

Interobserver variability in target volume delineation in definitive radiotherapy for thoracic esophageal cancer: A multi-center study from China

Xiao Chang

Department of Radiation Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100021, China.

Wei Deng

Department of Radiation Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100021, China.

Xin Wang

Department of Radiation Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100021, China.

Zongmei Zhou

Department of Radiation Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100021, China.

Jun Yang

Xinxiang Medical University First Affiliated Hospital

Wenling Guo

Xinxiang Medical University First Affiliated Hospital

Miaoling Liu

Department of Radiation Oncology, Affiliated Hospital of Hebei University, Baoding, 071000, China

Xiaolu Qi

Department of Radiation Oncology, Affiliated Hospital of Hebei University, Baoding, 071000, China

Ling Li

Department of Oncology, Tengzhou Central People's Hospital, Tengzhou, 277599, China.

Kaixian Zhang

Department of Oncology, Tengzhou Central People's Hospital, Tengzhou, 277599, China.

Min Zhang

Department of Oncology, Tengzhou Central People's Hospital, Tengzhou, 277599, China.

Yonggang Shi

Department of Radiation Oncology, The First Affiliated Hospital of Zhengzhou University, Henan Cancer Hospital, Zhengzhou, Henan, China

Ke Liu

Department of Radiation Oncology, The First Affiliated Hospital of Zhengzhou University, Henan Cancer Hospital, Zhengzhou, Henan, China

Yidian Zhao

Department of Radiation Oncology, Anyang tumor hospital, The fourth affiliated hospital of Henan University of science and technology, Anyang, 455000, China.

Huitao Wang

Department of Radiation Oncology, Anyang tumor hospital, The fourth affiliated hospital of Henan University of science and technology, Anyang, 455000, China.

Zhilong Yu

Department of Radiation Oncology, The Affiliated Hospital of Inner Mongolia Medical University Tongdao Road, Hohhot 10050, Inner Mongolia, P.R. China

Jihong Zhang

Department of Radiation Oncology, The Affiliated Hospital of Inner Mongolia Medical University Tongdao Road, Hohhot 10050, Inner Mongolia, P.R. China

Lihua Wang

Department of Radiation Oncology, The Affiliated Hospital of Inner Mongolia Medical University Tongdao Road, Hohhot 10050, Inner Mongolia, P.R. China

Xueying Qiao

Department of Radiation Oncology, the Fourth Hospital of Hebei Medical University, Shijiazhuang, 050011, China.

Chun Han

Department of Radiation Oncology, the Fourth Hospital of Hebei Medical University, Shijiazhuang, 050011, China.

Shuchai Zhu

Department of Radiation Oncology, the Fourth Hospital of Hebei Medical University, Shijiazhuang, 050011, China.

Ruohui Zhang

Department of Radiation Oncology, the Fourth Hospital of Hebei Medical University, Shijiazhuang, 050011, China.

Junqiang Chen

Department of Radiation Oncology, Fujian Cancer Hospital/Fujian Medical University Cancer Hospital, Fuzhou, 350014, China

Cairong Hu

Department of Radiation Oncology, Fujian Cancer Hospital/Fujian Medical University Cancer Hospital, Fuzhou, 350014, China

Fuquan Zhang

Department of Radiation Oncology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, No. 1 Shuaifuyuan, Wangfujing Dongcheng District, Beijing, China

Xiaorong Hou

Department of Radiation Oncology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, No. 1 Shuaifuyuan, Wangfujing Dongcheng District, Beijing, China

Qingsong Pang

Department of Radiation Oncology, Tianjin Medical University Cancer Institute and Hospital/ National Clinical Research Center for Cancer, Tianjin, 300060, China.

Wencheng Zhang

Department of Radiation Oncology, Tianjin Medical University Cancer Institute and Hospital/ National Clinical Research Center for Cancer, Tianjin, 300060, China.

Gaofeng Li

Department of Radiation Oncology, Beijing Hospital, National Center of Gerontology, Beijing, 100730, China

Hailei Lin

Department of Radiation Oncology, Beijing Hospital, National Center of Gerontology, Beijing, 100730, China

Xinchen Sun

Department of Radiation Oncology, the First Affiliated Hospital of Nanjing Medical University, Nanjing, 210029, China.

Xiaolin Ge

Department of Radiation Oncology, the First Affiliated Hospital of Nanjing Medical University, Nanjing, 210029, China.

Caihong Li

Department of Radiation Oncology, the First Affiliated Hospital of Nanjing Medical University, Nanjing, 210029, China.

Hong Ge

Department of Radiation Oncology, Affiliated Cancer Hospital of Zhengzhou University, Zhengzhou, Henan 450008, P.R. China.

Dingjie Li

Department of Radiation Oncology, Affiliated Cancer Hospital of Zhengzhou University, Zhengzhou, Henan 450008, P.R. China.

Yadi Wang

Department of Radiation Oncology, the Seventh Medical Center of PLA General Hospital, Beijing 100700, China.

Na Lu

Department of Radiation Oncology, the Seventh Medical Center of PLA General Hospital, Beijing 100700, China.

Xianshu Gao

Department of Radiation Oncology, Peking University First Hospital, Peking University, Beijing, China.

Shangbin Qin

Department of Radiation Oncology, Peking University First Hospital, Peking University, Beijing, China.

Yuan Tian

Department of Radiation Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100021, China.

