의학석사 학위논문

Comparison of Endoscopically Determined Gross Tumor Volume and Metabolic Tumor Volume in Esophageal Cancer

식도암에서의 내시경적으로 결정된 육안적 종양 체적과 대사성 종양 체적의 비교

2021년 2월

서울대학교 대학원

의학과 방사선종양학 전공

이 태 훈

A thesis of the Degree of Master of Science

식도암에서의 내시경적으로 결정된 육안적 종양 체적과 대사성 종양 체적의 비교

Comparison of Endoscopically Determined Gross Tumor Volume and Metabolic Tumor Volume in Esophageal Cancer

February 2021

Department of Medicine, Radiation Oncology Major
Seoul National University Graduate School
Tae Hoon Lee

Comparison of Endoscopically Determined Gross Tumor Volume and Metabolic Tumor Volume in Esophageal Cancer

지도교수 김 학 재

이 논문을 의학석사 학위논문으로 제출함

2020년 10월

서울대학교 대학원

의학과 방사선종양학 전공

이태훈

이태훈의 의학석사 학위논문을 인준함

2021년 1월

위 원 장 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

Comparison of Endoscopically Determined Gross Tumor Volume and Metabolic Tumor Volume in Esophageal Cancer

by

Tae Hoon Lee

A thesis submitted to the Department of Medicine in partial fulfillment of the requirements for the Degree of Master of Science in Medicine at Seoul National University Graduate School

January 2021

Approved by Thesis Committee:

Professor Eui Kyu Chie Chairman

Professor Nak Jae Lee

Professor Gi Jeong Cheon

ABSTRACT

Objective: The purpose of this study is to compare the longitudinal location of endoscopically-defined gross tumor volume (GTV) and positron emission tomography-based metabolic tumor volume (MTV) of esophageal cancer, and determine the optimal measures to delineate MTV using GTV as a reference.

Methods: A retrospective review of medical records was performed of the nine patients who underwent endoscopic placement of fiducial markers for radiotherapy of esophageal squamous cell carcinoma. Endoscopic hemoclips were used as the fiducial markers, and were placed at the superior and inferior borders of the endoscopically visible lesions. GTV was newly delineated solely based on the locations of the fiducial markers. The standardized uptake value (SUV) threshold corresponding to the superior and inferior borders of GTV was defined as the highest threshold that made MTV reach each border of GTV. Both the fixed relative and absolute threshold methods were used. The coefficients of variation of the threshold values from both thresholding methods were compared to establish which method would be more consistent to determine the threshold corresponding to the GTV borders.

Results: The median fixed relative and absolute thresholds were 32% and 3.8, respectively. The coefficient of variation was 0.781 for the fixed relative threshold method and 0.400 for the fixed absolute threshold method, indicating more

consistent results from the fixed absolute threshold method. All but two GTV borders were included in MTV with a SUV threshold of 2.5, which was used in previous studies. Esophageal tumors with a maximum SUV > 20 tended to have closer threshold values corresponding to the GTV borders to 2.5 (median 2.8 vs. 3.6, p = 0.069).

Conclusion: The fixed absolute threshold method was more suitable than the fixed relative threshold method for determining the MTV for esophageal lesions. A SUV of 2.5 was appropriate for esophageal tumors with a maximum SUV > 20. More study is needed to suggest a feasible threshold for all esophageal tumors.

Keywords: Esophageal cancer, PET scan, endoscopy, tumor volume

Student number: 2018-29944

CONTENTS

LIST OF TABLES AND FIGURES

LIST OF ABBREVIATIONS

RT: Radiotherapy

GTV: Gross tumor volume

CT: Computed tomography

PET: Positron emission tomography

MTV: Metabolic tumor volume

SUV: Standardized uptake value

CV: Coefficient of variation

CI: Concordance index

INTRODUCTION

Radiotherapy, often combined with chemotherapy, has an essential role as a definitive or neoadjuvant treatment in the management of esophageal cancer [1, 2]. After implementing conformal radiotherapy (RT), delineating an accurate target volume is an important stage of RT planning. There is no single definitive imaging modality to delineate precise gross tumor volume (GTV). As esophageal cancer can spread through the mucosa and submucosa, using a computed tomography (CT) scan alone to determine GTV of an esophageal tumor is challenging, particularly in the longitudinal direction, though a simulated CT scan is still essential for planning purposes. It is recommended that radiation oncologists incorporate information from multiple studies [3].

Endoscopy is used to access esophageal tumors, and an endoscopic description of the location and the length of the tumor correlates well with the pathological tumor extension and clinical features, such as prognosis [4]. Translation of the endoscopic description to a simulation CT scan is not intuitive [3], but the use of fiducial markers can be helpful in this process. Machiels et al. showed that endoscopy-guided implantation of a fiducial marker reduces variation in the inter- and intra-observer GTV delineation [5]. The same group compared endoscopically defined tumor borders and pathological findings using fiducial markers, and concluded that they are well-correlated [6]. Although not a routine process, placing

a fiducial marker can be very accurate to determine GTV for esophageal cancer RT planning, particularly for tumors with mucosal or submucosal spread.

^{18}F-Fluorodeoxyglucose (FDG) positron emission tomography/computed tomography (PET/CT) is a recommended study for staging and target delineation purposes, and introducing PET/CT to GTV delineation decreases uncertainty, such as interobserver variability [7]. PET/CT has higher sensitivity and specificity for detecting regional or distant metastases, but an esophageal primary tumor is not always accessible by PET/CT, particularly when the tumor is small or superficial [8]. It can be problematic to determine the cranial and caudal borders of the GTV based on PET/CT, as the tumor can spread through the superficial layer of the esophageal wall and this superficial spread might not be detected by PET/CT. The purpose of this study is to compare the longitudinal location of the GTV of esophageal squamous cell carcinoma based on fiducial markers placed by endoscopy. We also investigated metabolic tumor volume (MTV) based on hypermetabolic uptake on PET, and determined the optimal measures to delineate MTV using endoscopy-based GTV as a reference.

METHODS

1. Study population

This study was approved by the Institutional Review Board of Seoul National University Hospital (IRB number 1912-137-1091) before collecting patient information. The medical records of 100 patients with esophageal squamous cell carcinoma who had undergone RT between 2018 and 2019 were retrospectively reviewed. Among them, 27 patients had fiducial markers placed via endoscopy. Eight patients were excluded because the endoscope could not be passed through the esophageal tumor due to narrowing of the lumen, so the fiducial markers were placed only in the proximal margin of the esophageal tumor and the longitudinal margin was not determined by endoscopy. Two patients did not undergo pre-RT PET, and six patients had no or minimal hypermetabolism in the primary esophageal tumor, so these patients were excluded. Two other patients were excluded as they had diffuse malignant lesions of the esophagus. After these exclusions, nine patients remained and were included in this study.

