



# Inter-Observer and Intra-Observer Variability in Gross Tumor Volume Delineation of Primary Esophageal Carcinomas Based on Different Combinations of Diagnostic Multimodal Images

## OPEN ACCESS

### Edited by:

Tao Li,  
Sichuan Cancer Hospital, China

### Reviewed by:

Signe Friesland,  
Karolinska University Hospital,  
Sweden  
Diane Schott,  
University of Nebraska Medical Center,  
United States

### \*Correspondence:

Fengxiang Li  
lifengxiangli@aliyun.com  
Jianbin Li  
lijianbin@msn.com

<sup>†</sup>These authors share first authorship

### Specialty section:

This article was submitted to  
Radiation Oncology,  
a section of the journal  
Frontiers in Oncology

**Received:** 18 November 2021

**Accepted:** 28 February 2022

**Published:** 01 April 2022

### Citation:

Li F, Li Y, Wang X, Zhang Y,  
Liu X, Liu S, Wang W, Wang J,  
Guo Y, Xu M and Li J (2022) Inter-  
Observer and Intra-Observer Variability  
in Gross Tumor Volume Delineation  
of Primary Esophageal Carcinomas  
Based on Different Combinations  
of Diagnostic Multimodal Images.  
*Front. Oncol.* 12:817413.  
doi: 10.3389/fonc.2022.817413

Fengxiang Li<sup>1\*†</sup>, Yankang Li<sup>1†</sup>, Xue Wang<sup>2</sup>, Yingjie Zhang<sup>1</sup>, Xijun Liu<sup>1</sup>, Shanshan Liu<sup>1</sup>,  
Wei Wang<sup>1</sup>, Jinzhi Wang<sup>1</sup>, Yanluan Guo<sup>1</sup>, Min Xu<sup>1</sup> and Jianbin Li<sup>1\*</sup>

<sup>1</sup> Department of Radiation Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China, <sup>2</sup> Department of Radiation Oncology, Linyi Cancer Hospital, Linyi, China

**Background and Purpose:** This study aimed to investigate inter-/intra-observer delineation variability in GTVs of primary esophageal carcinomas (ECs) based on planning CT with reference to different combinations of diagnostic multimodal images from endoscopy/EUS, esophagography and FDG-PET/CT.

**Materials and Methods:** Fifty patients with pathologically proven thoracic EC who underwent diagnostic multimodal images before concurrent chemoradiotherapy were enrolled. Five radiation oncologist independently delineated the GTVs based on planning CT only (GTV<sub>C</sub>), CT combined with endoscopy/EUS (GTV<sub>CE</sub>), CT combined with endoscopy/EUS and esophagography (X-ray) (GTV<sub>CEX</sub>), and CT combined with endoscopy/EUS, esophagography, and FDG-PET/CT (GTV<sub>CEXP</sub>). The intra-/inter-observer variability in the volume, longitudinal length, generalized CI (CI<sub>gen</sub>), and position of the GTVs were assessed.

**Results:** The intra-/inter-observer variability in the volume and longitudinal length of the GTVs showed no significant differences ( $p > 0.05$ ). The mean intra-observer CI<sub>gen</sub> values for all observers was  $0.73 \pm 0.15$ . The mean inter-observer CI<sub>gen</sub> values for the four multimodal image combinations was  $0.67 \pm 0.11$ . The inter-observer CI<sub>gen</sub> for the four combined images was the largest, showing significant differences with those for the other three combinations. The intra-observer CI<sub>gen</sub> among different observers and inter-observer CI<sub>gen</sub> among different combinations of multimodal images showed significant differences ( $p < 0.001$ ). The intra-observer CI<sub>gen</sub> for the senior radiotherapists was larger than that for the junior radiotherapists ( $p < 0.001$ ).

**Conclusion:** For radiation oncologists with advanced medical imaging training and clinical experience, using diagnostic multimodal images from endoscopy/EUS, esophagography, and FDG-PET/CT could reduce the intra-/inter-observer variability and increase the accuracy of target delineation in primary esophageal carcinomas.

**Keywords:** esophageal carcinoma, diagnostic multimodal images, target delineation, intra-observer variability, inter-observer variability

## HIGHLIGHTS

- There is large variability in target volume delineation for esophageal carcinoma.
- Evaluation of inter-/intra-observer delineation variability based on diagnostic multimodal imaging.
- Multimodal diagnostic image combinations can reduce the intra-/inter-observer variability and increase delineation accuracy.

## INTRODUCTION

Esophageal carcinoma is the seventh most commonly diagnosed cancer and the third leading cause of cancer deaths worldwide in 2018 (1). Preoperative and definitive chemoradiation therapies have played a key role in the treatment of esophageal carcinoma (2–5). The proportion of residual esophageal carcinoma after chemoradiation is significantly correlated with locoregional and distant failure (6–9). Reliable delineation of the target gross tumor volume (GTV) is required for accurate radiation dose delivery and successful radiation therapy (10, 11). There is generally large variability in the target volume delineation for esophageal carcinoma, which might be primarily derived from the geometric uncertainties of different images and inherent variability among different observers based on the studies on other malignancies (12, 13).

Conventional three-dimensional CT (3DCT) has been the workhorse modality used to delineate the esophageal tumor target volume. However, it is difficult to determine the proximal and distal extension of tumors and differentiate the layers of the esophageal wall (14–16). An esophagography has shown a higher accuracy in assessing the tumor length (59% of cases, compared with 32% with CT), with tumor morphology influencing the accuracy (14–16). Although endoscopy and endoscopic ultrasonography (EUS) might present the tumor length more accurately (17, 18), it is difficult to transform the imaging to radiotherapy (RT) planning (19). Fluorine-18 fluorodeoxyglucose positron emission tomography/computed tomography (18F-FDG-PET/CT) has proved useful for diagnosing and staging esophageal carcinoma. However, there is limited evidence supporting the validity of FDG-PET/CT for target volume delineation (20, 21). The false-positive FDG uptake in areas of inflammation reduces the specificity of tumor extent visualization (22). Therefore, the combination of multimodal images is critical for determining the GTV of esophageal cancer

(EC) accurately. Several studies have focused on the inter-observer variability of target volume delineation in FDG-PET/CT compared with pure CT imaging (10, 23). As CT imaging has proved indispensable for the visualization/detection of esophageal tumors, the use of multimodality imaging including esophagography, endoscopy/EUS and FDG-PET/CT for target volume delineation has not received sufficient attention.

In general, patients scheduled to receive radiotherapy or chemoradiotherapy undergo diagnostic multimodal imaging including enhanced CT, endoscopy/EUS, esophagography, or FDG-PET/CT. In clinical practice, radiation oncologists generally delineate the target volumes based on the planning CT images, with reference to various preexisting diagnostic images. However, the outcome of using different combinations of diagnostic multimodal images on the inter-observer and intra-observer delineation variability remains unclear. The purpose of this study was to investigate the inter-observer and intra-observer delineation variability in the GTVs of primary esophageal tumors with reference to different combinations of multimodal images from endoscopy/EUS, esophagography, and FDG-PET/CT. This study indicated the influence of the addition of different multimodal images on the GTVs delineation variability, which may contribute to making clinical decision on acquire different multimodal images.