Zefen Xiao (✉ xiaozefen@sina.com)

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100021, China <https://orcid.org/0000-0003-0503-6814>

Research

Keywords: Cancer of the esophagus, Radiation therapy, Target volume delineation, Interobserver variability

Posted Date: October 16th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-35873/v3>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Purpose: To investigate the interobserver variability (IOV) in target volume delineation of definitive radiotherapy for thoracic esophageal cancer (TEC) among cancer centers in China, and ultimately improve contouring consistency as much as possible to lay the foundation for multi-center prospective studies.

Methods: Sixteen cancer centers throughout China participated in this study. In Phase 1, three suitable cases with upper, middle, and lower TEC were chosen, and participants were asked to contour a group of gross tumor volume (GTV-T), nodal gross tumor volume (GTV-N) and clinical target volume (CTV) for each case based on their routine experience. In Phase 2, the same clinicians were instructed to follow a contouring protocol to re-contour another group of target volume. The variation of the target volume was analyzed and quantified using dice similarity coefficient (DSC).

Results: Sixteen clinicians provided routine volumes, whereas ten provided both routine and protocol volumes for each case. The IOV of routine GTV-N was the most striking in all cases, with the smallest DSC of 0.37 (95% CI: 0.32-0.42), followed by CTV, whereas GTV-T showed high consistency. After following the protocol, the smallest DSC of GTV-N was improved to 0.64 (95% CI: 0.45-0.83, P=0.005) but the DSC of GTV-T and CTV remained constant in most cases.

Conclusion: Variability in target volume delineation was observed, but it could be significantly reduced and controlled using mandatory interventions.

Introduction

Definitive radiotherapy (dRT) concurrent with chemotherapy has been recognized as standard treatment for patients with locally advanced or unresectable thoracic esophageal cancer¹, and accurate target volume delineation was a prerequisite for three-dimensional conformal and intensity-modulated radiotherapy (IMRT) techniques, especially when using simultaneous-integrated boost (SIB) radiotherapy to deliver a boost dose to the gross tumor volume (GTV-T) and nodal gross tumor volume (GTV-N)^{2,3}. In 1998, Tai et al.⁴ observed interobserver variability (IOV) of target volume delineation in cervical esophageal cancer among 48 radiation oncologists, and the same team further discovered that the variation could be controlled with the help of special training⁵. However, the delineation variation in dRT for thoracic esophageal cancer has not been evaluated.

Traditionally, definitive radiotherapy (dRT) field borders for esophageal cancer were designated by 3–5-cm expansions proximally and distally beyond the primary lesion along the esophagus, based on 2-dimensional planning^{6,7}. Recently, based on the intensity modulated radiation therapy (IMRT) technique, an expert consensus on contouring guidelines⁸ compiled by radiation oncologists from cancer centers throughout the United States was published, which recommends that the CTV should include the GTV and GTV-N with at least 1-cm margin in all directions. In China, there are presently no consensus

reference contouring guidelines, and hence, some variance in dRT field is likely among different cancer centers. An investigation to address IOV in target volume delineation seemed appropriate, as the IOV appeared to have an impact on clinical outcomes in multi-center studies and could potentially be minimized with refined consensus guidelines^{9,10}.

This study aimed to investigate the IOV in target volume delineation in dRT for thoracic esophageal cancer among cancer centers in China, and ultimately improve contouring consistency as much as possible to lay the foundation for the multi-center prospective study.

Materials And Methods

Patients

The following clinical examinations of three cases were completed in the primary center: The barium meal films and esophagogastroduodenoscopy (EGD) helped in locating the site and length of the tumor; further, endoscopic ultrasound (EUS) and computed tomography (CT) were mainly used to determine invasive depth and the relationship of surrounding tissues. Besides, nodal status was comprehensively judged by EUS, CT, and 18-fluorodeoxyglucose-positron emission tomography computed tomography (PET/CT). Brain magnetic resonance imaging and PET/CT were performed to exclude distant metastasis.

Case 1: The primary lesion was in the upper thoracic esophagus, and its boundary to the surrounding tissue was unclear with suspicion of tracheal invasion. Suspicious lymph nodes were in Station 1R, 1L, 2L, 4L, and 5¹¹.

Case 2: The primary lesion was in the middle thoracic esophagus with a limited range of suspected lymphatic metastasis in Station 2R and 4R.

Case 3: The primary lesion was in the lower thoracic esophagus with a wide range of lymphatic metastasis. Suspicious lymph nodes were in Station 2R, 8L, and near the course of the left gastric artery.

Study layout

The invited radiation oncologists from sixteen cancer centers were members of the Jing-Jin-Ji Esophageal and Esophagogastric Cancer Radiotherapy Oncology Group (3JECROG).

A flow chart giving an overview of the study is shown in Figure 1. In Phase 1, all branch centers received patient history, clinical examinations and planning CT fused with planning PET (slice thickness: 3.0 mm), and heads of the radiotherapy department were asked to identify their specialists in thoracic oncology to delineate the first group of GTV-T, GTV-N, and clinical target volume (CTV) based on their own routine experience, which was sent back to the primary center after completion and recorded as the routine group (RG). In Phase 2, differences and consistency of these target volumes between the centers were fully discussed at the second 3JECROG annual conference, and finally the contouring protocol was drafted and referential target volumes (RTVs) were drawn based on the expert opinions. Then, RTVs along with the contouring protocol (Appendix A) and an atlas for target volume delineation¹² were sent to each center. The same specialists were asked to and give their opinions on RTVs and follow the protocol to re-delineate the second set of target volumes, which was recorded as the protocol group (PG).