2. Fiducial marker placement

All patients underwent endoscopy to examine the esophageal lesions before treatment. Endoscopic procedures for placing the fiducial markers were performed

separately from diagnostic (echo) endoscopy. Two patients underwent endoscopy with midazolam-based sedation. Endoscopic stainless steel hemoclips (EZ Clip, HX-610-090L, Olympus, Tokyo, Japan) with a 10-mm open width were used as the fiducial markers. At least one fiducial marker was placed in each superior and inferior border of the endoscopically visible esophageal lesion, with a 0.5–1.0 cm interval from the actual margins of the observed lesion. An experienced gastroenterologist performed the procedure, and explanations of the exact location of the fiducial markers relative to the esophageal tumor were written for each patient.

3. GTV delineation

A simulation CT scan was acquired after placing the fiducial markers on the same day, while maintaining nil per os before the CT scan, due to potential hemoclip displacement from food intake. Patients were positioned with both arms abducted over their head and immobilized using a wing board. Patients with cervical or high upper thoracic esophageal lesions were immobilized in the supine position with IMRT Aquaplast to reduce RT set-up error. The thickness of the axial cut for the simulation CT scan was 3 mm. GTVs for this study were newly delineated by a contouring system (ARIA Oncology Information System 13.6, Varian Medical Systems, Palo Alto, CA, USA) for each eligible patient based on the simulation CT scan, independently from the GTVs and other volumes used in the actual treatment. An experienced radiation oncologist delineated the GTVs based on the appearance

of the radio-opaque fiducial markers on the simulation CT scans only. A fiducial marker often appeared in several axial cuts of the CT scan, and the fiducial marker was considered placed in the middle of these axial cuts. The cranial and caudal borders of GTV were determined based on the location of the fiducial markers and the gastroenterologist's description about the spatial relationship between the fiducial markers and mucosal spread of the esophageal tumor. After determining the cranial and caudal borders, the whole esophagus, which appeared in axial cuts between these borders, was included in the GTV. Body contour was acquired using the Search Body function in the contouring software during RT planning, and this contour was used for fusing the simulation CT and PET images.

4. PET image acquisition

Patients fasted for at least 6 hours, and PET/CT was performed 1 hour after intravenous injection of FDG (5.18 MBq/kg) using a dedicated PET/CT scanner (Biograph mCT40 or mCT64, Siemens Healthcare, Erlangen, Germany). A low-dose CT scan for attenuation correction and anatomical localization was acquired first, followed by acquisition of the PET images from the vertex to the proximal thigh (1 min/bed position). While acquiring images, patients were positioned with both arms abducted over their head, except for one patient who maintained supine position. The PET images were reconstructed using an iterative algorithm and displayed by fusing with the CT image.

5. Metabolic tumor volume and threshold corresponding to the GTV borders

GTV based on placement of the fiducial markers via endoscopy was transferred from the contouring system to the PET imaging system (MIM 6.1.7, MIM Software Inc., Beachwood, OH, USA) in the DICOM RTstruct format. The simulation CT scan and PET images were fused by contour-based alignment offered by the PET imaging system using the body contour acquired in the contouring system. For the purpose of determining the location of the primary esophageal tumor based on PET, MTV, which relies on hypermetabolism that appeared on the pre-RT PET of each patient, was delineated on the PET imaging system. The threshold-based method was applied to delineate MTV. Both fixed relative and fixed absolute thresholds were used. The fixed relative threshold is defined as a certain percentage of the maximum standardized uptake value (SUV) of a tumor, while the fixed absolute threshold is defined as the absolute value of the SUV. To determine the thresholds corresponding to each end of the GTV, the highest threshold that could make MTV to reach the most superior or most inferior axial plane of the GTV on the CT scan was established. The highest threshold was found by changing the threshold of MTV by 1% for fixed relative threshold, and by 0.1 for fixed absolute threshold. Thus, two threshold values, each corresponding to the superior and inferior borders of the GTV were obtained for each patient and each thresholding method. Threshold values by location were compared using the paired *t*-test.

The threshold values were verified to determine any correlations with other covariates, by linear regression when the covariate was continuous, and by the Kruskal–Wallis rank sum test when the covariate was categorical. The coefficient of variation (CV), which is defined as standard deviation divided by the average, was calculated from the threshold values to compare the consistency of the different methods to obtain a threshold.

6. Concordance index analysis and linear approximation

The concept of a concordance index (CI) is mainly used to measure a discrepancy between different volumes. In many other studies, the CI is defined as the ratio of the volume of the intersection and the volume of the union of two volumes [9]. In this study, longitudinal lengths rather than volume measurements were compared, so we defined the CI as the ratio of the length of intersection and the length of the union of GTV and each MTV, because the measured MTV tended to be smaller than GTV, as GTV is cylindrical while MTV is spherical due to a different delineation principle. The longitudinal length of GTV and MTV was measured based on how many axial planes were occupied by each volume on the simulation CT scan. An example of the CI calculation is illustrated in Figure 1. A CI > 0.8 was defined as good concordance.

To delineate MTV for CI analysis, both fixed relative and fixed absolute thresholds were used, too. Percentages of 10%, 20%, 30%, 40%, 50%, 60%, and 70%

of maximum SUV of the esophageal tumor were used as thresholds for the fixed absolute threshold method, and SUVs of 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, and 6.0 were used as thresholds for the fixed absolute threshold method. In some patients, MTV could not be determined for a low threshold, primarily due to diffuse background hypermetabolism of the esophagus. The threshold that had the highest CI for a particular patient was defined as the optimal threshold for the patient. The threshold that produced the highest CI for each patient and the highest CI of each patient from a certain threshold method were verified to determine if there were any correlations with other covariates by linear regression and by the Kruskal–Wallis rank sum test, as with the previous analysis. The CV was also calculated using the optimal threshold values.

Linear approximation to find the SUV threshold that would allow MTV to have the same volumetric properties as GTV was used in other studies [10, 11]. Therefore, we compared the results from the CI analysis and linear approximation. For each patient and the fixed relative and absolute threshold methods, a threshold that would make the longitudinal length of MTV to that of GTV was calculated by linear approximation, using the mentioned thresholds and corresponding longitudinal length of MTV. If longitudinal length of MTV was measured at the same multiple thresholds, the average values of these thresholds were used for approximation. Extrapolation derived from the closest values was applied when the resulting threshold was not expected to be between the mentioned thresholds. Figure 2 illustrates an example of linear approximation. The CV was calculated from the

optimal threshold values obtained from the CI analysis and linear approximation, to compare the consistency of the results. All statistical analyses were performed using R 3.6.0 (The R Foundation for Statistical Computing, Vienna, Austria).

Figure 1. Example of a concordance index (CI) calculation. Radiopaque fiducial markers are visible on the simulation CT scan (white arrow), and the gross tumor volume (GTV) (red line, A) was delineated based on these markers. Metabolic tumor volume (MTV) (magenta line, B) was delineated based on hypermetabolic uptake on a PET scan. The intersection of GTV and MTV was determined (cyan line). The CI was defined as the ratio of longitudinal lengths of the intersection of two volumes (C) and longitudinal lengths of the union of two volumes (D).