## MATERIALS AND METHODS

### Patient Selection and Characteristics

This study was approved by the institutional research ethics board and informed consent has been obtained from the participants involved. Fifty-one patients with pathologically proven thoracic EC who had undergone preoperative or definitive concurrent chemoradiotherapy between May 2015 and June 2017 at the institutional hospital were enrolled. Among the selected patients, there were seventeen cases each of upper, middle, and lower EC. One patient with lower EC was excluded due to the lack of PET-CT imaging data. All patients underwent a diagnostic imaging examination that included an endoscopy/EUS, esophagography, and FDG-PET/CT before receiving chemoradiotherapy. The average time for acquiring the diagnostic images was within the two-week period before chemoradiotherapy. **Table 1** presents the patient characteristics.

### Multimodal Imaging

**Endoscopy/EUS examination:** All patients underwent diagnostic endoscopy examinations using an electronic gastroscope

**TABLE 1 |** Patient's characteristics.

Characteristics	Number
Sex, n (%)	
M	40 (80%)
F	10 (20%)
Age, median, y (range)	63 (44-88)
Tumor histology, n (%)	
Squamous cell carcinoma	50 (100%)
SUV <sub>max</sub> , mean, median, y (range)	17.1, 15.2 (2.8-49.5)
TNM* stage, n (%)	
T <sub>2</sub> N <sub>0-2</sub> M <sub>0</sub>	4 (8%)
T <sub>3</sub> N <sub>1-3</sub> M <sub>0-1</sub>	34 (68%)
T <sub>4a</sub> N <sub>0-2</sub> M <sub>0-1</sub>	12 (24%)
Tumor location, n (%)	
Upper	17 (34%)
Mid-	17 (34%)
Distal	16 (32%)

(Olympus GIF-Q260J) before treatment. Seven patients did not undergo EUS examinations due to esophageal stenosis. The ultrasonic probe (Olympus EVIS EUS EU-ME2) was inserted into the patient's esophagus along the track of the biopsy forceps to detect the depth of tumor infiltration in the esophageal wall and the extent of proximal and distal tumor infiltration. The distances from the proximal and distal ends of the tumor to the incisors were recorded.

**Esophagography (X-ray) image acquisition:** Esophagography was performed before treatment using a digital radiography machine (Siemens Luminos dRT Max). All barium examinations were performed under fasting conditions, followed by a standard protocol (drinking 200 ml of diluted barium, in the upright, supine, and prone positions, with and without the gas powder).

**PET/CT image acquisition:** The PET-CT scan was performed within the two-week period prior to the planning CT scan as a part of the routine diagnostic management for EC. An 18F-FDG PET/CT scan of the chest was performed with an integrated PET/CT system (Philips Gemini TF Big Bore). The PET images were reconstructed with the CT-derived attenuation correction using an ordered subset expectation maximization algorithm with post-reconstruction Gaussian filtering, with a full width at half maximum of 5 mm.

**Planning CT image acquisition:** During the simulation, all patients were immobilized using a thermoplastic mask in the supine position with the arms placed along the side of the body. Each patient underwent an enhanced planning CT scan of the thoracic region on a 16-slice CT scanner (Philips Brilliance Bore CT) under free-breathing conditions. The planning CT images were reconstructed using a thickness of 3 mm and subsequently transferred to an Eclipse treatment planning system (Varian Eclipse 11).

## Target Volume Delineation

A treatment planning system (Eclipse; Varian Medical Systems, Inc., Palo Alto, CA, USA) was used to contour the GTVs of the primary EC. The visualization parameter for delineation included the mediastinal window set to +40/400 HU. Before contouring, some clinical information such as the physical examination, pathological findings, and diagnostic CT image

data were made available to the observers, while they were blind to the diagnostic endoscopy/EUS, esophagography, and FDG-PET/CT data. If the positive lymph nodes could not be separated from the primary tumor visually, they were delineated together with the primary tumor.

Five radiation oncologists (observers), who were blind to the diagnostic endoscopy/EUS, esophagography, and FDG-PET/CT patient data, were asked to independently delineate the GTVs with reference to different combinations of the multimodal images, including planning CT only (GTV<sub>C</sub>), CT combined with endoscopy/EUS (GTV<sub>CE</sub>), CT combined with endoscopy/EUS and esophagogram (X-ray) (GTV<sub>CEX</sub>), and CT combined with endoscopy/EUS, esophagogram, and FDG-PET/CT (GTV<sub>CEXP</sub>) (**Figure 1**). All observers were blind to the contours delineated by the other oncologists and their own former/previous contours. Observers 1 and 2 with clinical experience within five years were regarded as junior observers, while observers 3, 4, and 5 with more than ten years of clinical experience were regarded as senior observers. All contours were delineated in about two years. A delay of at least two months existed between each contouring of the tumor to eliminating a recall of the previous contouring for observers 1, 2, 3, and 5. The time interval for observer 4 was only one month, as the former observer 4 dropped out of the delineation process due to parturition.

## Inter-/Intra-Observer Variability Analysis

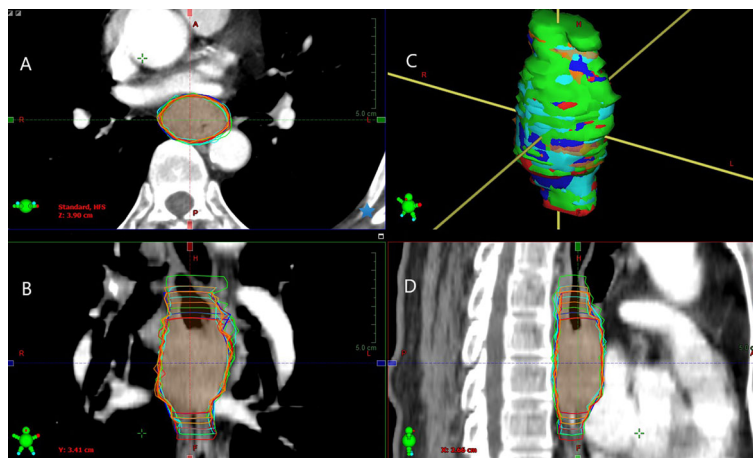
Inter-/intra-observer variability in the volume, longitudinal length, generalized conformity index (CI<sub>gen</sub>), and position of the GTVs was assessed. The intra-observer variability can be generally regarded as the variability of the same observer when re-contouring a single case. However, in this study, it is defined as the variability of the contours on the four multimodal imaging/image combinations for one observer (23).