Contour analysis

We introduced the dice similarity coefficient (DSC)¹³ as a direct measure of the degree of target volume matching (Figure 2), which had the ability to comprehensively evaluate the similarity in both volume and location. The method was used to calculate the spatial overlap between RTVs and target volumes from branch-centers. The value of DSC varies from 0 (completely disjoined) to 1 (absolutely overlapped). DSC was defined as follows:

$$DSC = \frac{2 * V_{(RTVs \cap branch)}}{V_{(branch)} + V_{(RTVs)}}$$

where $V_{(RTVs)}$, $V_{(branch)}$, and $V_{(RTVs \cap branch)}$ are the volume of RTVs, target volumes from branch-centers, and their overlapping region, respectively.

Statistical analysis

The Shapiro-Wilk test¹⁴ was used to check the distribution of continuous data for normality. The intergroup differences with normal distribution were evaluated using the paired t test; the ones with skewed distribution were evaluated using the rank test. All tests were two-sided, and a $P < 0.05$ was considered to indicate statistical significance. All statistical analyses were performed using R, version 3.5.1 (<http://www.r-project.org/>).

Results

Number of datasets received

A total of 16 datasets was retrieved from 15 branch-centers in Phase 1, and one of the branch-centers included two radiation oncologists delineating target volumes separately. CTVs delineation in the RG was presented in Figure 3A. In all three cases, the RG and PG were available from 10 clinicians; however, one of them did not comply with the protocol for the second delineation. Three participating clinicians returned their agreement on RTVs instead of contouring the new one, and the other three clinicians did not submit the protocol group of target volumes before the deadline. Figure 3B shows the CTVs in the PG, and a total of nine pairs of target volumes were included into the following comparison analysis.

Interobserver variability

According to the RG, the result of IOV in routine clinical practice was shown in Table 1. The maximum volume of CTV in case 3 was nearly seven times that of the smallest (range, 95.9–652.9 cc) volume. In general, GTV-T showed a higher degree of consistency, of which the DSC >0.75 in all three cases. In contrast, variability in GTV-N was larger, of which the DSC <0.55.

Efficiency of protocol

Detailed results of the comparison between the paired groups are presented in Table 2. The use of protocol had almost improved the DSC of all target volumes, and the most significant improvement was in GTV-N, which increased from 0.51, 0.38, and 0.37 to 0.67 ($P=0.022$), 0.55($P=0.260$), and 0.72 ($p=0.005$) in case 1, case 2, and case 3, respectively. In addition, it could be observed that the CTV of case 3 had a significantly better consistency with its DSC increasing from 0.63 to 0.72 ($P=0.004$).

Discussion

In the era of precision radiotherapy, the accuracy of target volume delineation plays a significant role in planning and execution of radiotherapy. However, owing to variance in the location of primary lesions and the range of lymph node metastasis, there is variability among radiation oncologists and radiation centers with respect to the target volumes of dRT, which may lead to delineation bias in multicenter research studies. Therefore, it is important to ensure the consistency of target volume delineation before

conducting a prospective, multi-center study. Our study found that IOV existed in routine delineating practice, and the contouring protocol could help in improving the contouring consistency.

According to the RG data, the consistency in delineation of GTV-T is generally high with basic DSC above 0.75; no obvious IOV existed among clinicians and centers regardless of the location of the primary lesion. Similar results were reported by a QA program of PRODIGE 26/CONCORDE phase 2/3 trial¹⁵ that the GTV delineation was almost respected in all centers. As also reported by Nowee et al.¹⁶, GTV delineation consistency seemed difficult to further improved.

For GTV-N, although PET-CT, EUS, and other auxiliary examinations were provided to help diagnose metastatic lymph nodes, and the contouring protocol was applied to improve its consistency in diagnosis, its DSC value was generally lower than 0.70. The possible reasons for these results are: first, there is no clear standardized definition of metastatic lymph nodes in esophageal cancer, which may have resulted in instances of either missed or over contoured GTV-N. Second, some clinical studies^{17–19} have shown that a large proportion of metastatic lymph nodes diagnosed by preoperative imaging is clinically over- or under-estimated compared with the postoperative pathological results. As reported by Mantziari et al.²⁰ in a study of 193 patients with esophageal cancer (clinical stage: T3N0), though the patients were enrolled into the single surgery group, pathological N0 cases accounted for only 35.8%, which indicated that more than 60% cases were under-estimated in the clinical assessment. Finally, according to the analysis by Gockel²¹, there was a 27–55% rate of lymph node metastasis when the primary lesion invades the submucosa. In addition, the lymph node metastasis in esophageal cancer is very extensive. For the cases receiving three-field lymph node dissection, as reported by Isono²², the rate of metastases in cervical nodes was 27.5% among patients with middle thoracic esophageal cancer. Therefore, it is challenging to assess lymph node metastasis in clinical practice. We reviewed the multi-center target volumes and found that variance in GTV-N is mainly due to the first reason, that is, clinicians' cautious overestimation of suspected metastasis in lymph nodes. In reality, the introduction of protocol allows clinicians to comprehensively combine multiple diagnostic methods for judgment, which may be the reason for the increased consistency of GTV-N.

The CTV field mainly relies on clinical examinations that mainly serves to provide a reference for clinicians by improving the accuracy of judgment of metastatic lymph nodes and determination of the range of radiation treatment. In the RG, IOV in CTV were observed among branch-centers. However, for those cases with relatively limited and proven range of lymph node metastasis, the IOV was relatively small regardless of the study group, which indicates that radiation oncologists reached an agreement in delineation of such target volumes. However, in case 3, with relatively more extensive lymph node metastasis, the IOV in target volume delineation becomes an issue. The efficacy of involved field irradiation (IFI) versus elective nodal irradiation (ENI) is still debatable^{23–25}, and a meta-analysis²⁶ shows that there is no survival difference between IFI and ENI. Thus, more prospective, comparative studies should be conducted for validation. Our study suggests that regardless of the location of the primary lesion, the consistency of delineating CTV was significantly improved according to the

requirements of the protocol. Therefore, the findings of this multi-center study are important because it emphasizes that a different center could achieve a more consistent target volume delineation.