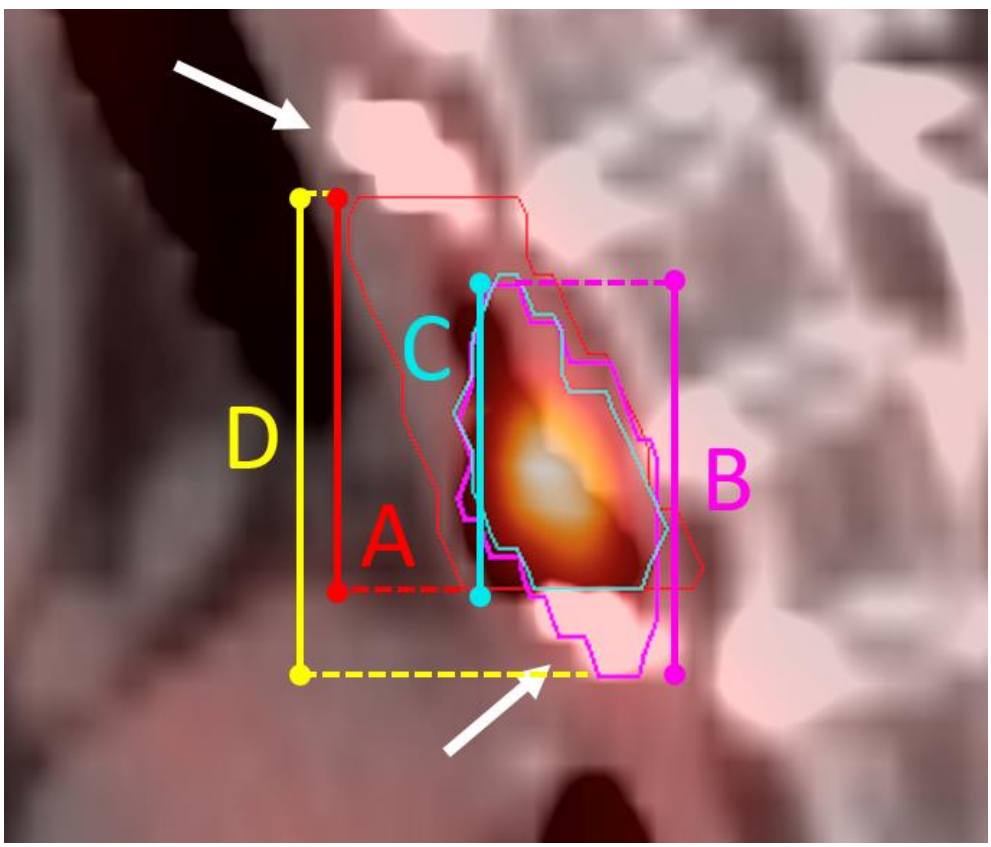
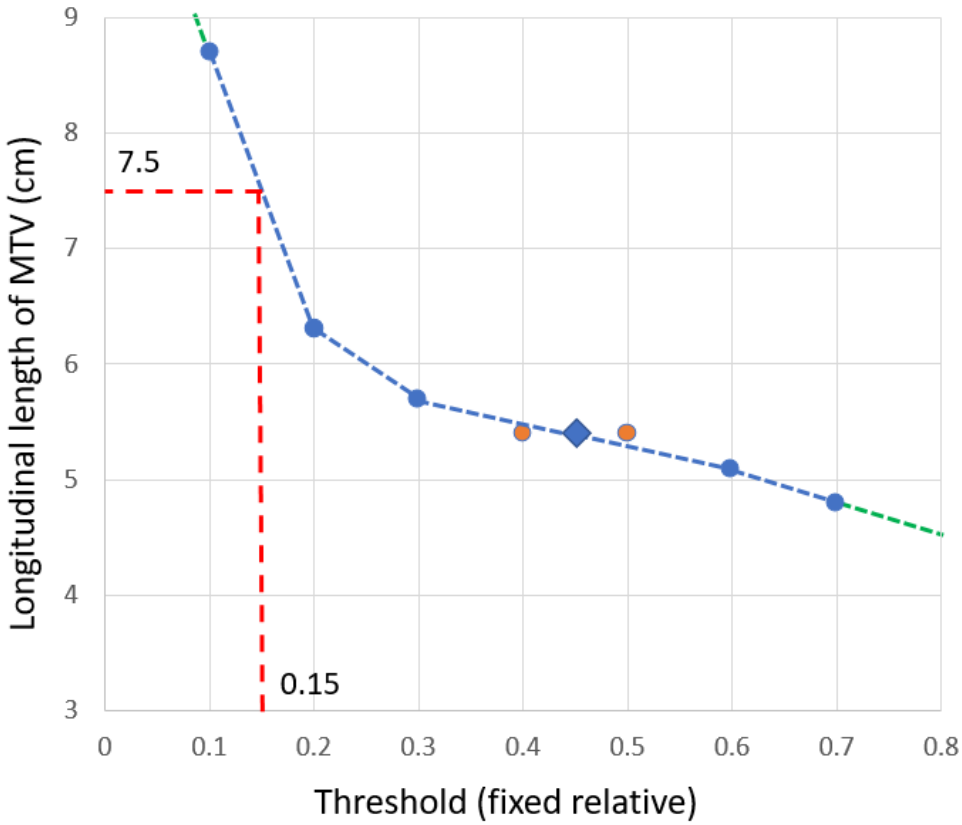


Figure 2. An example of linear approximation. Blue dots represent acquired longitudinal length of metabolic tumor volume (MTV) according to each threshold. Linear approximation was done between these points (blue dashed lines). In this case, the longitudinal length of gross tumor volume (GTV) was 7.5 cm, and the linear approximation shows that the corresponding threshold was 0.15 (15%). Thus, the optimal threshold for this patient was 15% by linear approximation. If the longitudinal length of MTV was measured in multiple thresholds (orange dots), the average values of these thresholds were used for approximation (blue diamonds). When the resulting threshold was not expected to be in between the mentioned thresholds, extrapolation derived from closest values was applied (green dashed lines).


Longitudinal length of MTV (cm)
9
8
7
6
5
4
3
7.5
0.15
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Threshold (fixed relative)

RESULTS

1. Patient characteristics

The characteristics of the nine patients included in this study are summarized in the left columns of Table 1. Seven patients were male and two were female. All but one patient was > 60 years. The range of T stages for the esophageal primary lesion was T1b to T3. Five patients had suspected or pathologically confirmed lymph node metastasis. The borders of the cervical and the upper, middle, and lower thoracic cancer were the thoracic inlet, the azygos vein, and the inferior pulmonary veins, respectively. Two patients had cervical esophageal lesions, three had upper thoracic, two had middle thoracic, and two had lower thoracic lesions. Three patients underwent the simulation CT scan in the supine position, due to the superior location of their esophageal lesions, as stated before. The range of endoscopically defined GTV was 0.9 to 8.0 cm, and the average was 4.7 cm. The range of maximum SUV from the PET scan for each patient was 7.1 to 28.6, and the average was 16.2.