The mean volume and longitudinal length of the GTVs based on different multimodal imaging combinations for different observers were calculated. The inter-observer variability in the volume and longitudinal length on different multimodal imaging combinations and the intra-observer variability for different observers were measured. The tumor length was measured using CT, endoscopy/EUS (43 cases), esophagography, and FDG-PET/CT, with the difference between the tumor length and corresponding longitudinal length of the GTVs subsequently evaluated.

Conformity index (CI) was defined as the ratio of the common volume to encompassing volume (13, 24). The generalized CI (CI<sub>gen</sub>) was used to assess the overall consistency of all volume combinations delineated by different observers on the same imaging-modality combination, and that delineated by the same observer on different imaging-modality combinations. The formula is given by (13, 25):

$$CI_{gen} = \frac{\sum_{pairsij} |A_i \cap A_j|}{\sum_{pairsij} A_i \cup A_j}$$

CI<sub>gen</sub> is a good parameter for revealing the difference in the volumes delineated based on the size, shape, and location (10, 23).



**FIGURE 1** | Example of GTVs delineated based on different combinations of multimodal images by observer 1 (green segment), observer 2 (red segment), observer 3 (blue segment), observer 4 (orange segment), and observer 5 (cyan segment) in transverse (A), frontal (B), surface (C), and sagittal (D) planes for one patient (Patient 5). Inter-/intra-observer variability in the volume and longitudinal length on different combinations of multimodal images exhibiting significant differences.

The use of  $CI_{gen}$  tends to decrease the bias in the number of delineations (13). The lower is the  $CI_{gen}$  value for the same imaging-modality combination, the greater is the inter-observer variability. Similarly, a lower  $CI_{gen}$  for the same observer suggests a greater intra-observer variability.

In addition, the x (right-left), y (anterior-posterior), and z (superior-inferior) axes of the center of mass (COM) of the volume were measured. The centroid shifts between the different volumes were then obtained. Finally, the three dimensional (3D) centroid shifts were calculated using the followed equation (24, 26):

$$3D \text{ centroid shifts} = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$$

## Statistical Analysis

Statistical analysis was performed using the SPSS software package (SPSS 25.0). All the data had an approximately normal distribution. The one-way ANOVA test was applied to detect the inter-/intra-observer variability in the volume, longitudinal length,  $CI_{gen}$ , and position of the GTVs among different observers and different multimodal imaging combinations. The paired *t*-test was used to compare the volume, longitudinal length,  $CI_{gen}$ , and position of the GTVs between two observers or two multimodal imaging combinations. A  $P < 0.05$  was considered significant.

## RESULTS

### GTV Volume

**Table 2** shows the primary GTV delineated based on four different multimodal imaging combinations for each observer. No significant inter-observer differences in the volume were observed for  $GTV_C$ ,  $GTV_{CE}$ ,  $GTV_{CEX}$ , or  $GTV_{CEXP}$  ( $p = 0.904$ ,

0.987, 0.984, and 0.97, respectively). The intra-observer variability in the volume of the GTVs derived from four different multimodal imaging combinations for observers 1–5 also showed no significant differences ( $p = 0.926, 0.997, 0.908, 0.943$ , and  $0.99$ , respectively). However, the paired comparisons indicated significant differences in the GTV volume between observers 1 and 2, observers 1 and 4, and observers 3 and 4 ( $t = 3.154, 6.368$ , and  $3.342$ ,  $p = 0.002, <0.001$ , and  $0.001$ , respectively). Approximate statistical differences in the GTV volume were found between observers 1 and 3, and observers 2 and 4 ( $t = 3.342$  and  $1.869$ ,  $p = 0.061$  and  $0.063$ , respectively).

### Esophageal Tumor Length

**Table 3** shows the mean tumor lengths measured by CT, endoscopy/EUS, esophagography, and FDG-PET/CT. No significant differences were found between any two image-based tumor lengths. **Table 3** presents the mean longitudinal lengths measured by the five observers corresponding to  $GTV_C$ ,  $GTV_{CE}$ ,  $GTV_{CEX}$ , and  $GTV_{CEXP}$ . The mean longitudinal length for  $GTV_{CEXP}$  was larger than the tumor length measured by FDG-PET/CT ( $p = 0.0035$ ). The intra-observer variability in the longitudinal length of the GTVs based on four multimodal imaging combinations for observers 1–5 showed no significant differences ( $p = 0.751, 0.794, 0.115, 0.962$ , and  $0.753$ , respectively). **Table 2** shows the tumor lengths measured based on the four different multimodal imaging combinations for each observer. No significant inter-observer differences in the longitudinal length were recorded for  $GTV_C$ ,  $GTV_{CE}$ ,  $GTV_{CEX}$ , and  $GTV_{CEXP}$  ( $p = 0.286, 0.503, 0.997$ , and  $0.749$ , respectively). The two-related-samples tests indicated significant differences in the longitudinal lengths of the four GTVs between observers 1 and 2 ( $t = 2.776$ ,  $p = 0.006$ ), observers 1 and 5 ( $t = 1.98$ ,  $p = 0.049$ ), observers 3 and 2 ( $t = -3.166$ ,  $p = 0.002$ ), and observers 3 and 5 ( $t = 2.992$ ,  $p = 0.003$ ).

**TABLE 2** | The volume and longitudinal length of GTVs based on different combinations of multimodal imaging for different observers.

Parameter		Observer 1	Observer 2	Observer 3	Observer 4	Observer 5	Mean±SD
<b>GTV<sub>C</sub></b>	Volume(cm <sup>3</sup> )	37.57 ± 26.68	33.78 ± 27.42	36.42 ± 26.91	33.48 ± 28.07	33.28 ± 26.87	34.91 ± 27.19
	Length(mm)	5.7 ± 2.6	5.3 ± 2.5	6.3 ± 2.5	5.7 ± 2.6	5.3 ± 2.3	5.7 ± 2.5
<b>GTV<sub>CE</sub></b>	Volume(cm <sup>3</sup> )	34.06 ± 25.90	34.47 ± 26.65	36.33 ± 27.10	32.68 ± 25.79	35.13 ± 27.50	34.68 ± 26.82
	Length(mm)	5.7 ± 2.5	5.4 ± 2.2	6.3 ± 2.4	5.8 ± 2.4	5.7 ± 2.3	5.8 ± 2.4
<b>GTV<sub>CEX</sub></b>	Volume(cm <sup>3</sup> )	35.43 ± 25.79	35.00 ± 26.40	33.57 ± 26.68	34.50 ± 26.84	35.03 ± 28.37	34.34 ± 26.81
	Length(mm)	5.6 ± 2.2	5.6 ± 2.2	5.5 ± 2.3	5.6 ± 2.2	5.6 ± 2.5	5.6 ± 2.3
<b>GTV<sub>CEXP</sub></b>	Volume(cm <sup>3</sup> )	36.30 ± 27.41	34.17 ± 26.19	33.48 ± 27.13	33.28 ± 26.67	36.73 ± 28.27	35.04 ± 27.17
	Length(mm)	6.1 ± 2.5	5.7 ± 2.2	5.5 ± 2.3	5.8 ± 2.4	5.8 ± 2.3	5.8 ± 2.3
<b>Mean ± SD</b>	Volume(cm <sup>3</sup> )	35.84 ± 26.45	34.36 ± 26.67	34.95 ± 27.20	33.52 ± 26.91	35.04 ± 27.75	34.37 ± 27.29
	Length(mm)	5.8 ± 2.4	5.5 ± 2.3	5.9 ± 2.4	5.7 ± 2.4	5.6 ± 2.4	5.7 ± 2.4

Gross target volumes (GTV) delineated on planning CT only (GTV<sub>C</sub>), CT combined with endoscopy/EUS (GTV<sub>CE</sub>), CT combined with endoscopy/EUS and esophagography (X-ray) (GTV<sub>CEX</sub>), and CT combined with endoscopy/EUS, esophagography, and FDG-PET/CT (GTV<sub>CEXP</sub>).