A similar observation has been documented for other tumors such as nasopharyngeal, cervical, pulmonary, and gastric carcinomas^{27–30}. Factors accounting for the variance in GTV-N and CTV definition in this study were similar to those found by Weiss et al.,³¹ who suggested that causes are multifactorial, including image- and observer-related factors. A previous study¹⁰ suggested that refined contouring guidelines should be provided to better reduce the IOV. Accordingly, the present protocol proposed a consensus on involved lymph nodes and strict definition in the expansion criteria of CTV, leading to higher consistency in delineation of GTV-N and CTV. According to an investigation on head and neck cancers by Peters et al.³², protocol compliance did improve the radiotherapy quality assurance to achieve optimal treatment outcomes in the combined modality (chemoradiotherapy) treatment.

To our best knowledge, this is a minority of multicenter study of the variability in the target volumes of dRT for thoracic esophageal cancer, including 16 centers throughout China. Besides, we investigated IOV with respect to both volume and spatial relationship. In addition, unlike the previous trials that included dummy runs for QA analysis^{33,34}, our study showed, similar to Spoelstra et al.,²⁸ that IOV both before and after using the protocol was used to evaluate its efficiency.

Our study explored the variability in target volumes of dRT for thoracic esophageal cancer and compensated for the gap in this field. Furthermore, our study enforced that a contouring protocol could contribute to the consistency of target volumes, making the results of multi-center studies more reliable. Except the protocol, the improvement in the contouring consistency depends on advances in diagnostics. The department of imaging diagnosis in our center indicated that combining both the lymph node size and axial ratio relationship could improve the sensitivity in diagnosis³⁵. Besides, the addition of PET-CT and EUS will further increase the accuracy of N-stage^{36,37}, thereby improving the consistency of target volumes definition.

One limitation of our study is that the results were based on the combination of multiple modality examinations, while patients could only receive several of them in clinical practice, which could result in more striking contouring variance. In addition, we did not ask participating centers to design treatment plans for their target volumes, and therefore variances in dosimetric parameters could not be evaluated. Instead, we planned to further assess the variability in treatment plans designed for dRT of esophageal cancer and evaluate the impact of the dose-restriction protocol on dose–volume histogram parameters.

Conclusion

The IOV was observed in target volume delineation, and no available uniform consensus may account for it, which likely illustrates the different contouring philosophies of the participating centers and emphasizes the need for standardization. The consistency of target volumes delineation in different

centers could be improved through mandatory procedure, to lay a solid foundation for the reliability of multi-center prospective studies.

Abbreviations

dRT: Definitive Radiotherapy; IMRT: intensity-modulated radiation therapy; SIB: simultaneously integrated boost; GTV-T: Gross Tumor Volume; GTV-N: Metastatic Regional Nodes; IOV: Interobserver Variability; OAR: organ at risk; EGD: esophagogastroduodenoscopy; EUS: endoscopic ultrasound; CT: computed tomography; PET: Positron Emission Tomography; CTV: Clinical Target Volume; RG: routine group; RTV :referential target volume; PG: protocol group; DSC: dice similarity coefficient; IFI: involved field irradiation; ENI: elective nodal irradiation;

Declarations

Ethics approval and consent to participate:

The study conformed to the Declaration of Helsinki and was approved by the institutional ethics committee (National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital). Informed consent was obtained with the opt-out method.

Consent for publication:

Our study did not contain any individual person's data in any form.

Data availability:

All data generated or analyzed during this study are included in this published article and its supplementary information data.

Competing interests:

The authors declare no competing interests.

Funding:

This work was supported by the Beijing Hope Run Special Fund of Cancer Foundation of China (LC2016L04).

Authors' contributions:

Z.X. participated in design of the study and revised the article critically; Y.T. participated in design of the study, coordinated study procedure and revised the article critically; X.C. participated in statistical analyses and draft the manuscript; W.D., X.W., Z.Z., J.Y., W.G., M.L., X.Q., L.L., K.Z., M.Z., Y.S., K.L., Y.Z., H.W., Z.Y., J.Z., L.W., X.Q., C.H., S.Z., R.Z., J.C., C.H, F.Z., X.H., Q.P., W.Z., G.L., H.L., X.S., X.G., C.L, H.G., D. L.,

Y.W., N.L., X.G., S.Q. participated in the procedure of the target volumes delineation, data transmission or review of their own cancer centers, respectively.

Acknowledgements:

The authors thank all participating centers for their effort to send their target volumes for comparison with other centers.

