatial features extracted from the GTV boundaries, leading to greater accuracy.

Quantitative evaluation metrics

The Dice similarity coefficient (DSC) [19] and 95th percentile Hausdorff distance (95HD) [20] were used to quantify the contouring accuracy. The DSC measured the relative volumetric overlap between two segmented masks, with a higher value indicating a higher overlap ratio; 1 means the two masks are identical. The 95HD reflected the alignment between two contours, with a higher value indicating a larger

difference. The DSC and 95HD were calculated at a 3D level, and mean ± standard deviation (SD) values were calculated by averaging the values.

Clinical Evaluation

Oncologist evaluation

Two oncologists who did not participate in the LN stations and GTVnd contouring performed the subject evaluation on 20 randomly selected patients. The manual reference contours were separated into the MC group, whereas the corresponding contours generated by the proposed model were placed in the artificial intelligence (AI) group. Then, 10% of the AI and GT results were randomly selected for each LN station (485) and CTV evaluation (46) and randomly labelled as 1 or 2. If AI was labelled 1, then MC was 2. Two oncologists were asked to blindly evaluate these results slice-by-slice and grade each automatic contouring result together based on the curative effect with a 4-grade criterion (Supplementary Table S2).

Consistency Test

Another 10% of the slices from each LN zone and CTV in each case were randomly selected to simultaneously mark the contours of the AI and MC, of which only one structure was presented for each slice, the colours were randomly swapped, and one was blindly selected by each oncologist that was better for clinical applications. If the AI group was better, it was recorded as a positive result; otherwise, it was negative.

Time Cost

The processing time was measured for AI, pre-, and post-AI assistance in the delineation of LN stations and formation of CTV for post-modified radical mastectomy RT.

Statistical Analysis

An independent sample t-test was performed to determine if the differences in the scores provided by the oncologists were statistically significant between AI and MC. Furthermore, McNemar's and consistency tests were performed for consistency evaluation.

Results

The full CT image dataset was randomly divided into training-validation (180 patients) and test (20 patients) datasets at a ratio of 9:1. Statistical analysis of significant differences in the location of the primary tumour, tumour anatomic site, and GTVnd cases between the training-validation and test datasets was performed using the chi-square test. Age was tested using an independent sample t-test with statistical significance set at a two-tailed $P < 0.05$. The general characteristics of the patients in the training and test datasets showed no significant differences (Table S3). The average DSC and 95HD results for each LN region of the DiUNet are shown in Table 1.

Table 1
DSC and 95HD values of proposed model in each LN station.

Lymph node station	DiUNet	
	DSC ± STD	95HD ± STD(mm)
1R	0.79 ± 0.18	4.04 ± 6.40
1L	0.83 ± 0.14	2.72 ± 3.12
2R	0.88 ± 0.06	2.36 ± 1.43
2L	0.89 ± 0.04	1.92 ± 0.99
3A	0.86 ± 0.04	2.36 ± 0.87
3P	0.75 ± 0.11	4.02 ± 3.95
4R	0.92 ± 0.04	1.61 ± 0.55
4L	0.81 ± 0.08	2.86 ± 1.59
5	0.78 ± 0.15	2.21 ± 1.48
6	0.78 ± 0.08	2.40 ± 1.06
7	0.79 ± 0.09	3.35 ± 2.44
8	0.65 ± 0.13	4.68 ± 3.44
10R	0.86 ± 0.08	2.35 ± 0.99
10L	0.89 ± 0.05	2.00 ± 0.92
11R	0.72 ± 0.12	3.50 ± 2.06
11L	0.77 ± 0.13	3.17 ± 2.04

In the test dataset, DiUNet provided a suitable CTV for clinical application. The LN station boundaries were accurate even for bulky lump GTVnd (Fig. 3).