## Generalized CI (CI<sub>gen</sub>)

**Table 4** lists the mean CI<sub>gen</sub> values for the four GTVs derived from different multimodal imaging combinations (mean intra-observer CI<sub>gen</sub>) for each observer. The mean intra-observer CI<sub>gen</sub> values for all observers was 0.73 ± 0.15. The mean intra-observer CI<sub>gen</sub> was the largest for observer 4, exhibiting significant differences with that for the other observers. The mean intra-observer CI<sub>gen</sub> for observer 1 was the lowest, exhibiting significant differences with that for observers 3, 4, and 5. The mean intra-observer CI<sub>gen</sub> among different observers was statistically significant ( $F=32.493$ ,  $p<0.001$ ). **Table 5** lists the mean CI<sub>gen</sub> values for the five GTVs derived from different observers (mean inter-observer CI<sub>gen</sub>) for each multimodal imaging combination. The mean inter-observer CI<sub>gen</sub> values for the four multimodal imaging combinations was 0.67 ± 0.11. The mean inter-observer CI<sub>gen</sub> was the largest for the fourth multimodal imaging combination, which exhibited significant differences with that for the other three combinations. The mean inter-observer CI<sub>gen</sub> among the different multimodal imaging combinations showed a significant difference ( $F=6.872$ ,  $p<0.001$ ).

## Three-Dimensional (3D) Centroid Shifts

**Table 4** lists the mean 3D centroid shifts of the four GTVs derived from different multimodal imaging combinations (mean intra-observer 3D centroid shifts) for each observer. The mean intra-observer 3D centroid shifts for all observers was 3.67 ± 4.62 mm. The mean intra-observer 3D centroid shifts for observer 4 showed significant differences compared with the other observers. The mean intra-observer 3D centroid shifts among different observers was significant ( $F=3.898$ ,  $p=0.004$ ). **Table 5**

presents the 3D centroid shifts of the five GTVs derived from different observers (mean inter-observer 3D centroid shifts) for each multimodal imaging combination. The mean inter-observer 3D centroid shifts for all four multimodal imaging combinations was 3.81 ± 4.7 mm. The mean inter-observer 3D centroid shifts among the different multimodal imaging combinations showed no significant difference ( $F=0.327$ ,  $p=0.806$ ).

## DISCUSSION

Uncertainties in volume delineation for esophageal carcinomas is a well-recognized potential cause of treatment failure in radiotherapy (27, 28). Minimizing the inter-/intra-observer delineation variability in volume delineation is regarded as an effective alternative method to define the GTV accurately (29, 30), since the gold standard of a pathological reference volume is rarely attainable (31, 32). The significance of quantifying the degree of variability or uncertainty in volume delineation is that the resulting impact on dosimetry and clinical outcomes (29, 30).

Accurate target delineation for esophageal cancer is often restricted by the poor discriminative value of current imaging modalities (23), particularly CT, and the inability to relate diagnostic endoscopy/EUS, esophagography, or FDG-PET/CT information to the planning CT images (13–17, 23). However, reasonable pretreatment staging assessments are essential to determine a rational treatment strategy. In each patient with newly diagnosed esophageal cancer, the acquired diagnostic imaging information should identify the feasibility of delineating the GTVs of the primary based on the planning CT

**TABLE 3** | Comparison the tumor length measured by CT, endoscopy/EUS, esophagography, and FDG-PET/CT with the mean longitudinal length measured by five observers for GTV<sub>C</sub>, GTV<sub>CE</sub>, GTV<sub>CEX</sub>, and GTV<sub>CEXP</sub>.

Imaging modality		CT	Endoscopy/EUS	Esophagography	PET-CT
<b>Tumor length(cm)</b>		5.5 ± 2.2	5.1 ± 2.0	5.3 ± 2.0	5.4 ± 2.2
<b>Target volume</b>		<b>GTV<sub>C</sub></b>	<b>GTV<sub>CE</sub></b>	<b>GTV<sub>CEX</sub></b>	<b>GTV<sub>CEXP</sub></b>
<b>Longitudinal length(cm)</b>		5.7 ± 2.5	5.8 ± 2.4	5.6 ± 2.3	5.8 ± 2.3
<b>Paired comparison</b>	<b>t-value</b>	-0.704	-1.759	-1.272	-2.172
	<b>p-value</b>	0.485	0.086	0.209	0.035

Gross target volumes (GTV) delineated on planning CT only (GTV<sub>C</sub>), CT combined with endoscopy/EUS (GTV<sub>CE</sub>), CT combined with endoscopy/EUS and esophagography (X-ray) (GTV<sub>CEX</sub>), and CT combined with endoscopy/EUS, esophagography, and FDG-PET/CT (GTV<sub>CEXP</sub>).

**TABLE 4** | The  $CI_{gen}$  values and 3D centroid shifts (Mean  $\pm$  SD) of the four GTVs derived from different combinations of multimodal imaging for each observer.

Parameter	Observer 1	Observer 2	Observer 3	Observer 4	Observer 5
$CI_{gen}$	0.68 $\pm$ 0.12	0.69 $\pm$ 0.14	0.75 $\pm$ 0.15	0.80 $\pm$ 0.15	0.74 $\pm$ 0.17
<b>Paired t-test</b>	0.173 (vs Obs2)	–	–	–	–
<b>(p value)</b>	<0.001 (vs Obs3)	<0.001 (vs Obs3)	–	–	–
	<0.001 (vs Obs4)	<0.001 (vs Obs4)	<0.001 (vs Obs4)	–	–
	<0.001 (vs Obs5)	<0.001 (vs Obs5)	0.264 (vs Obs5)	<0.001 (vs Obs5)	–
<b>CI (<math>G_C</math>, <math>G_{CEXP}</math>)</b>	0.66 $\pm$ 0.12	0.70 $\pm$ 0.18	0.75 $\pm$ 0.13	0.78 $\pm$ 0.17	0.73 $\pm$ 0.18
<b>CI (<math>G_C</math>, <math>G_{CE}</math>)</b>	0.69 $\pm$ 0.15	0.68 $\pm$ 0.14	0.77 $\pm$ 0.15	0.84 $\pm$ 0.15	0.78 $\pm$ 0.17
<b>Paired t-test</b>	-1.718	1.247	-0.848	-2.592	-2.666
<b>(p value)</b>	0.092	0.218	0.4	0.013	0.01
<b>3D shifts(mm)</b>	3.69 $\pm$ 4.47	4.34 $\pm$ 4.24	3.84 $\pm$ 3.93	2.85 $\pm$ 4.24	3.65 $\pm$ 6.19
<b>Paired t-test</b>	0.023 (vs Obs2)	–	–	–	–
<b>(p value)</b>	0.572 (vs Obs3)	0.077 (vs Obs3)	–	–	–
	0.005 (vs Obs4)	<0.001 (vs Obs4)	0.001 (vs Obs4)	–	–
	0.911 (vs Obs5)	0.084 (vs Obs5)	0.59 (vs Obs5)	0.039 (vs Obs5)	–