References

1. Cooper JS, Guo MD, Herskovic A, Macdonald JS, Martenson JA, Al-Sarraf M *et al.* Chemoradiotherapy of locally advanced esophageal cancer: Long-term follow-up of a prospective randomized trial (RTOG 85-01). *J Am Med Assoc* 281, 1623–1627 (1999).
2. Chang JY, Gomez DR, Allen PK, Younes AI, Bhutani M, Komaki RU *et al.* Local Control and Toxicity of a Simultaneous Integrated Boost for Dose Escalation in Locally Advanced Esophageal Cancer: Interim Results from a Prospective Phase I/II Trial. *J Thorac Oncol* 12, 375–382 (2016).
3. Welsh J, Palmer MB, Ajani JA, Liao Z, Swisher SG, Hofstetter WL *et al.* Esophageal cancer dose escalation using a simultaneous integrated boost technique. *Int J Radiat Oncol Biol Phys* 82, 468–474 (2012).
4. Tai P, Van Dyk J, Yu E, Battista J, Stitt L, Coad T. Variability of target volume delineation in cervical esophageal cancer. *Int J Radiat Oncol Biol Phys* 42, 277–288 (1998).
5. Tai P, Van Dyk J, Battista J, Yu E, Stitt L, Tonita J *et al.* Improving the consistency in cervical esophageal target volume definition by special training. *Int J Radiat Oncol Biol Phys* 53, 766–774 (2002).
6. Herskovic A, Martz K, Al-Sarraf M, Leichman L, Brindle J, Vaitkevicius V *et al.* Combined Chemotherapy and Radiotherapy Compared with Radiotherapy Alone in Patients with Cancer of the Esophagus. *N Engl J Med* 326, 1593–1598 (1992).
7. Krasna MJ, Willett C, Goldberg R, Sugarbaker D, Tepper J, Hollis D *et al.* Phase III Trial of Trimodality Therapy With Cisplatin, Fluorouracil, Radiotherapy, and Surgery Compared With Surgery Alone for Esophageal Cancer: CALGB 9781. *J Clin Oncol* 26, 1086–1092 (2008).
8. Wu AJ, Bosch WR, Chang DT, Hong TS, Jabbour SK, Kleinberg LR *et al.* Expert consensus contouring guidelines for intensity modulated radiation therapy in esophageal and gastroesophageal junction cancer. *Int J Radiat Oncol Biol Phys* 92, 911–920 (2015).
9. Vinod SK, Jameson MG, Min M, Holloway LC. Uncertainties in volume delineation in radiation oncology: A systematic review and recommendations for future studies. *Radiother Oncol* 121, 169–179 (2016).
10. Joye I, Macq G, Vaes E, Roels S, Lambrecht M, Pelgrims A *et al.* Do refined consensus guidelines improve the uniformity of clinical target volume delineation for rectal cancer? Results of a national review project. *Radiother Oncol* 120, 202–206 (2016).

11. El-Sherief AH, Wu CC, Abbott GF, Drake RL, Rice TW, Lau CT. International Association for the Study of Lung Cancer (IASLC) Lymph Node Map: Radiologic Review with CT Illustration. *RadioGraphics* 34, 1680–1691 (2014).
12. Xiao Z, Zhou Z, Li Y. *Esophageal Cancer target volume delineation and treatment guidance for radiation therapy (in Chinese)*, 1st ed. Beijing:People's Medical Publishing House. (2017).
13. Dice LR. Measures of the Amount of Ecologic Association Between Species. *Ecology* 26, 297–302 (1945).
14. Royston P. Remark AS R94: A Remark on Algorithm AS 181: The W-test for Normality. *Appl Stat* 44, 547 (2006).
15. Boustani J, Rivin Del Campo E, Blanc J, Peiffert D, Benezery K, Pereira R et al. Quality Assurance of Dose-Escalated Radiation Therapy in a Randomized Trial for Locally Advanced Oesophageal cancer. *Int J Radiat Oncol Biol Phys* 105, 329–337 (2019).
16. Nowee ME, Voncken FEM, Kotte ANTJ, Goense L, van Rossum PSN, van Lier ALHMW et al. Gross tumour delineation on computed tomography and positron emission tomography-computed tomography in oesophageal cancer: A nationwide study. *Clin Transl Radiat Oncol* 14, 33–39 (2019).
17. Samson P, Puri V, Robinson C, Lockhart C, Carpenter D, Broderick S et al. Clinical T2N0 Esophageal Cancer: Identifying Pretreatment Characteristics Associated with Pathologic Upstaging and the Potential Role for Induction Therapy. *Ann Thorac Surg* 101, 2102–2111 (2016).
18. Speicher PJ, Ganapathi AM, Englum BR, Hartwig MG, Onaitis MW, D'Amico TA et al. Induction therapy does not improve survival for clinical stage T2N0 esophageal cancer. *J Thorac Oncol* 9, 1195–1201 (2014).
19. Zhang JQ, Hooker CM, Brock M V, Shin J, Lee S, How R et al. Neoadjuvant chemoradiation therapy is beneficial for clinical stage T2 N0 esophageal cancer patients due to inaccurate preoperative staging. *Ann Thorac Surg* 93, 429–437 (2012).
20. Mantziari S, Gronnier C, Renaud F, Duhamel A, Théreaux J, Brigand C et al. Survival Benefit of Neoadjuvant Treatment in Clinical T3N0M0 Esophageal Cancer: Results From a Retrospective Multicenter European Study. *Ann Surg* 266, 805–813 (2017).
21. Gockel I, Sgourakis G, Lyros O, Polotzek U, Schimanski CC, Lang H et al. Risk of lymph node metastasis in submucosal esophageal cancer: A review of surgically resected patients. *Expert Rev Gastroenterol Hepatol* 5, 371–384 (2011).
22. Isono K, Sato H, Nakayama K. Results of a nationwide study on the three-field lymph node dissection of esophageal cancer. *Oncol* 48, 411–420 (1991).
23. Onozawa M, Nihei K, Ishikura S, Minashi K, Yano T, Muto M et al. Elective nodal irradiation (ENI) in definitive chemoradiotherapy (CRT) for squamous cell carcinoma of the thoracic esophagus. *Radiother Oncol* 92, 266–269 (2009).
24. Ji K, Zhao L, Yang C, Meng M, Wang P. Three-dimensional conformal radiation for esophageal squamous cell carcinoma with involved-field irradiation may deliver considerable doses of incidental nodal irradiation. *Radiat Oncol* 7, 1–8 (2012).