The oncologist evaluation results of contouring 418 slices of LN stations are listed in Table 2. Using the scoring criteria for contour evaluation, most LN station contours were deemed clinically acceptable (≥ 2) by both oncologists. The percentages of clinically acceptable scores were 99.18% (481/485) and 98.35% (477/485) for AI and MC, respectively, for oncologist A, 98.56% (478/485) and 97.94% (475/485) respectively, for oncologist B. The total percentage of clinically acceptable scores for both oncologists was 98.87% (959/970) in AI group.

The overall average scores were 2.914 and 2.899 for AI and MC, respectively, which showed no significant difference ($P > 0.05$). However, a significant difference was observed in the evaluations of both

oncologists (P < 0.05).

Table 2
Graded oncologist evaluations for AI and MC LN station contours.

Oncologist	**A**		**B**	
Score	AI	MC	AI	MC
0	0	0	0	0
1	4	8	7	10
2	35	36	27	26
3	446	441	451	449
Mean score	2.911	2.893	2.916	2.905
P value*	0.541		0.338	

*** Tested by an independent sample t-test**

To build the CTV for the deep learning model, a total of 46 (10%) of the CTV slices in the test dataset were randomly selected to simultaneously mark the contours of AI and MC. The graded oncologist evaluations for AI contours were compared to those of MC in CTV (Table 3); there were no significant differences in the evaluations of both oncologists (P0.05).

Table 3
Graded oncologist evaluations for CTV contours in AI and MC.

Oncologist	**A**		**B**	
Score	AI	MC	AI	MC
0	0	0	0	0
1	0	0	1	0
2	8	7	13	11
3	38	39	32	35
Mean score	2.826	2.848	2.674	2.761
P value*	0.577		0.069	

*** Tested by an independent sample t-test**

The merits of the DiUNet model were separately compared with those of the MC group. The overall positivity rates for oncologists A and B were 53.20% (258/485) and 52.37% (254/485), respectively. This demonstrated that the proposed DiUNet performed equally as well as human delineation. McNemar's test was statistically significant ($P < 0.05$), and the kappa consistency index was 0.276 ($P < 0.05$), which indicated that the oncologists had different positivity rates and poor consistency, respectively (Table S4).

The CTV with involved LN stations was contoured using DiUNet, and the oncologists (slice-by-slice) in the test dataset. The positivity rates produced by the two oncologists were 50% (23/46) and 54.35% (25/46) (Table S5). McNemar's test was statistically significant ($P < 0.05$), and the kappa consistency index was 0.261 ($P < 0.05$), which indicated that the oncologists had different positivity rates and poor consistency, respectively.

The proposed method took approximately 5 min per test case from DICOM file parsing to contour generation. It took an oncologist more than 80 min to completely delineate the CTV but with the assistance of AI that was reduced to 25 min. Therefore, the proposed model can efficiently shorten the contouring time for oncologists.

Discussion

Currently, there is no accurate definition of LN station boundaries, and oncologists mainly delineate CTV in lung cancer based on their clinical experience, resulting in great heterogeneity. When the CTV region is described, neither the specific LN regions contained within nor their range boundaries can be accurately determined. Many studies have found this problem, but none have found an effective solution. Zhu Guangying and his colleague investigate the consensus and controversies on delineation of radiotherapy target volume for patients with NSCLC in 10 radiation departments in china and 2 departments in US. The

delineation of the CTV in mediastinal lymph nodes varied greatly [23]. In Femke O B Spoelstra' s study in 2010, seventeen thoracic radiation oncologists were invited to contour their routine CTV for 2 representative NSCLC patients, the delineation of the CTV was vary greatly between different doctors at different times. As showed in Supplementary Figure S3, CTV were contoured at an interval of 2 weeks by six oncologists [21]. Another study in 2019 also confirmed this view, image interpretational differences can lead to large interobserver variation particularly when delineating the gross tumor volume lymph node [22]. In other words, even though the CTV containes same lymph node zones, the boundaries of each region vary greatly in different oncologist. The heterogeneity in the CTV is particularly evident when there is bulky lump GTVnd in the LN stations. Although automatic segmentation models for CTV in lung cancer have been proposed, they failed in cases involving bulky lump GTVnd. In this study, the boundaries of each LN region were accurately defined for the first time in RT, and a CNN-based model was trained to auto contour LN stations for accurate and consistent CTV formation, which is an important innovation point of this study.