Gross target volumes (GTV) delineated on planning CT only ( $G_C$ ), CT combined with endoscopy/EUS ( $G_{CE}$ ), and CT combined with endoscopy/EUS, esophagography, and FDG-PET/CT ( $G_{CEXP}$ ). The CI between  $G_C$  and  $G_{CEXP}$  [CI ( $G_C$ ,  $G_{CEXP}$ )], the CI between  $G_C$  and  $G_{CE}$  [CI ( $G_C$ ,  $G_{CE}$ )].

**TABLE 5** | The  $CI_{gen}$  values and 3D centroid shifts (Mean  $\pm$  SD) of the five GTVs delineated by different observers based on each combinations of multimodal imaging.

Parameter	GTV <sub>C</sub>	GTV <sub>CE</sub>	GTV <sub>CEX</sub>	GTV <sub>CEXP</sub>
$CI_{gen}$	0.66 $\pm$ 0.13	0.66 $\pm$ 0.12	0.67 $\pm$ 11	0.69 $\pm$ 0.10
<b>Paired t-test</b>	0.443 (vs $G_{CE}$ )	–	–	–
<b>(p value)</b>	<0.088 (vs $G_{CEX}$ )	0.269 (vs $G_{CEX}$ )	–	–
	<0.001 (vs $G_{CEXP}$ )	<0.001 (vs $G_{CEXP}$ )	<0.001 (vs $G_{CEXP}$ )	–
<b>3D shifts (mm)</b>	3.78 $\pm$ 4.04	3.78 $\pm$ 3.79	3.98 $\pm$ 5.03	3.68 $\pm$ 5.94
<b>Paired t-test</b>	0.981 (vs $G_{CE}$ )	–	–	–
<b>(p value)</b>	0.463 (vs $G_{CEX}$ )	0.463 (vs $G_{CEX}$ )	–	–
	0.762 (vs $G_{CEXP}$ )	0.744 (vs $G_{CEXP}$ )	0.218 (vs $G_{CEXP}$ )	–

Gross target volumes (GTV) delineated on planning CT only (GTV<sub>C</sub>), CT combined with endoscopy/EUS (GTV<sub>CE</sub>), CT combined with endoscopy/EUS and esophagography (X-ray) (GTV<sub>CEX</sub>), and CT combined with endoscopy/EUS, esophagography, and FDG-PET/CT (GTV<sub>CEXP</sub>).

image with reference to the above-mentioned information. In this study, the geometric features of the GTVs derived from different observers and different planning CT image combinations were compared with the diagnostic imaging information. Furthermore, the value of the different planning CT image combinations in conjunction with diagnostic imaging information was evaluated for tumor delineation in esophageal carcinoma.

The results of this study showed no statistically significant inter-observer differences in the esophageal volume estimation based on different combinations of the CT, endoscopy/EUS, esophagography, and FDG-PET/CT data (Table 2). For a particular multimodal imaging combination, different observers reported similar estimates for the GTV based on a similar knowledge of multimodal imaging. Moreover, for each observer, the volumes of the four GTVs delineated on different multimodal imaging combinations showed no significant differences. This indicates that the GTV volume assessments on different multimodal imaging combinations did not transform/change for the same observer. The data presented here is similar to the results reported in other literature (33, 34). However, Choi et al. (13) reported that the number of observers and number of observations made might affect the level of significance. In this study, many significant differences were observed in the GTV volume between different observers in

the pairwise comparisons. Therefore, inter-observer variation in the target delineation could not be revealed/identified by merely comparing the volumes of the GTVs.

Similar to the observed variability in the volumes of the GTVs, the inter-observer and intra-observer variability in the longitudinal length showed no statistically significant differences (Table 3). However, some significant differences between different observers were identified in the pairwise comparisons. The main reason behind these differences might be a different understanding of the procedure of determining the tumor length on multimodal imaging by different observers. Radiation oncologists have always found the procedure to determine the proximal and distal extension of esophageal carcinoma based on different images challenging. Conventional images from CT, endoscopy/EUS, and esophagography, and MRI or FDG-PET/CT have their share of advantages and limitations for determining the tumor length (14–17, 22, 35, 36). It is critical to familiarize radiation oncologists with these advantages and limitations before selecting the different image combinations. In this study, the tumor length determined by the multimodal images tended to be larger than that measured by a single image. In particular, the longitudinal length of GTV<sub>CEXP</sub> was significantly larger than the tumor length measured by FDG-PET/CT. Therefore, the use of the multimodal images to determine the target length contributes to reducing the

limitation of a single image, and improving the accuracy of target delineation; however, this is based on the precondition that the observers develop a good knowledge of the features of the multimodal images *via* unified training.

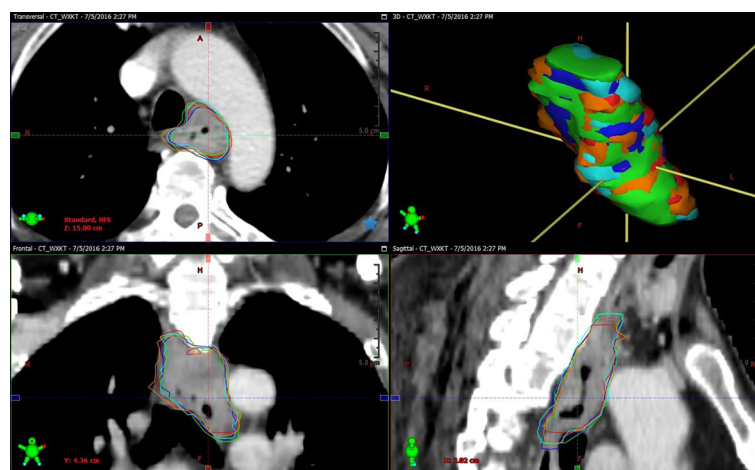
The  $CI_{gen}$  values for  $GTV_C$ ,  $GTV_{CE}$ ,  $GTV_{CEX}$ , and  $GTV_{CEXP}$  for each observer represent the intra-observer variations, which include the random and inherent variations derived from different multimodal imaging combinations for the same observer. Here, the mean  $CI_{gen}$  for intra-observer variability (0.73) was larger than that for inter-observer variability (0.67). This indicates that the intra-observer variability in delineating esophageal tumors was lower than the inter-observer variability, which shows agreement with the results reported in other studies (33, 34). Machiels et al. (33) reported the mean  $CI_{gen}$  values for intra-observer delineation variability and inter-observer variability in ten patients without endoscopically implanted fiducial markers versus those with markers to be 0.54 versus 0.68 and 0.68 versus 0.75, respectively. Vollenbrock et al. (34) reported the mean  $CI_{gen}$  over six patients as 0.68 on FDG-PET/CT, 0.66 on  $T_2W$ -MRI, and 0.68 on  $T_2W+DW$ (diffusion-weighted)-MRI. Compared with the above studies, fifty patients with upper, middle, and lower thoracic esophageal carcinoma were enrolled in this study. Moreover, different multimodal imaging combinations, including CT, endoscopy/EUS, esophagography, and FDG-PET/CT, were employed to eliminate any bias from a single imaging technique.