Table 1. Patient characteristics and threshold corresponding to the gross tumor volume (GTV) borders

Patient number	Sex	Age at diagnosis	T stage	LN metastasis	Disease location	Position at simulation CT	Longitudinal length of GTV (cm)	Maximum SUV	Fixed relative threshold method		Fixed absolute threshold method	
									Superior border	Inferior border	Superior border	Inferior border
1	M	52	T3	+	Middle thoracic	Arm abducted	7.5	26.5	10%	16%	2.8	4.6
2	F	69	T1b	-	Cervical	Supine†	3.0	10.6	27%	70%	2.9	7.4
3	M	74	T2	-	Upper thoracic	Supine	2.4	13.6	24%	53%	3.3	7.3
4	M	71	T1b	-	Upper thoracic	Arm abducted	3.0	6.5	59%	36%	3.8	2.4
5	M	67	T3	+	Lower thoracic	Arm abducted	8.1	13.0	51%	69%	6.7	9.0
6	M	74	T3	+	Middle thoracic	Arm abducted	7.2	28.6	13%	8%	3.7	2.5
7	M	80	T1b	-	Upper thoracic	Arm abducted	0.9	7.1	84%	80%	6.0	5.7
8	F	66	T2	+	Cervical	Supine	5.7	23.6	7%	11%	1.7	2.8
9	M	79	T3	+	Lower thoracic	Arm abducted	4.5	16.4	16%	44%	2.7	7.3

Abbreviations: LN – lymph node, GTV – gross tumor volume, SUV – standardized uptake value.

† The patients had a PET scan in the supine position.

2. Thresholds corresponding to the GTV borders

The thresholds corresponding to the superior and inferior GTV borders are summarized in the right columns of Table 1. The median threshold value for the fixed relative threshold method was 24% for the superior border, and 44% for the inferior border. The median threshold value of all threshold values was 32%. No significant difference was observed in the threshold values by location ($p = 0.159$). The median threshold value using the fixed absolute threshold method was 3.3 for superior borders and 5.7 for inferior borders. The median threshold value of all threshold values was 3.8. No significant difference in the threshold value was observed by location ($p = 0.061$). The distribution of the threshold values is summarized in Figure 3. The CV was 0.781 for every threshold value from the fixed relative threshold method, and it was 0.400 for the fixed absolute threshold method. Thus, the fixed absolute threshold method showed more consistent results.

Several thresholds for delineating MTV for esophageal cancer have been suggested, and a fixed absolute threshold of SUV 2.5 is one of them [12]. All but two superior and inferior GTV borders were included by MTV with a SUV threshold of 2.5. Threshold values from the patients with a maximum SUV > 20 had a tendency to be closer to 2.5. The median threshold of patients with a maximum SUV > 20 was 2.8, while the median threshold of the others was 3.6. Nevertheless, no significant difference was observed between threshold values with a maximum SUV of 20 ($p = 0.069$).

Linear regression was performed for the fixed relative and absolute thresholds to check the correlation between the threshold values and the continuous covariates, which were the longitudinal length of GTV and the maximum SUV. Figure 4 illustrates dot plots from these analyses, and a significant linear regression model was constructed for the fixed absolute thresholds and both continuous covariates. Significant differences in fixed relative thresholds were observed by T stage ($p = 0.038$), lymph node positivity ($p = 0.011$), and location of the tumor ($p = 0.044$). No significant difference in the fixed absolute thresholds or these covariates were found.

Figure 3. Histogram of (A) fixed relative threshold, (B) fixed absolute threshold corresponding to the gross tumor volume (GTV) borders.

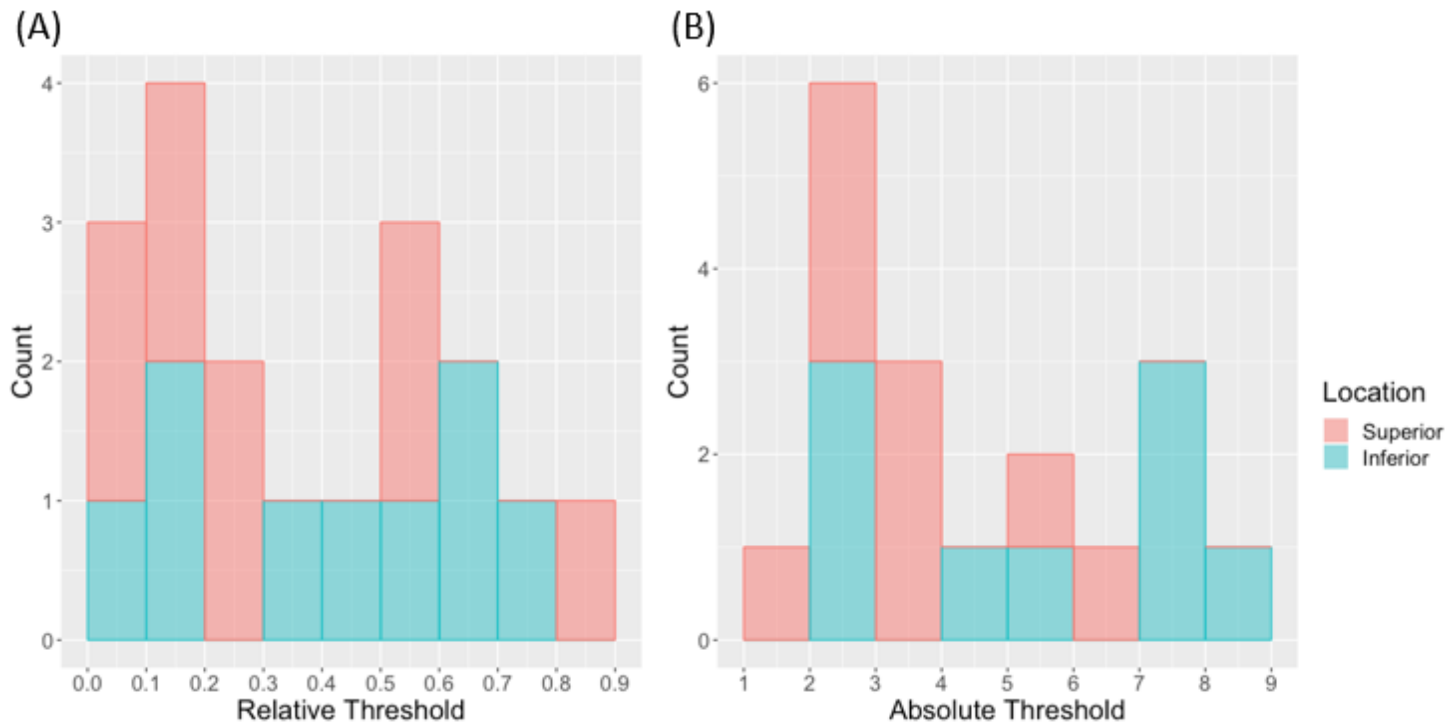
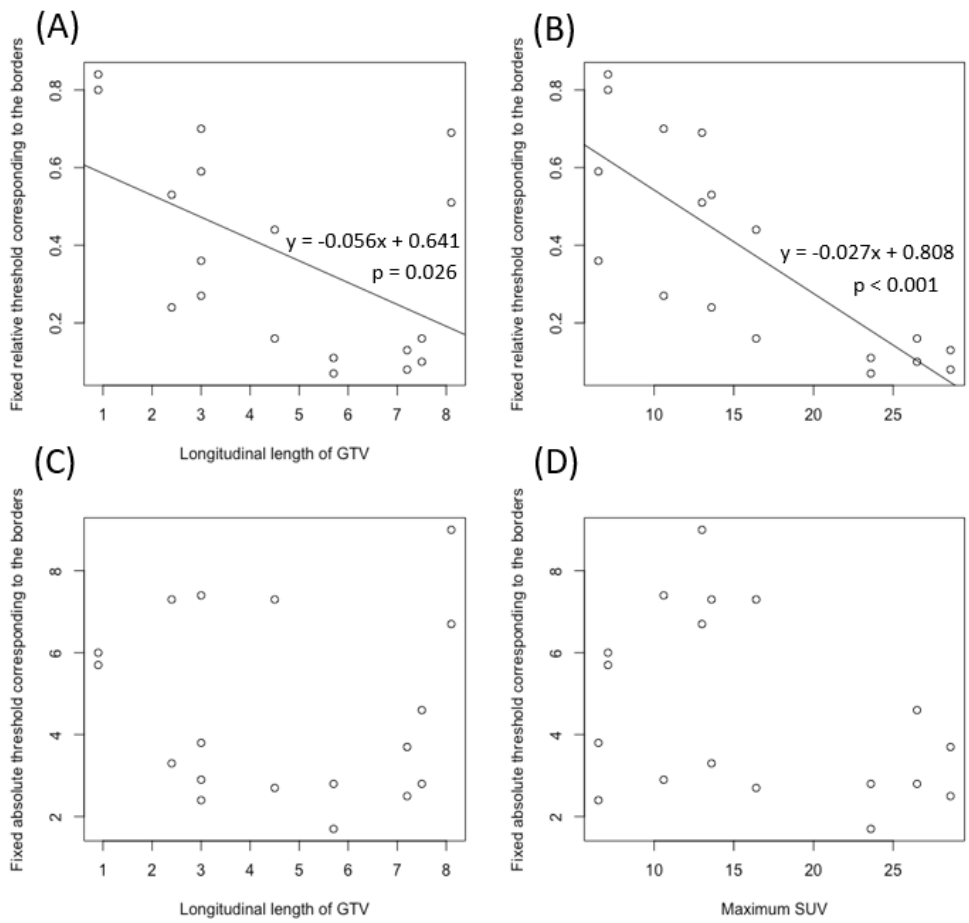