The boundaries of LN stations vary widely in shape, location, and size of GTVnd changes. Therefore, the prior GTVnd expert knowledge was incorporated as an additional training condition of the proposed DiUnet model, which constrained the output mask boundaries using the spatial features extracted from the GTVnd boundaries, leading to more accurate results.

Examples of the LN stations contoured using DiUNet are shown in Fig. 3, in which LN boundaries were contoured accurately even for bulky lump GTVnd. This has great value in CTV contouring, especially for central lung carcinoma patients. Therefore, it was first applied to auto-contouring with machine learning models for lung cancer.

The DSC values of DiUNet in each LN station were greater than 0.7, except for Station 8 (Fig. 4). The inaccuracy of contouring the oesophagus may explain this lower value.

The percentage of clinically acceptable scores was greater than 98% and the average score was higher than 2.91 in the DiUNet model (Table 2). LN stations contoured by AI were acceptable compared with human-generated structures. The CTV evaluation results are presented in Table 3. Despite the differences in the evaluation systems among physicians, it is believed that the CTV contoured by DiUNet can be clinically acceptable. DiUNet passed the consistency test with an overall positivity rate higher than 50% in the comparison of LN stations and CTV. The concordance of data between the two oncologists was poor, but they both considered that LN stations and the CTV contoured with AI models were not inferior to those of MC (Tables S4 and S5).

Accurate definition and complete delineation of lymph node stations' boundaries are an important basis for the formation of standardized CTV. This is the first study to delineate lymph node stations to promote CTV standardization.

It should be noted that most slices that required corrections were located at the edges adjacent to the blood vessel or at the border of the LN station. This may be because accurate delineation needs to

integrate information from multiple upper and lower slices, but the information available at the boundary is limited. In the future, we are planning to explore the impact of CTV on survival prognoses that are comprised of intact partially involved LN regions, local recurrence, and rate of nodal failures.

Conclusions

The purpose of this study was to define LN region boundaries in RT for lung cancer and automatically contour LN stations based on the IASLC LN map. This is the first study to propose this approach, and qualitative and quantitative experiments demonstrated the effectiveness and efficiency of the proposed DiUNet model for CTV consisting of involved LN stations, especially for bulky lump GTVnd cases.

Declarations

Authors' contributions

B.W., H.L., J.M., J.S., L.H., M.D., Q.C., Q.G., S.W., T.D., X.L., Y.C., Y.P., Y.Z. and Z.L. collected the data and finished the quality control of data. F.Z. and J.S. provided the study concepts, designed the study, M.D. performed statistical analysis. J.S. and M.D. wrote the manuscript. All the authors read and approved the final manuscript.

Availability of data and materials

The datasets generated and/or analysed during the current study are not publicly available due the need of protecting patients' privacy and patent right, but are available from the corresponding author on reasonable request.

Additional information

Competing interests

The authors declare that they have no competing interests

Funding

This work was supported by grants from the National 135 Major Project of China (2016YFC0105207) and the National Foundation for Education Sciences Planning (grant number BLA200216).

Board Approval

All experimental protocols were approved by The Institutional Review Board(IRB) of Peking Union Medical College Hospital (PUMCH). All methods were carried out in accordance with relevant guidelines and regulations.

Written consent

Informed consent was given by all patients.