In addition, the CT is a basic image ( $GTV_C$ ). CT combined with endoscopy/EUS is a simple combination ( $GTV_{CE}$ ), while CT combined with endoscopy/EUS, esophagography, and FDG-PET/CT ( $GTV_{CEXP}$ ) is regarded as an effective alternative method to define the GTV accurately (Table 5). Therefore, The CI between  $GTV_C$  and  $GTV_{CEXP}$ , was significantly less than the CI between  $GTV_C$  and  $GTV_{CE}$  for all observers ( $t = -3.018$ ,  $p = 0.003$ ), which suggested that a comprehensive

combination of multimodal images was more conducive to influence the target delineation compared a simple combination.

In the ANOVA analysis, the intra-observer  $CI_{gen}$  for the GTVs derived from different multimodal imaging combinations among the five observers was statistically significant ( $p < 0.001$ ). The intra-observer  $CI_{gen}$  for the senior radiation oncologists (observers 3, 4, and 5) was larger than that for the junior radiation oncologists (observers 1 and 2). An optimum intra-observer  $CI_{gen}$  was obtained for the senior radiotherapist who spent minimal time delineating the GTVs (observer 4) (Table 4). The senior radiotherapists, who were generally familiar with the multimodal imaging features for distinguishing the tumors from the normal structures and the location subject to relapse, might not be easily affected when only a single imaging modality is used/available for target contouring (37, 38). In addition, the shorter repeating delineation intervals did not eliminate the record of previous delineations, which might have improved the consistency of the target delineation. This suggests that background knowledge in medical imaging, clinical experience, and repeating delineation intervals might affect the intra-observer variability of the target  $CI_{gen}$ . Strengthening the target delineation and medical imaging knowledge training contributes to improve the accuracy of target delineation for EC.

While the inter-observer  $CI_{gen}$  calculated for the different multimodal imaging combinations did not increase for the combined CT and endoscopy/EUS data, as compared with CT only,  $CI_{gen}$  tended to increase for the combined CT and endoscopy/EUS and esophagography information ( $p = 0.088$ ). Furthermore, the addition of FDG-PET/CT to the endoscopy/EUS, and esophagography data significantly improved the inter-observer  $CI_{gen}$ . The use of multimodal imaging, including CT, endoscopy/EUS, esophagography, and FDG-PET/CT, for target delineation reduced the inter-observer variability.



**FIGURE 2** | Example of GTVs delineated based on CT combined with endoscopy/EUS, esophagogram, and FDG-PET/CT ( $GTV_{CEXP}$ ) by observer 1 (green segment), observer 2 (red segment), observer 3 (blue segment), observer 4 (orange segment), and observer 5 (cyan segment) in transversal, frontal, surface, and sagittal planes for one patient (Patient 10). The volume and longitudinal length of  $GTV_{CEXP}$  delineated on CT combined with endoscopy/EUS, esophagogram, and FDG-PET/CT exhibiting a good consistency.

**Figure 2** showed example of GTVs delineated based on CT combined with endoscopy/EUS, esophagogram, and FDG-PET/CT ( $GTV_{CEXP}$ ) by five oncologists, and the volume and longitudinal length of  $GTV_{CEXP}$  exhibited a good consistency. The effect of FDG-PET/CT on the intra- and inter-observer variability of target volume delineation in patients with gastro-esophageal cancer remains controversial. Vesprini et al. (39) reported that the combined use of FDG-PET/CT based on CT for GTV delineation significantly decreased both intra- and inter-observer variability, while Schreurs et al. (40) did not find PET/CT to have a significant effect on the inter-observer variability. Therefore, besides FDG-PET/CT, the additional use of endoscopy/EUS and esophagography for target delineation might prove beneficial in reducing the inter-observer variability. Recent studies have shown that the use of endoscopically implanted fiducial markers and MRI might reduce the variability of target volume delineation (33, 34). The use of multimodal imaging has proved increasingly valuable in improving the accuracy of target definition in esophageal carcinoma.

The target conformity index (CI) is mainly influenced by positional and morphological difference of targets. Here, the intra-observer variability in the 3D centroid shifts of the GTVs among different observers showed a significant difference. In the case of no significant variability in the volume of the GTVs, the intra-observer variability in the position could have mainly contributed to the statistical significance in the  $CI_{gen}$  value. The intra-observer variability in the shape also tends to affect the intra-observer  $CI_{gen}$ . The inter-observer variability in the 3D centroid shifts and volume of the GTVs showed no significant differences, suggesting that the inter-observer variability in the shape had a critical influence on the inter-observer  $CI_{gen}$ . Thus, this study indirectly implies that using different multimodal image combinations might transform/change different observers' visual perception of tumors. In addition, we found the 3D centroid shifts were 3-4mm either between the observers themselves or between the observers. Therefore, whether it is necessary to expand an extra margin to include this error is a clinical problem and deserves further thinking.

## CONCLUSION

In conclusion, for radiation oncologists with advanced medical imaging training and clinical experience, the use of diagnostic multimodal images from endoscopy/EUS, esophagography, and FDG-PET/CT for target delineation based on planning CT

reduced the intra- and inter-observer variability and increased the accuracy of target delineation in primary thoracic esophageal carcinomas. The use of the combination of multimodal images would reduce uncertainties in volume delineation for esophageal carcinomas, and potentially increase the success rate of radiotherapy. We also found the inter/intra observer variability in the 3D centroid shifts of GTVs were about 3-4mm, whether it is necessary to expand an extra margin to include this error deserves further thinking.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Shandong Cancer Hospital and Institute. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

FL and YL contributed to the study design, the patient enrollment, the data statistics and analysis, and writing the manuscript. JL participated in the study design. XW contributed to the patient enrollment. YZ, XL, SL, JW, and MX contributed to the delineation. WW and YG made important contributions in collecting the data and revising the content. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by Shandong Cancer Hospital and Institute, China. National Natural Science Foundation of China (817732870); Beijing Medical Award Foundation (YXJL-2020-0785-0616); Taishan Scholars Program of Shandong Province (NO.ts 20190982); Wu Jieping Medical Foundation (320.6750.2021-02-79).