Figure 4. Dot plot of (A) fixed relative thresholds corresponding to the gross tumor volume (GTV) borders and GTV longitudinal length, (B) fixed relative thresholds and the maximum standardized uptake value (SUV), (C) fixed absolute thresholds and longitudinal length of GTV, (D) fixed absolute thresholds and maximum SUV. A significant linear regression model was constructed in plots (A) and (B).



3. Concordance index analysis

The calculated CIs of endoscopically defined GTV and PET-based MTV defined by the fixed relative threshold are summarized in Table 2 and Figure 5A. The highest median CI was 0.850 with a threshold of 10%, but MTV was only delineated in three patients using this threshold, as background hypermetabolism was too high to determine MTV with the low threshold in the other six patients. All three patients had a maximum SUV > 20. Seven patients had their optimal relative threshold as their lowest relative threshold that MTV could be delineated, and the range of their optimal threshold was 10% to 40%. Five of nine patients had their highest CI > 0.8, indicating good concordance between GTV and optimal MTV regardless of the threshold value.

Linear regression was performed for the highest CI and the optimal thresholds from each patient to check the correlation between these values and the continuous covariates. Figure 6 illustrates dot plots and a significant linear regression model. A significant correlation was found only between the optimal threshold of each patient and maximum SUV, indicating that the lower fixed relative threshold would be suitable for patients with a higher maximum SUV. No significant differences were observed in the highest CI or optimal threshold by T stage, lymph node positivity, or location of the tumor.

The calculated CIs of GTV and MTV defined by the fixed absolute threshold are summarized in Table 3 and Figure 5B. The highest median CI was

0.794 with a threshold of 3.0, and three patients had their optimal MTV with this threshold. Four patients achieved a CI > 0.8 with a threshold of 3.0. Six of the nine patients had their highest CI > 0.8, indicating good concordance between GTV and optimal MTV regardless of the threshold value. All but one patient had an optimal threshold ≥ 3.0, indicating that although only three patients had their optimal MTV with this threshold, a threshold with a SUV of 3.0 could cover the optimal MTV of eight of the nine patients.

Linear regression was performed to check the correlations, and Figure 7 illustrates the results. No significant correlation was detected for the longitudinal length of GTV or maximum SUV. Also, no significant differences were observed in the highest CI or the optimal threshold by T stage, lymph node positivity, or location of the tumor.

Table 2. Calculated concordance indices for the fixed relative threshold method and adaptive method

Patient number	Threshold relative to maximum SUV						
	10%	20%	30%	40%	50%	60%	70%
1	0.862	0.840	0.760	0.720	0.720	0.680	0.640
2	*	*	0.692	0.545	0.600	0.500	0.500
3	*	0.667	0.600	0.556	0.556	0.500	0.500
4	*	*	*	0.900	0.700	0.700	0.500
5	*	0.771	0.871	0.931	0.964	0.963	0.778
6	0.733	0.667	0.583	0.542	0.500	0.458	0.458
7	*	*	*	0.500	0.600	1.000	1.000
8	0.850	0.737	0.737	0.632	0.579	0.579	0.263
9	*	0.765	0.688	0.733	0.600	0.600	0.467
Median	0.850	0.751	0.692	0.632	0.600	0.600	0.500

* Metabolic tumor volume could not be delineated as hypermetabolic uptake background was higher than the threshold.

The highest concordance indices of each patient for fixed relative threshold method are underlined.

Abbreviations: SUV – standardized uptake value.

Table 3. Calculated concordance indices for the fixed absolute threshold method

Patient number	Absolute SUV value of the threshold								
	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0
1	0.500	0.658	0.828	<u>0.852</u>	0.846	0.846	0.840	0.840	0.840
2	*	0.358	0.600	0.583	<u>0.636</u>	0.545	0.545	0.600	0.600
3	*	0.421	<u>0.636</u>	0.545	0.600	0.600	0.500	0.556	0.556
4	*	0.818	<u>0.900</u>	0.700	0.600	0.500	0.500	0.500	0.200
5	0.643	0.771	0.794	0.818	0.871	0.931	0.931	<u>0.964</u>	<u>0.964</u>
6	0.522	0.667	<u>0.821</u>	0.815	0.792	0.667	0.667	0.667	0.667
7	*	0.300	0.600	0.600	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	<u>1.000</u>	0.667
8	0.750	<u>0.850</u>	0.842	0.842	0.789	0.789	0.737	0.737	0.737
9	*	0.600	0.684	<u>0.750</u>	<u>0.750</u>	0.688	0.688	0.733	0.733
Median	0.582	0.658	0.794	0.750	0.789	0.688	0.688	0.733	0.667

* Metabolic tumor volume could not be delineated as background hypermetabolic uptake was higher than the threshold.