References

1. S Dewas, C Vautravers-Dewas, P Blanchard, Y Pointreau, F Denis, T Lacornerie et al. Delineation variation of lung cancer in conformal radiotherapy among radiation oncology residents after an educational course: a prospective study. Cancer radiother **14**(**2**), 103–110 (2010).
2. Yatman Tsang, Peter Hoskin, Emiliano Spezi, David Landau, Jason Lester, Elizabeth Miles et al. Assessment of contour variability in target volumes and organs at risk in lung cancer radiotherapy. Tech Innov Patient Support Radiat Oncol **10**, 8–12 (2019).
3. Liang J, Li M, Zhang T, Han W, Chen D, Hui Z. The effect of image-guided radiation therapy on the margin between the clinical target volume and planning target volume in lung cancer. J Med Radiat Sci **61**(**1**):30–37 (2014).
4. Ettinger DS, Wood DE, Akerley W, Bazhenova LA, Borghaei H, Camidge DR et al. Non-small cell lung cancer, version 6.2015. J Natl Compr Canc Netw **13**(**5**), p 515–524 (2015).
5. De Ruysscher D, Faivre-Finn C, Nestle U, Hurkmans CW, Le Péchoux C, Price A, Senan S. European Organisation for Research and Treatment of Cancer recommendations for planning and delivery of high-dose, high-precision radiotherapy for lung cancer. J Clin Oncol **28**(**36**), 5301–5310 (2010).
6. Bradley JD, Paulus R, Komaki R, Masters G, Blumenschein G, Schild S et al. Standard-dose versus high-dose conformal radiotherapy with concurrent and consolidation carboplatin plus paclitaxel with or without cetuximab for patients with stage IIIA or IIIB non-small-cell lung cancer (RTOG 0617): A randomised, two-by-two factorial phase 3 study. Lancet Oncol **16**(**2**), 187–199 (2015).
7. Lustberg T, van Soest J, Gooding M, Peressutti D, Aljabar P, van der Stoep J et al. Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer. Radiother Oncol **126**(**2**), 312–317 (2018).
8. Giri MG, Cavedon C, Mazzarotto R, Ferdeghini M. A Dirichlet process mixture model for automatic ^{18}F-FDG PET image segmentation: Validation study on phantoms and on lung and esophageal lesions. Med Phys **43**(**5**), 2491 (2016).
9. Feng X, Qing K, Tustison NJ, Meyer CH, Chen Q. Deep convolutional neural network for segmentation of thoracic organs-at-risk using cropped 3D images. Med Phys **46**(**5**), 2169–2180 (2019).
10. Yang J, Veeraraghavan H, Armato SG III, Farahani K, Kirby JS, Kalpathy-Kramer J et al. Autosegmentation for thoracic radiation treatment planning: A grand challenge at AAPM 2017. Med Phys **45**(**10**), 4568–4581 (2018).
11. Zhang F, Wang Q, Li H. Automatic segmentation of the gross target volume in non-small cell lung cancer using a modified version of ResNet. Technol Cancer Res Treat **19**, 1533033820947484 (2020).
12. Bi N, Wang J, Zhang T, Chen X, Xia W, Miao J et al. Deep learning improved clinical target volume contouring quality and efficiency for postoperative radiation therapy in non-small cell lung cancer.

Front Oncol **9**,1192 (2019).