25. Yamashita H, Takenaka R, Omori M, Imae T, Okuma K, Ohtomo K *et al.* Involved-field radiotherapy (IFRT) versus elective nodal irradiation (ENI) in combination with concurrent chemotherapy for 239 esophageal cancers: A single institutional retrospective study. *Radiat Oncol* 10, 1–10 (2015).
26. Cheng YJ, Jing SW, Zhu LL, Wang J, Wang L, Liu Q *et al.* Comparison of elective nodal irradiation and involved-field irradiation in esophageal squamous cell carcinoma: a meta-analysis. *J Radiat Res* 59, 604–615 (2018).
27. Jansen EPM, Nijkamp J, Gubanski M, Lind PARM, Verheij M. Interobserver Variation of Clinical Target Volume Delineation in Gastric Cancer. *Int J Radiat Oncol Biol Phys* 77, 1166–1170 (2010).
28. Spoelstra FOB, Senan S, Le Péchoux C, Ishikura S, Casas F, Ball D *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, 1106–1113 (2010).
29. Eminowicz G, McCormack M. Variability of clinical target volume delineation for definitive radiotherapy in cervix cancer. *Radiother Oncol* 117, 542–547 (2015).
30. Peng Y lin, Chen L, Shen G zhu, Li Y ning, Yao J jin, Xiao W wei *et al.* Interobserver variations in the delineation of target volumes and organs at risk and their impact on dose distribution in intensity-modulated radiation therapy for nasopharyngeal carcinoma. *Oral Oncol* 82, 1–7 (2018).
31. Weiss E, Hess CF. The impact of gross tumor volume (GTV) and clinical target volume (CTV) definition on the total accuracy in radiotherapy: Theoretical aspects and practical experiences. *Strahlentherapie und Onkol* 179, 21–30 (2003).
32. Peters LJ, O'Sullivan B, Giralt J, Fitzgerald TJ, Trott A, Bernier J *et al.* Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: Results from TROG 02.02. *J Clin Oncol* 28, 2996–3001 (2010).
33. Foppiano F, Fiorino C, Frezza G, Greco C, Valdagni R. The impact of contouring uncertainty on rectal 3D dose-volume data: Results of a dummy run in a multicenter trial (AIROPROS01-02). *Int J Radiat Oncol Biol Phys* 57, 573–579 (2003).
34. Davis JB, Reiner B, Dusserre A, Giraud JY, Bolla M. Quality assurance of the EORTC trial 22911. A phase III study of post-operative external radiotherapy in pathological stage T3N0 prostatic carcinoma: The dummy run. *Radiother Oncol* 64, 65–73 (2002).
35. Liu J, Wang Z, Shao H, Qu D, Liu J, Yao L. Improving CT detection sensitivity for nodal metastases in oesophageal cancer with combination of smaller size and lymph node axial ratio. *Eur Radiol* 28, 188–195 (2018).
36. Choi J, Kim SG, Kim JS, Jung HC, Song IS. Comparison of endoscopic ultrasonography (EUS), positron emission tomography (PET), and computed tomography (CT) in the preoperative locoregional staging of resectable esophageal cancer. *Surg Endosc* 24, 1380–1386 (2010).
37. Tan R, Yao SZ, Huang ZQ, Li J, Li X, Tan HH *et al.* Combination of FDG PET/CT and contrast-enhanced MSCT in detecting lymph node metastasis of esophageal cancer. *Asian Pacific J Cancer Prev* 15, 7719–7724 (2014).

Tables

Table 1
Interobserver variability (IOV) among 16 centers in routine clinical practice

	Volume (cc)	Dice similarity coefficient
	Mean (smallest-largest)	Mean (95% CI)
Case1		
GTV-T	101.2(74.7-127.7)	0.86(0.83-0.88)
GTV-N	12.9(4.3-30.5)	0.50(0.41-0.58)
CTV	333.8(204.4-442.0)	0.80(0.76-0.82)
Case 2		
GTV-T	27.3(21.8-34.2)	0.81(0.79-0.83)
GTV-N*	2.2(0-19.4)	0.26(0.08-0.43)
CTV	179.8(94.7-307.5)	0.74(0.68-0.79)
Case 3		
GTV-T	37.9(30.1-50.8)	0.79(0.76-0.82)
GTV-N*	11.8(5.7-32.3)*	0.39(0.34-0.43)
CTV	367.6(95.9-652.9)	0.65(0.57-0.73)
CI, confidence interval; *GTV-Ns do not conform to normal distribution as per the Shapiro-Wilk test, so the median is used instead of the mean.		

Table 2
Interobserver variation between the routine
and protocol groups from nine clinicians

	Dice similarity coefficient (95% CI)		
	Routine	Protocol	P
Case 1			
GTV-T	0.86(0.82–0.89)	0.91(0.86–0.96)	.013
GTV-N	0.51(0.39–0.62)	0.67(0.53–0.80)	.022
CTV	0.83(0.79–0.86)	0.87(0.84–0.89)	.077
Case 2			
GTV-T	0.81(0.79–0.84)	0.83(0.78–0.89)	.130
GTV-N*	0.38(0.18–0.59)	0.55(0.39–0.71)	.260
CTV	0.74(0.66–0.83)	0.79(0.75–0.84)	.084
Case 3			
GTV-T	0.79(0.75–0.84)	0.84(0.78–0.91)	.190
GTV-N	0.37(0.32–0.42)	0.64(0.45–0.83)	.005
CTV	0.63(0.53–0.74)	0.72(0.62–0.81)	.004
CI, confidence interval.			
*GTV-N paired Mann–Whitney U test			
[]			