Highest concordance indices of each patient are underlined.

Abbreviations: SUV – standardized uptake value.

Figure 5. Box plot of concordance indices from the (A) fixed relative threshold analysis, (B) the fixed absolute threshold analysis.

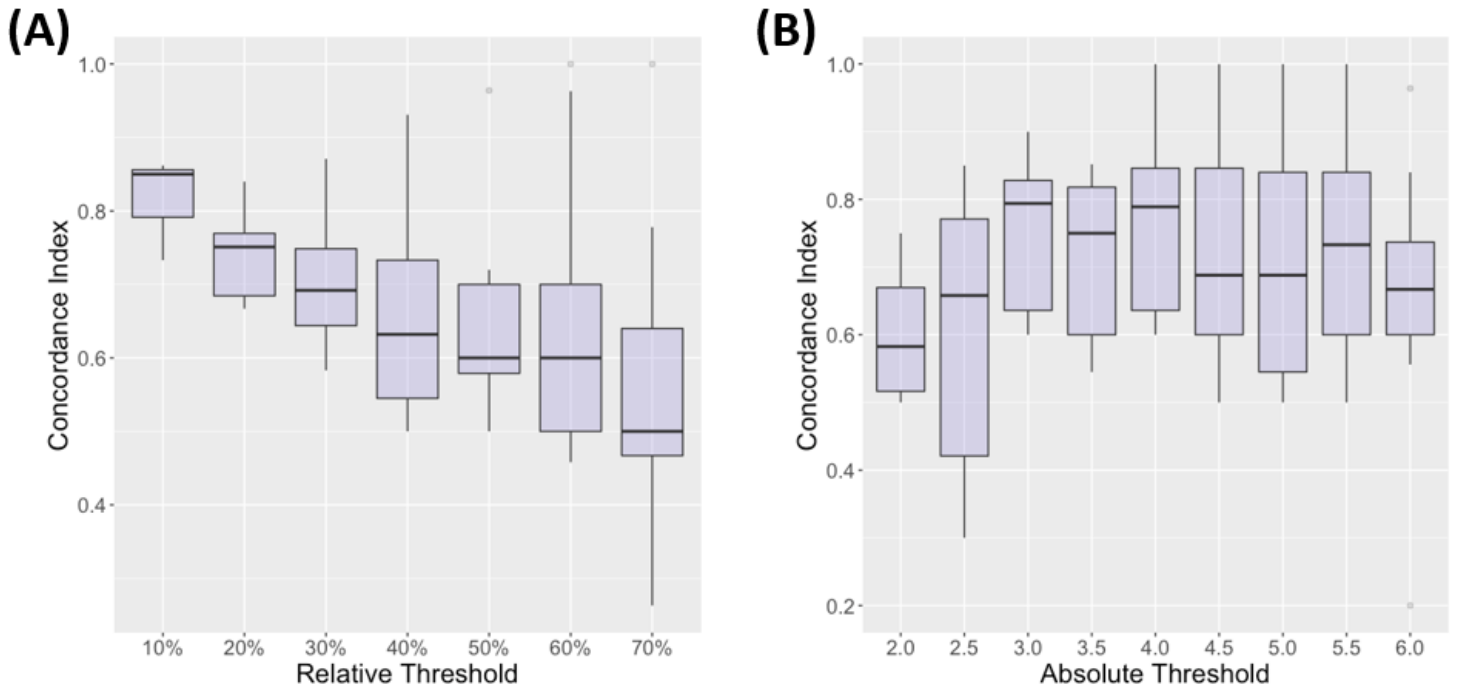


Figure 6. Dot plot of (A) highest concordance index (CI) of each patient from the fixed relative threshold method and longitudinal length of gross tumor volume (GTV), (B) highest CI and maximum standardized uptake value (SUV), (C) potential optimal threshold of each patient from the fixed relative threshold method and longitudinal length of GTV, (D) optimal threshold and maximum SUV. A significant linear regression model was constructed in plot (D).