13. Liu Z, Liu X, Guan H, Zhen H, Sun Y, Chen Q et al. Development and validation of a deep learning algorithm for auto-delineation of clinical target volume and organs at risk in cervical cancer radiotherapy. Radiother Oncol **153**, 172–179 (2020).
14. Jeremic B, Gomez-Caamano A, Dubinsky P, Cihoric N, Casas F, Filipovic N. Radiation therapy in extensive stage small cell lung cancer. Front Oncol **7**, 169 (2017).
15. Nestle U, De Ruysscher D, Ricardi U, Geets X, Belderbos J, Pöttgen C et al. ESTRO ACROP guidelines for target volume definition in the treatment of locally advanced non-small cell lung cancer. Radiother Oncol **127**(**1**), 1–5 (2018).
16. El-Sherief AH, Lau CT, Wu CC, Drake RL, Abbott GF, Rice TW. International Association for the Study of Lung Cancer (IASLC) lymph node map: Radiologic review with ct illustration. RadioGraphics **34**(**6**), 1680–1691 (2014).
17. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *In: Proc, 2016 IEEE conf. Comp Vis Pattern Recognit (CVPR)*, Las Vegas, NV:770–778 (2016).
18. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. *In: Proc IEEE Conf. Comput Vis Pattern Recognit (CVPR)*, **1**(**2**), p 4700–4708 (2017).
19. Dice LR. Measures of the amount of ecologic association between species. Ecology **26**(**3**), 297–302 (1945).
20. Huttenlocher DP, Klanderman GA, Rucklidge WJ. Comparing images using the Hausdorff distance. IEEE Trans Pattern Anal Mach Intell **15**(**9**), 850–863 (1993).
21. Femke OB Spoelstra, Suresh Senan, Cecile Le Péchoux, Satoshi Ishikura, Francesc Casas, David Ball et al. Variations in target volume definition for postoperative radiotherapy in stage III non-small-cell lung cancer: analysis of an international contouring study. Int J Radiat Oncol Biol Phys **76**(**4**), 1106–1113 (2010).
22. Susan Mercieca, José S A Belderbos, Angela van Baardwijk, Stefan Delorme, Marcel van Herk. The impact of training and professional collaboration on the interobserver variation of lung cancer delineations: a multi-institutional study. Acta Oncologica **58**(**2**), 200–208 (2019).
23. Zhu Guangying, Xia Tingyi, Wang Lvhua, Gao Xianshu, Wang Junjie, Li Gaofeng et al. Consensus and controversies on delineation of radiotherapy target volume for patients with non-small cell lung cancer. Chin J Rasiat Oncol **17**(**6**), 432–436 (2008).