Figures

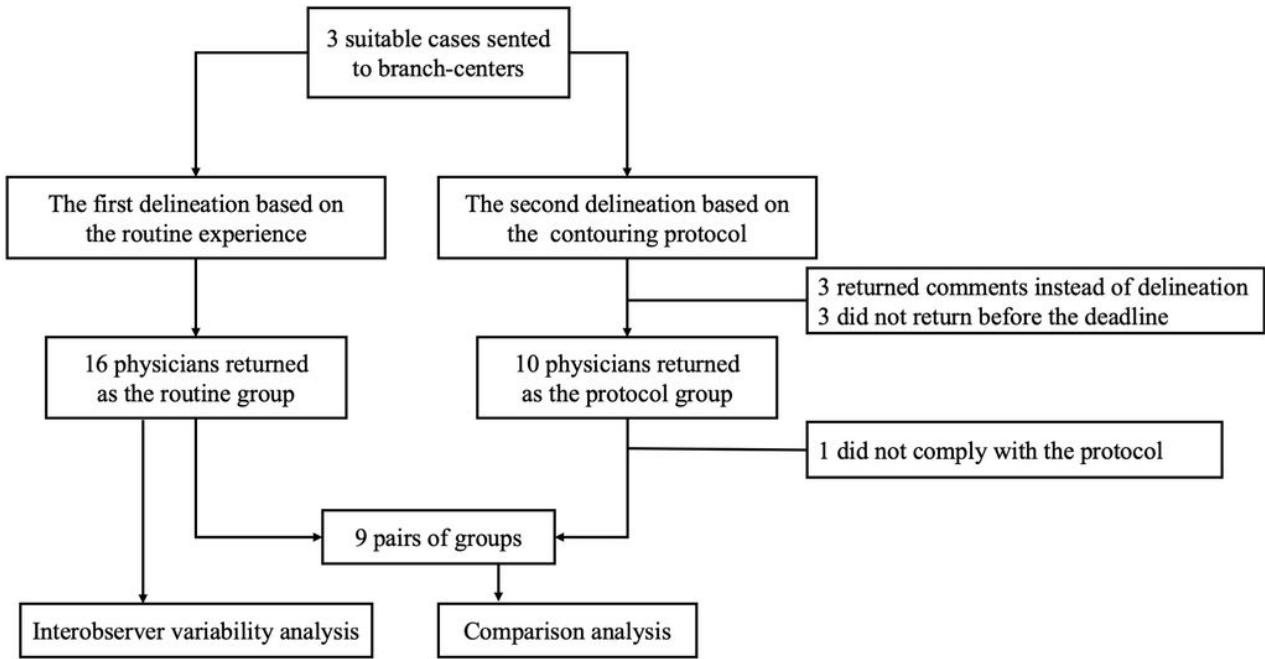
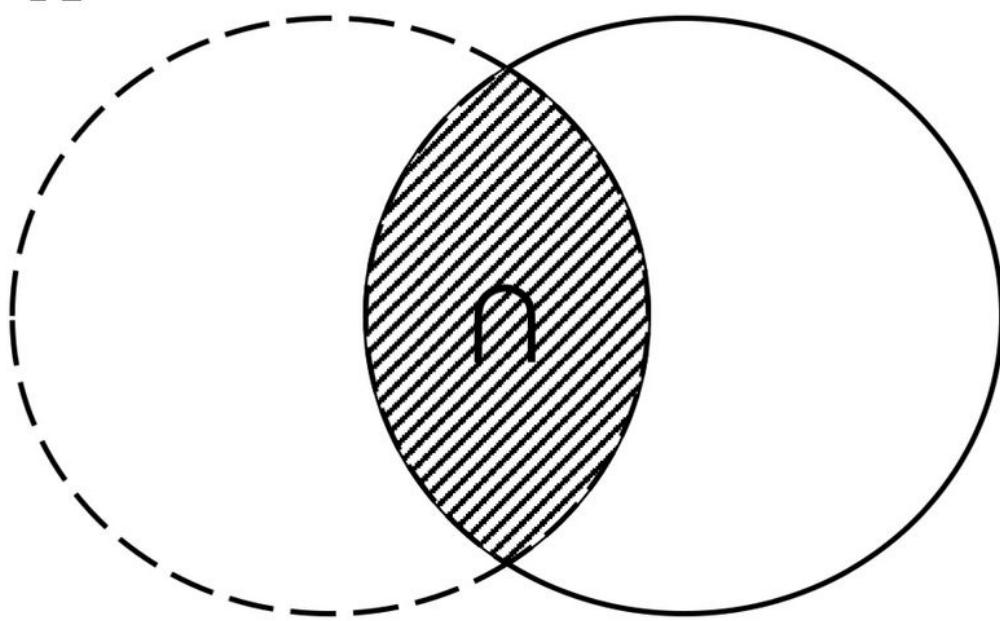