## REFERENCES

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2018) 68:394–424. doi: 10.3322/caac.21492
- Sjoquist KM, Burmeister BH, Smithers BM, Zalcberg JR, Simes RJ, Barbour A, et al. Survival After Neoadjuvant Chemotherapy or Chemoradiotherapy for Resectable Oesophageal Carcinoma: An Updated Meta-Analysis. *Lancet Oncol* (2011) 12(7):681–92. doi: 10.1016/S1470-2045(11)70142-5
- Shapiro J, van Lanschot JJB, Hulshof MCCM, van Hagen P, van Berge Henegouwen MI, Wijnhoven BPL, et al. Neoadjuvant Chemoradiotherapy Plus Surgery Versus Surgery Alone for Oesophageal or Junctional Cancer (CROSS): Long-Term Results of a Randomised Controlled Trial. *Lancet Oncol* (2015) 16(9):1090–8. doi: 10.1016/S1470-2045(15)00040-6



4. Lagergren J, Smyth E, Cunningham D, Lagergren P. Oesophageal Cancer. *Lancet* (2017) 390(10110):2383–96. doi: 10.1016/S0140-6736(17)31462-9
5. Cooper JS, Guo MD, Herskovic A, Macdonald JS, Martenson JA Jr, Al-Sarraf M, et al. Chemoradiotherapy of Locally Advanced Esophageal Cancer: Long-Term Follow-Up of a Prospective Randomized Trial (RTOG 85-01). *JAMA* (1999) 281(17):1623–7. doi: 10.1001/jama.281.17.1623
6. Rohatgi PR, Swisher SG, Correa AM, Wu TT, Liao Z, Komaki R, et al. Failure Patterns Correlate With the Proportion of Residual Carcinoma After Preoperative Chemoradiotherapy for Carcinoma of the Esophagus. *Cancer* (2005) 104(7):1349–55. doi: 10.1002/cncr.21346
7. Welsh J, Settle SH, Amini A, Xiao L, Suzuki A, Hayashi Y, et al. Failure Patterns in Patients With Esophageal Cancer Treated With Definitive Chemoradiation. *Cancer* (2012) 118(10):2632–40. doi: 10.1002/cncr.26586
8. Amini A, Ajani J, Komaki R, Allen PK, Minsky BD, Blum M, et al. Factors Associated With Local-Regional Failure After Definitive Chemoradiation for Locally Advanced Esophageal Cancer. *Ann Surg Oncol* (2014) 21(1):306–14. doi: 10.1245/s10434-013-3303-0
9. Xi M, Yang Y, Zhang L, Yang H, Merrell KW, Hallemeier CL, et al. Multi-Institutional Analysis of Recurrence and Survival After Neoadjuvant Chemoradiotherapy of Esophageal Cancer: Impact of Histology on Recurrence Patterns and Outcomes. *Ann Surg* (2019) 269(4):663–70. doi: 10.1097/SLA.0000000000002670
10. Nowee ME, Voncken FEM, Kotte ANTJ, Goense L, van Rossum PSN, van Lier ALHMW, et al. Dutch National Platform for Radiotherapy of Gastrointestinal Tumours (LPRGE) Group. Gross Tumour Delineation on Computed Tomography and Positron Emission Tomography-Computed Tomography in Oesophageal Cancer: A Nationwide Study. *Clin Transl Radiat Oncol* (2018) 14:33–9. doi: 10.1016/j.ctro.2018.10.003
11. Muijs C, Smit J, Karrenbeld A, Beukema J, Mul V, van Dam G, et al. Residual Tumor After Neoadjuvant Chemoradiation Outside the Radiation Therapy Target Volume: A New Prognostic Factor for Survival in Esophageal Cancer. *Int J Radiat Oncol Biol Phys* (2014) 88(4):845–52. doi: 10.1016/j.ijrobp.2013.11.009
12. Steenbakkers RJ, Duppen JC, Fitton I, Deurloo KE, Zijp LJ, Comans EF, 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):435–48. doi: 10.1016/j.ijrobp.2005.06.034
13. Choi HJ, Kim YS, Lee SH, Lee YS, Park G, Jung JH, et al. Inter- and Intra-Observer Variability in Contouring of the Prostate Gland on Planning Computed Tomography and Cone Beam Computed Tomography. *Acta Oncol* (2011) 50(4):539–46. doi: 10.3109/0284186X.2011.562916
14. Pollard JM, Wen Z, Sadagopan R, Wang J, Ibbott GS. The Future of Image-Guided Radiotherapy Will be MR Guided. *Br J Radiol* (2017) 90:20160667. doi: 10.1259/bjr.20160667
15. Drudi FM, Trippa F, Cascone F, Righi A, Iascone C, Ricci P, et al. Esophagogram and CT vs Endoscopic and Surgical Specimens in the Diagnosis of Esophageal carcinoma. *Radiol Med* (2002) 103(4):344–52.
16. Encaoua J, Abgral R, Leleu C, El Kabaj O, Caradec P, Bourhis D, et al. Radiotherapy Volume Delineation Based on (<sup>18</sup>F)-Fluorodeoxyglucose Positron Emission Tomography for Locally Advanced or Inoperable Oesophageal Cancer. *Cancer Radiother* (2017) 21(4):267–75. doi: 10.1016/j.canrad.2016.12.004
17. Caputo FM, Buquicchio GL. Esophageal Cancer Staging: The Role of Radiology. *Rays* (2005) 30(4):309–14.
18. Wang B, Liu C, Lin C, Hsu PK, Hsu WH, Wu YC, et al. Endoscopic Tumor Length Is an Independent Prognostic Factor in Esophageal Squamous Cell Carcinoma. *Ann Surg Oncol* (2012) 19:2149–58. doi: 10.1245/s10434-012-2273-y
19. Gao XS, Qiao X, Wu F, Cao L, Meng X, Dong Z, et al. Pathological Analysis of Clinical Target Volume Margin for Radiotherapy in Patients With Esophageal and Gastroesophageal Junction Carcinoma. *Int J Radiat Oncol Biol Phys* (2007) 67:389–96. doi: 10.1016/j.ijrobp.2006.09.015
20. Muijs CT, Beukema JC, Pruijm J, Mul VE, Groen H, Plukker JT, et al. A Systematic Review on the Role of FDG-PET/CT in Tumor Delineation and Radiotherapy Planning in Patients With Esophageal Cancer. *Radiother Oncol* (2010) 97(2):165–71. doi: 10.1016/j.radonc.2010.04.024
21. Thomas L, Lapa C, Bundschuh RA, Polat B, Sonke JJ, Guckenberger M, et al. Tumor Delineation in Oesophageal Cancer – A Prospective Study of Delineation in PET and CT With and Without Endoscopically Placed Clip Markers. *Radiother Oncol* (2015) 116:269–75. doi: 10.1016/j.radonc.2015.07.007
22. Metser U, Even-Sapir E. Increased (<sup>18</sup>F)-Fluorodeoxyglucose Uptake in Benign, Nonphysiologic Lesions Found on Whole-Body Positron Emission Tomography/Computed Tomography (PET/CT): Accumulated Data From Four Years of Experience With PET/Ct. *Semin Nucl Med* (2007) 37:206–22. doi: 10.1053/j.semnuclmed.2007.01.001
23. Louie AV, Rodrigues G, Olsthoorn J, Palma D, Yu E, Yaremko B, et al. Inter-Observer and Intra-Observer Reliability for Lung Cancer Target Volume Delineation in the 4D-CT Era. *Radiother Oncol* (2010) 95(2):166–71. doi: 10.1016/j.radonc.2009.12.028
24. Li FX, Li JB, Zhang YJ, Xu M, Shang D, Fan T, et al. Geometrical Differences in Gross Target Volumes Between 3DCT and 4DCT Imaging in Radiotherapy for Non-Small-Cell Lung Cancer. *J Radiat Res* (2013) 54(5):950–6. doi: 10.1093/jrr/rrt017
25. Kouwenhoven E, Gezien M, Struikmans H. Measuring the Similarity of Target Volume Delineations Independent of the Number of Observers. *Phys Med Biol* (2009) 54:2863 – 73. doi: 10.1088/0031-9155/54/9/018
26. Li FX, Li JB, Zhang YJ, Liu TH, Tian SY, Xu M, et al. Comparison of the Planning Target Volume Based on Three-Dimensional CT and Four-Dimensional CT Images of Non-Small-Cell Lung Cancer. *Radiother Oncol* (2011) 99(2):176–80. doi: 10.1016/j.radonc.2011.03.015
27. Thomas M, Mortensen HR, Hoffmann L, Moller DS, Troost EGC, Muijs CT, et al. Proposal for the Delineation of Neoadjuvant Target Volumes in Oesophageal Cancer. *Radiother Oncol* (2021) 156:102–12. doi: 10.1016/j.radonc.2020.11.032
28. Chang X, Deng W, Wang X, Zhou Z, Yang J, Guo W, et al. Interobserver Variability in Target Volume Delineation in Definitive Radiotherapy for Thoracic Esophageal Cancer: A Multi-Center Study From China. *Radiat Oncol* (2021) 16(1):102. doi: 10.1186/s13014-020-01691-4
29. 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* (2016) 121(2):169–79. doi: 10.1016/j.radonc.2016.09.009
30. 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. *Strahlenther Onkol* (2003) 179(1):21–30. doi: 10.1007/s00066-003-0976-5
31. Daisne JF, Duprez T, Weynand B, Lonnet M, Hamoir M, Reyckers H, et al. Tumor Volume in Pharyngolaryngeal Squamous Cell Carcinoma: Comparison at CT, MR Imaging, and FDG PET and Validation With Surgical Specimen. *Radiology* (2004) 233:93–100. doi: 10.1148/radiol.2331030660
32. Wanet M, Lee JA, Weynand B, De Bast M, Poncelet A, Lacroix V, et al. Gradient-Based Delineation of the Primary GTV on FDG-PET in Nonsmall Cell Lung Cancer: A Comparison With Thresholdbased Approaches, CT and Surgical Specimens. *Radiother Oncol* (2011) 98:117–25. doi: 10.1016/j.radonc.2010.10.006
33. Machiels M, Jin P, van Hooft JE, Gurney-Champion OJ, Jelvehgaran P, Geijsen ED, et al. Reduced Inter-Observer and Intra-Observer Delineation Variation in Esophageal Cancer Radiotherapy by Use of Fiducial Markers. *Acta Oncol* (2019) 58(6):943–50. doi: 10.1080/0284186X.2019.1588991
34. Vollenbrock SE, Nowee ME, Voncken FEM, Kotte ANTJ, Goense L, van Rossum PSN, et al. Gross Tumor Delineation in Esophageal Cancer on MRI Compared With 18f-FDG-PET/Ct. *Adv Radiat Oncol* (2019) 4(4):596–604. doi: 10.1016/j.adro.2019.04.004
35. Hou DL, Shi GF, Gao XS, Asaumi J, Li XY, Liu H, et al. Improved Longitudinal Length Accuracy of Gross Tumor Volume Delineation With Diffusion Weighted magnetic Resonance Imaging for Esophageal Squamous Cell Carcinoma. *Radiat Oncol* (2013) 8:169. doi: 10.1186/1748-717X-8-169
36. Machiels M, van Montfoort ML, Thuijs NB, Henegouwen MIVB, Hulshof MCCM. PO-0812: Pathological Validation of Endoscopically Placed Fiducials