Figures

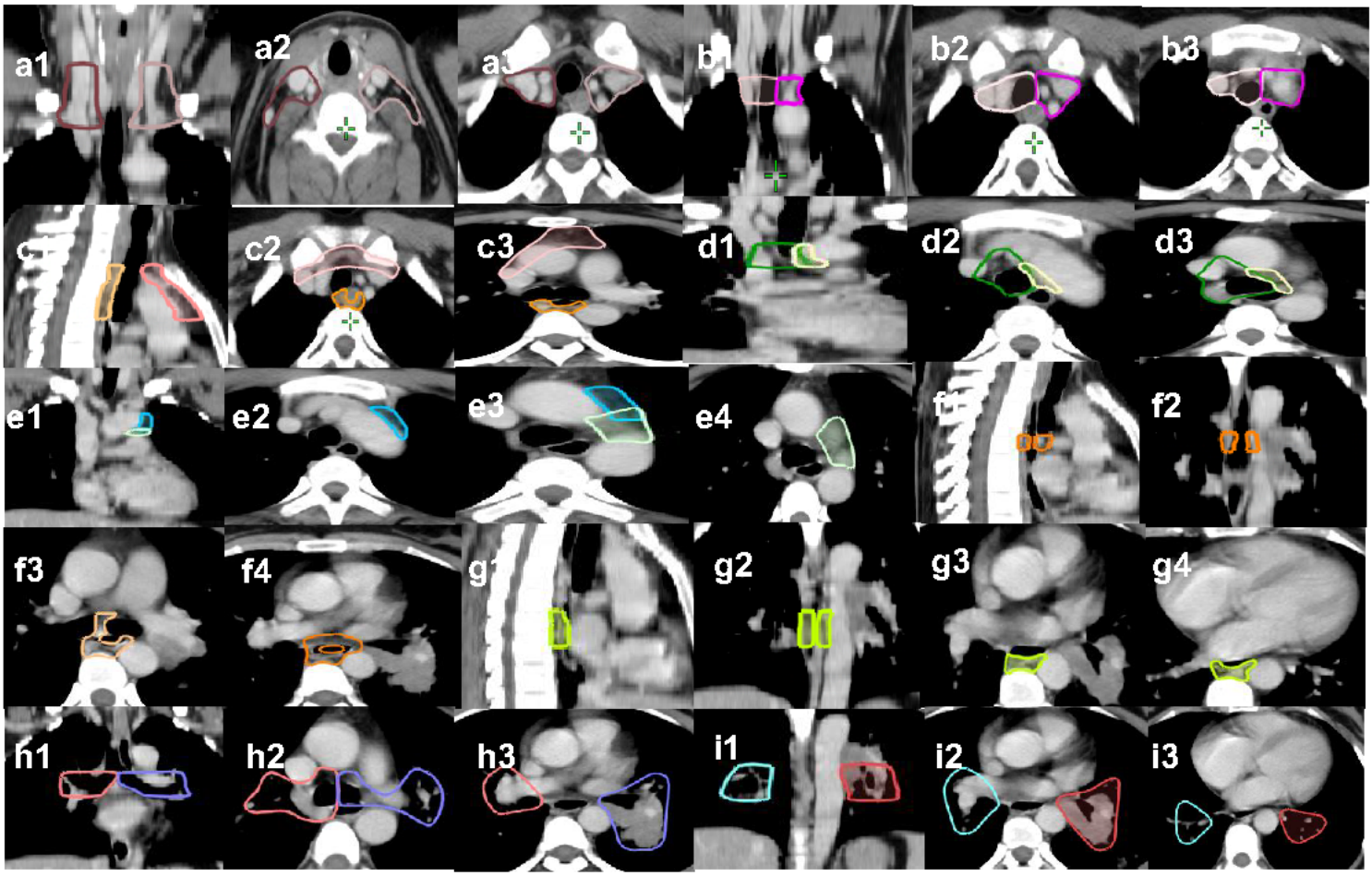

Figure 1

Examples of each LN station of the oncologists contours.

a1: RGB1286465: Station 1L; RGB1286368: Station 1R;

a2: Station 1R/L upper border;

a3: Station 1R/L lower border.

b1: RGB2380238:Station 2L; RGB255204204:Station 2R;

b2: Station 2R/L upper border;

b3: Station 2R/L lower border.

c1: RGB255144144:Station 3A; RGB2551500:Station 3P;

c2: Station 3A/P upper border;

c3: Station 3A/P lower border.

d1: RGB240240128: Station 4L; Dark green: Station 4R;

d2: Station 4R/L upper border;

d3: Station 4R/L lower border.

e1: RGB0200255: Station 6; RGB164255164: Station 5;

e2: Station 6 upper border;

e3: Station 5 upper border; Station 6 lower border; e4: Station 5 lower border.

f1, f2: Orange: Station 7;

f3: Station 7 upper border;

f4: Station 7 lower border.

g1, g2: RGB192255: Station 8;

g3: Station 8 upper border;

g4: Station 8 lower border.

h1: RGB00253: Station 10L; Pink: Station 10R;

h2: Station 10R/L upper border;

h3: Station 10R/L lower border.

i1: RGB2307676: Station 11L; RGB0254254: Station 11R;

i2: Station 11R/L upper border;

i3: Station 11R/L lower border.