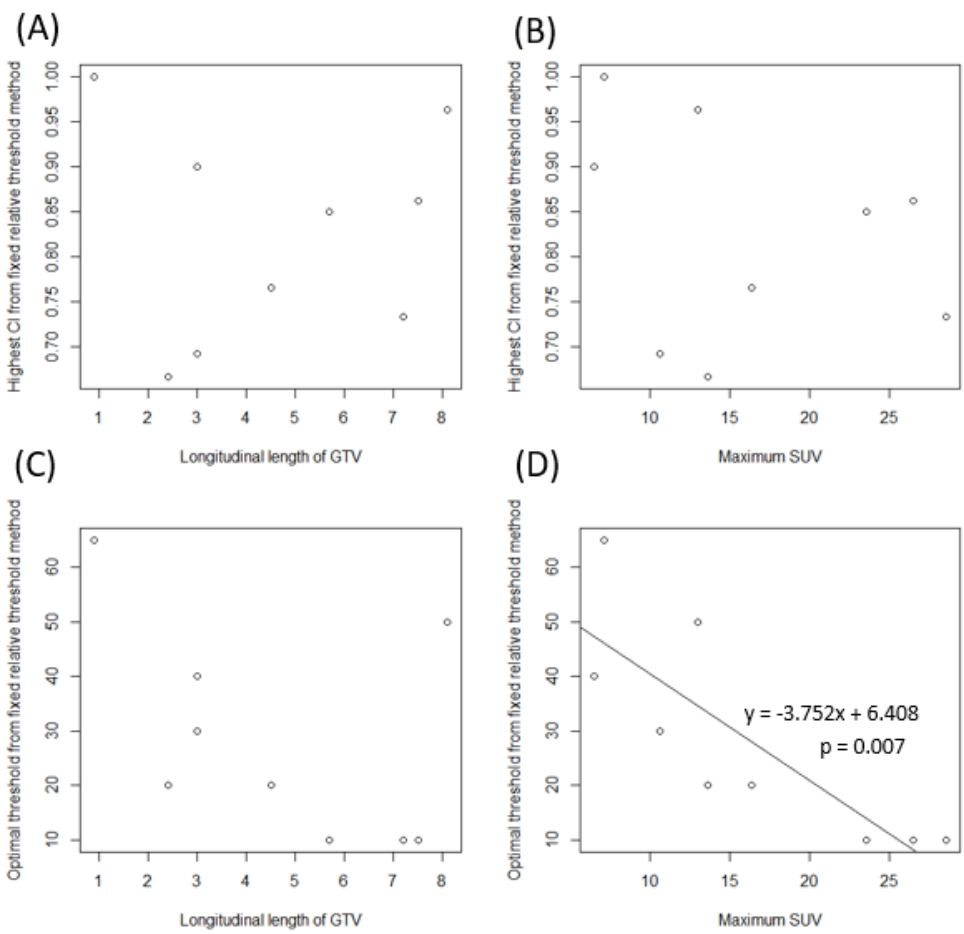
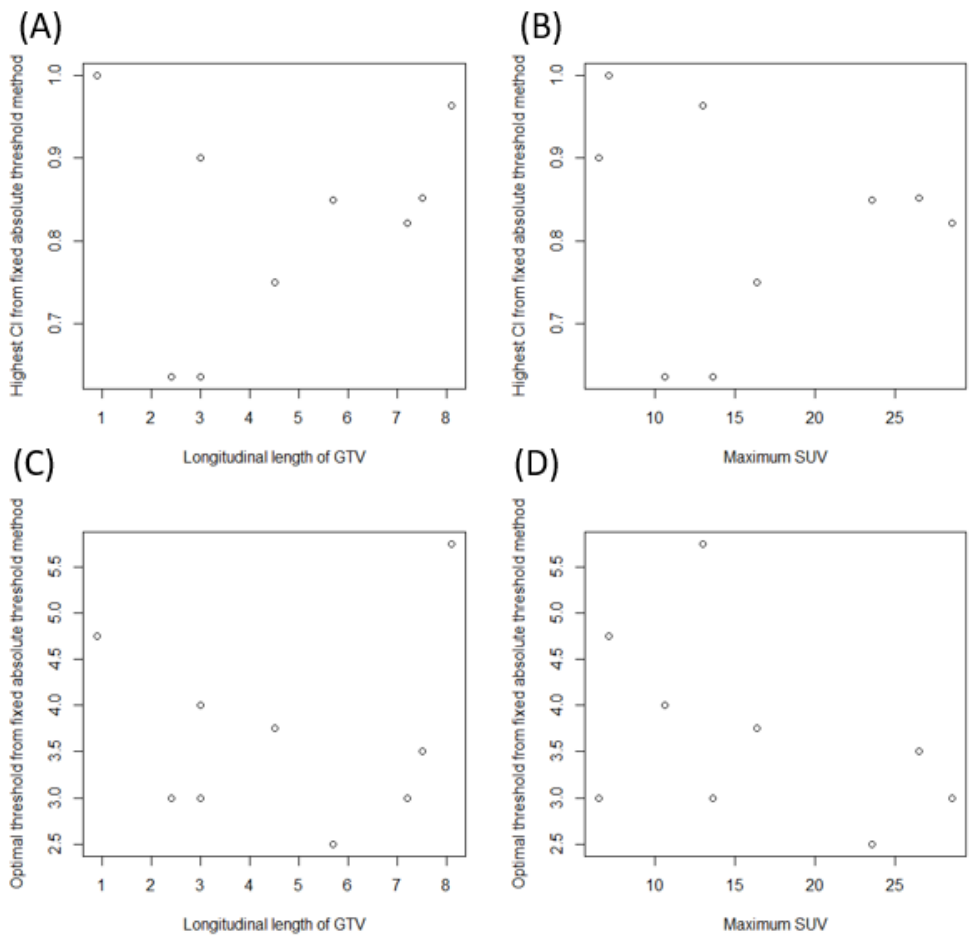


Figure 7. Dot plot of (A) highest concordance index (CI) of each patient from the fixed absolute threshold method and longitudinal length of gross tumor volume (GTV), (B) highest CI and maximum standardized uptake value (SUV), (C) potential optimal threshold of each patient from the fixed absolute threshold method and longitudinal length of GTV, (D) optimal threshold and maximum SUV.



4. Comparison of CI analysis and linear approximation

Optimal thresholds for each patient obtained by both threshold methods based on CI are summarized in the left column of Table 4. The CV was lower for the fixed absolute threshold method, indicating that this method produced more consistent results than the fixed relative threshold method. Optimal thresholds for each patient based on linear approximation are summarized in the right column of Table 4, and the results were generally similar with the CI analysis. Both the CI analysis and linear approximation determined a median optimal threshold from the fixed absolute threshold method of 3.5, while the median optimal threshold from the fixed relative threshold method was slightly different. The CV was lower for the fixed absolute threshold method than that for the fixed relative threshold method in the linear approximation.

Table 4. Comparisons of different methods to obtain the optimal threshold

Patient number	Optimal threshold obtained based on concordance index		Optimal threshold obtained based on linear approximation	
	Fixed relative threshold method	Fixed absolute threshold method	Fixed relative threshold method	Fixed absolute threshold method
1	10%	3.5	15%	3.5
2	30%	4.0	34%	3.4
3	20%	3.0	30%	4.25
4	40%	3.0	35%	2.5
5	50%	5.75	55%	6.5
6	10%	3.0	13.3%	3.58
7	65%	4.75	65%	4.75
8	10%	2.5	7.5%	2.4
9	20%	3.75	20%	3.38
Median	20.0%	3.5	30%	3.5
Average	28.3%	3.69	30.5%	3.81
SD	0.197	1.021	0.194	1.251
CV	0.695	0.276	0.634	0.329

Abbreviations: SD – standard deviation, CV – coefficient of variation

DISCUSSION

PET/CT has been integrated into RT planning and utilized by many clinicians. As superficial tumors often cannot be detected by PET [8], PET-based tumor borders may not be accurate if the tumor tends to spread superficially. Konski et al. compared the lengths of esophageal primary tumors measured by CT scan, PET, and endoscopic ultrasonography and found no significant difference between tumor lengths measured by PET and endoscopic ultrasonography, while tumor lengths measured by CT scan were significantly longer than those measured by PET [12]. This previous study revealed the superiority of implementing PET/CT to delineate GTV, but the analysis was only based on measured tumor length, and no actual geometric comparison of hypermetabolism appeared on PET/CT, which is different from the current study.

The delineation of MTV has been studied primarily in lung cancer, and various segmentation methods and cut-off values for thresholds have been used [13]. An early study by Erdi et al. showed that a fixed relative threshold of 36–44% was well-correlated with lung lesion volumes > 4 mL [14]. Also, Yu et al. examined PET images and lobectomy specimens of 15 non-small-cell lung cancer patients, and proposed that a relative threshold of 31% and an absolute threshold of 3.0 would be appropriate to define the target volume [10]. Previous studies on esophageal cancer used similar absolute cut-off SUV values to distinguish benign lesions from

malignant lesions. Konski et al. used a SUV of 2.5 to determine the extent of esophageal tumors by PET [12]. We propose that a fixed absolute threshold of SUV 2.5 is sufficient to set MTV, and this is consistent with other reports that proposed SUV thresholds of 2.5–3.0 [10, 15]. The current study also shows that not every lesion had a threshold corresponding to the borders close to a SUV of 2.5. This result is consistent with the report of Biehl et al. which concluded that no single threshold could be determined to define lung lesion volume, compared with GTV based on a CT scan [11]. We propose that esophageal lesions with a maximum SUV > 20 are more suitable for MTV with a threshold SUV of 2.5. More studies are needed to suggest a feasible threshold for every esophageal tumor.