- on Tumor Borders in Esophageal Cancer. *Radiother Oncol* (2019) 133:S423–S24. doi: 10.1016/S0167-8140(19)31232-0
37. Van de Steene J, Linthout N, de Mey J, Vinh-Hung V, Claassens C, Noppen M, et al. Definition of Gross Tumor Volume in Lung Cancer: Inter-Observer Variability. *Radiother Oncol* (2002) 62(1):37–49. doi: 10.1016/S0167-8140(01)00453-4
38. Schimek-Jasch T, Troost EG, Rücker G, Prokic V, Avlar M, Duncker-Rohr V, et al. A Teaching Intervention in a Contouring Dummy Run Improved Target Volume Delineation in Locally Advanced Non-Small Cell Lung Cancer: Reducing the Interobserver Variability in Multicentre Clinical Studies. *Strahlenther Onkol* (2015) 191(6):525–33. doi: 10.1007/s00066-015-0812-8
39. Schreurs LM, Busz DM, Paardekooper GM, Beukema JC, Jager PL, van der Jagt EJ, et al. Impact of 18-Fluorodeoxyglucose Positron Emission Tomography on Computed Tomography Defined Target Volumes in Radiation Treatment Planning of Esophageal Cancer: Reduction in Geographic Misses With Equal Inter-Observer Variability. *Dis Esophagus* (2010) 23(6):493–501. doi: 10.1111/j.1442-2050.2009.01044.x
40. Vesprini D, Ung Y, Dinniwel R, Breen S, Cheung F, Grabarz D, et al. Improving Observer Variability in Target Delineation for Gastro-Oesophageal Cancer – The Role of (18F)Fluoro-2-Deoxy-Dglucose Positron Emission Tomography/Computed Tomography. *Clin Oncol (R Coll Radiol)* (2008) 20:631–8. doi: 10.1016/j.clon.2008.06.004
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.
- Copyright © 2022 Li, Li, Wang, Zhang, Liu, Liu, Wang, Wang, Guo, Xu and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.