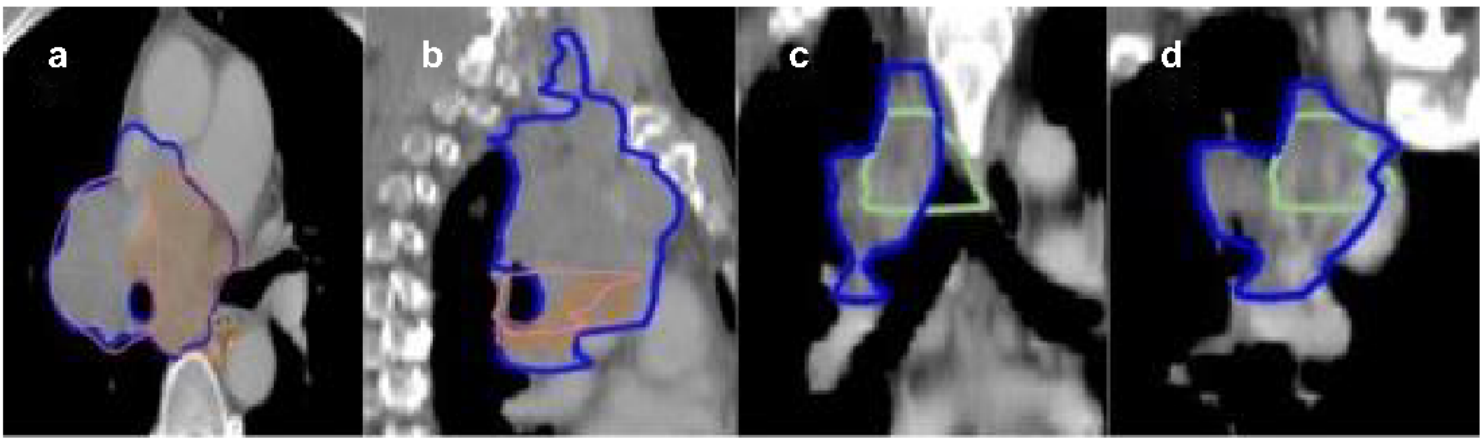

Figure 2

Contouring of GTVnd (blue) and LN Stations 10R (pink), 7 (orange), and 4R (dark green).

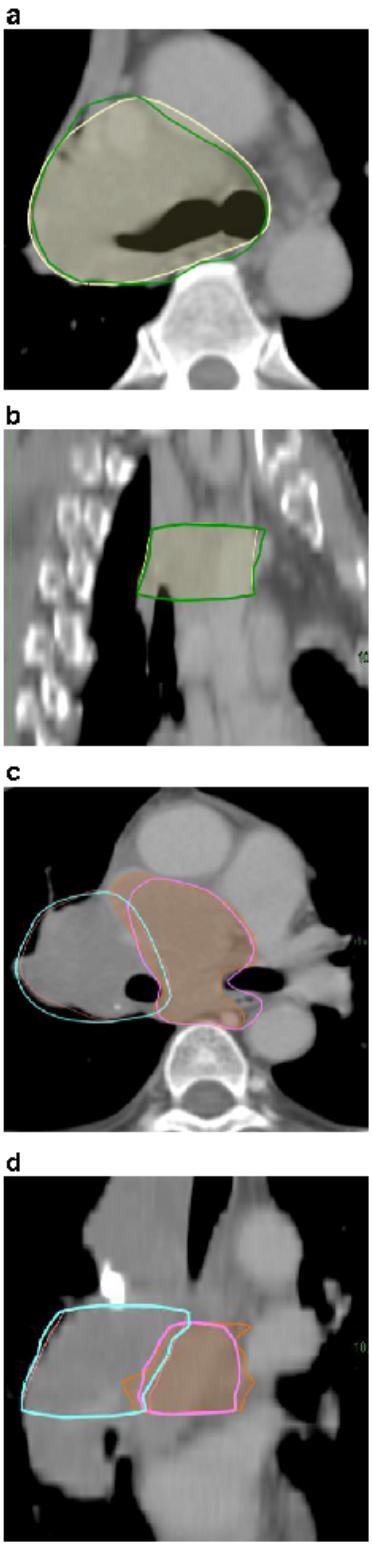

Figure 3

Example of accurate boundary delineation in cases with large GTVnd presented in the LN regions.

(a) and (b) Station 4R contoured by oncologists (Dark green) and DiUNet (RBG240240128). (c) and (d) Station 10R contoured by oncologists (pink) and DiUNet (cyan), and Station 7 contoured by oncologists (orange) and DiUNet (purple).