Figure 1

Flow chart of the study

A



B

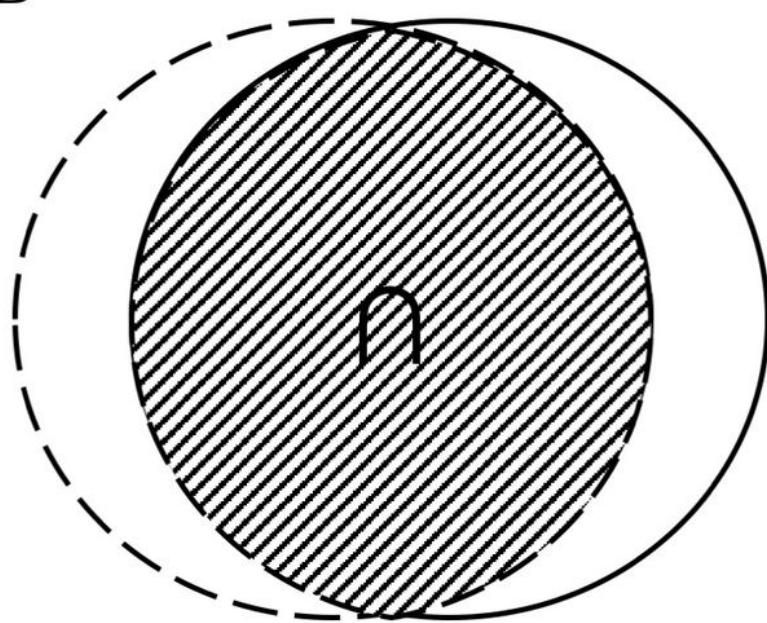


Figure 2

Contouring variability in spatial location evaluated by dice similarity coefficient (DSC). Sample “ n ” with gray region represents overlapping region. Panels A and B are examples of poor and good spatial consistency, respectively.

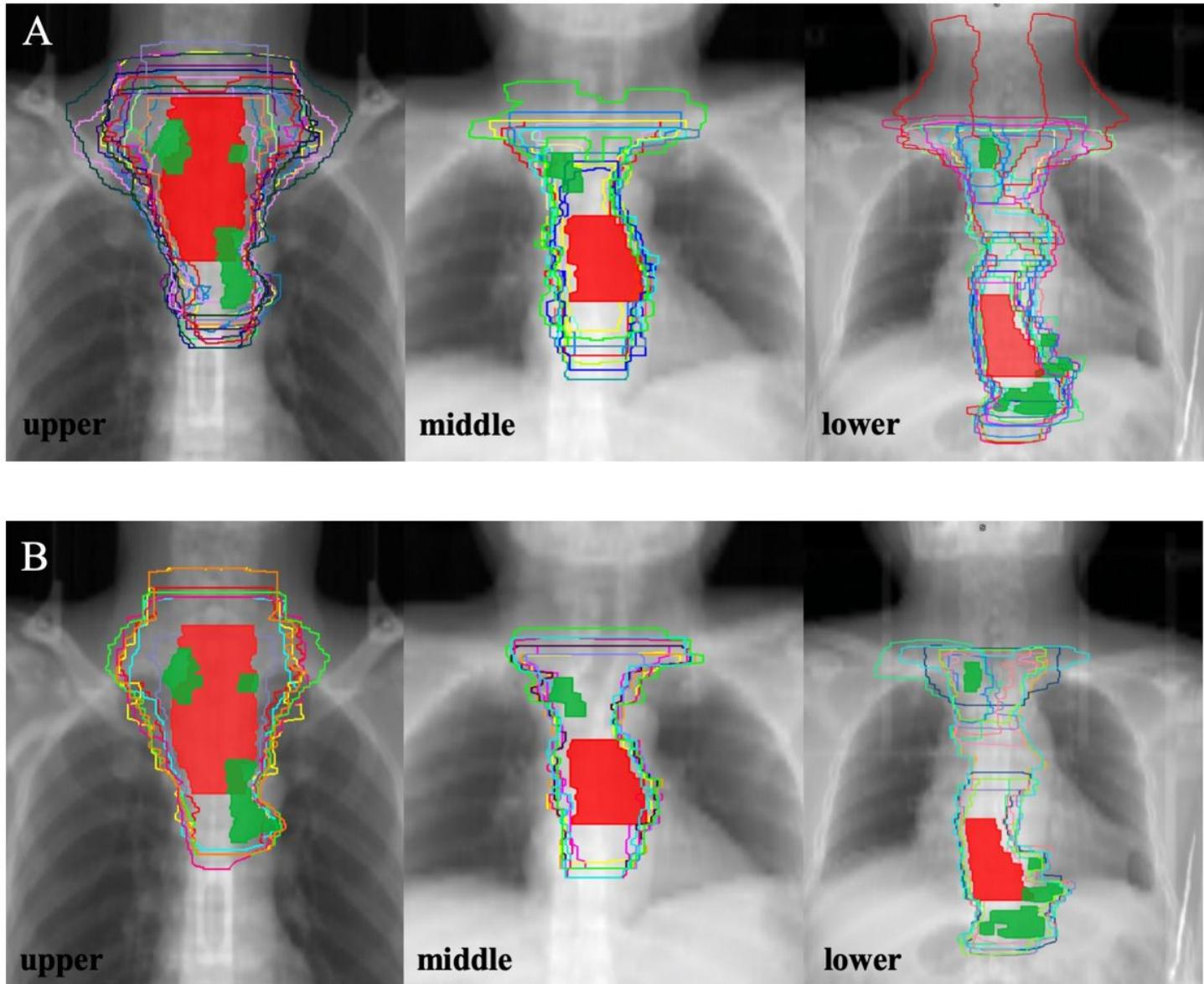


Figure 3

The routine group (RG; panel A) of clinical target volumes (CTVs) from 16 clinicians and the protocol group (PG; panel B) of CTVs from nine clinicians projected on one digitally reconstructed radiograph (DDR) of a CT dataset. Red and green areas indicate the primary tumor and lymph nodes, respectively.

Supplementary Files

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

- [AppendixA.doc](#)