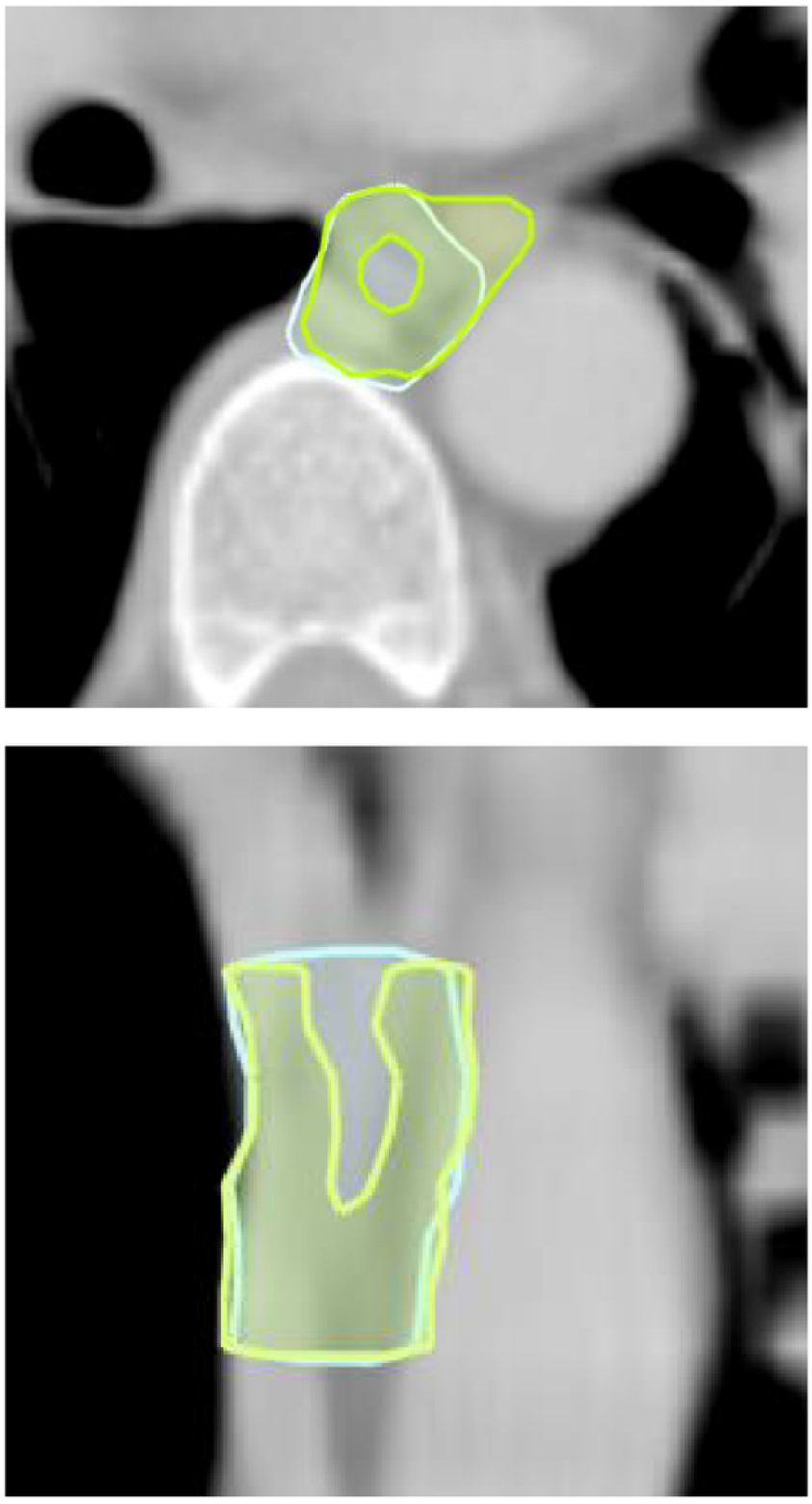

Figure 4

Example of inaccurate contouring for Station 8: (a) RGB1922550 contoured by oncologists and (b) RGB196255255 contoured by DiUNet.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementary412.docx
- supplementaryFigureS1.pdf
- supplementaryFigureS2.pdf
- supplementaryFigureS3.jpg