Endoscopy is widely accepted as useful tool to access esophageal tumors, particularly those with superficial spread. Previous studies have shown that placing fiducial markers can be very helpful to interpret endoscopic findings for volume delineation [5, 6]. In this present study, GTV was delineated solely based on the location of fiducial markers appearing on the simulation CT scan, and various MTV delineation methods were compared with this endoscopically defined GTV. These methods are based on the assumption that a fiducial marker placed by endoscopy accurately reflects mucosal spread of an esophageal tumor, as the above-mentioned study stated. Geometric differences in MTV and GTV were analyzed using fiducial markers, and the result showed discordance between GTV based on the endoscopic fiducial markers and MTV in some patients.

Contrary to previous studies from other groups, endoscopic hemoclips, which are used for hemostasis, were used as fiducial markers in the present study. Hemoclips are useful as fiducial markers because they are readily available in many medical centers. When the simulation CT scan was conducted after placing the hemoclips while maintaining nil per os, the hemoclips were stable enough to locate the esophageal tumors on the simulation CT scan. As most gastroenterologists are familiar with hemoclips, we expect that this procedure could be accurately performed to locate esophageal tumors. Nevertheless, there is no previous confirmation for hemoclips to be as precise as dedicated fiducial marker. Also, no validation for how exactly translate the appearance of hemoclips in simulation CT to target volume has been made, and there is no standard for the process. The hemoclips as fiducial markers on the CT scan usually appeared on three or four axial planes with an axial cut thickness of 3.0 mm, and this could add inconsistency of GTV delineation. Furthermore, hemoclips were not stable enough for the pathologic examination, as no hemoclip was left in a surgical specimen in the present study, contrary to dedicated fiducial markers [16]. These potential concerns about hemoclips require further examinations.

The concept of a CI is mostly used for volumetric comparisons in radiation oncology. As mentioned in a previous section, it is a measure of the overlap of different volumes. A CI of 1 means the structures are perfectly overlapped, while 0 means no overlap is present [9]. In this study, the measured volumes of the structures were substituted by the longitudinal length of the volumes, considering the patterns

of local spread of esophageal tumors and the different delineation principles applied to GTV and MTV. As previously stated, CI analysis and the longitudinal length of the volumes produced similar results with linear approximation, which was used in previous studies. Compared with linear approximation, which only considers longitudinal length of tumor volumes, the CI analysis included the actual location and relationship of the tumor volumes. A low CI means discordance between two structures, but does not indicate why this discordance happens, which is a shortcoming of the CI. Thus, the CI value is useful to compare the amount of concordance but may have limited clinical implications. Also, there is no CI limit that is generally agreed to be clinically significant. In this study, a CI > 0.8 was defined as good concordance, and a CI of 0.8 means there was an overestimate or an underestimate as large as 20% for the union of two longitudinal lengths, though this definition of good concordance based on CI is arbitrary.

Both CI analysis and linear approximation resulted in median optimal absolute threshold of 3.5. In finding MTV thresholds corresponding to the borders of GTV, we proposed that fixed absolute threshold of SUV 2.5 would encompass most of GTV borders, and we also emphasized that this result is consistent with some previous studies. SUV 3.5 had better result for making MTV to resemble endoscopically defined GTV, as showed in CI analysis and linear approximation, which are primarily volume comparisons. Though in treatment perspective, not missing the actual tumor is important for determination of target volumes. Therefore, SUV 2.5 would be more clinically relevant for delineation purpose.

Although there is a study which showed pathologic tumor spread and borders defined by dedicated fiducial markers were correlated well [6], and the current study was based on the assumption derived from the previous study, no confirmation for using hemoclips as fiducial markers has been made. Direct comparison of pathologic features of resected tumor and tumor borders defined by hemoclips is not feasible, due to low stability of hemoclips. Instead, constructing a validation cohort consists of new or previously excluded patients, or checking failure patterns of patients who had radiotherapy plan based on target volumes defined by hemoclips may be helpful to confirm the conclusion of the present study. Also, to identify how using hemoclips as fiducial markers can affect the actual planning process, calculating inter- or intra-observer variability from GTV without fiducial marker and GTV with hemoclips as fiducial markers would be feasible.

The present study had some limitations. This study was conducted retrospectively and PET/CT scans were not intended to be fused with a simulation CT scan. Therefore, geographic differences existed between PET scan and simulation CT scan and could have influenced the analyses. The longitudinal resolution of the simulation CT scan was too low to have sufficient accuracy for the longitudinal length comparisons. The current study was based on assumption that PET scan and endoscopy would represent pathologic GTV, but subclinical disease or regional metastasis can also influence hypermetabolism on PET scan. The patients in this study might be less representative due to the small number of patients. Nevertheless, this study hypothesized that integrating the placement of endoscopic

fiducial markers was helpful to delineate esophageal GTV.

In conclusion, the fixed absolute threshold method resulted in more consistent threshold values than the fixed relative threshold method, and the fixed relative threshold method tended to be more influenced by other properties of the tumor. A SUV of 2.5, which was validated in previous studies, allowed the MTV to include all but two borders of the GTV, but thresholds corresponding to the GTV borders tended to be higher for esophageal tumors with maximum SUVs < 20. We proposed that a SUV of 2.5 is more suitable for esophageal lesions with a maximum SUV of 20. Some discordance between PET-based MTV and endoscopy-based GTV was detected by the CI analysis, and integrating endoscopic features by using fiducial markers may be useful for the correction. Additional studies with a larger patient population and various other MTV delineation methods are warranted.

REFERENCES

1. Stahl M, Walz MK, Stuschke M, et al. Phase III comparison of preoperative chemotherapy compared with chemoradiotherapy in patients with locally advanced adenocarcinoma of the esophagogastric junction. J Clin Oncol. 2009;27:851-856.

2. Shapiro J, van Lanschot JJB, Hulshof MCCM, 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;16:1090-1098.

3. Hazard L, Yang G, McAleer MF, Hayman J, Willett C. Principles and techniques of radiation therapy for esophageal and gastroesophageal junction cancers. J Natl Compr Cancer Netw. 2008;6:870-878.

4. Wang BY, Liu CY, Lin CH, et al. Endoscopic tumor length is an independent prognostic factor in esophageal squamous cell carcinoma. Ann Surg Oncol. 2012;19:2149-2158.

5. Machiels M, Jin P, van Hooft JE, et al. Reduced inter-observer and intra-observer delineation variation in esophageal cancer radiotherapy by use of fiducial markers. Acta Oncol. 2019;58:943-950.

6. Machiels M, Van Montfoort ML, Thuijs NB, et al. Microscopic tumor spread

beyond (echo)endoscopically determined tumor borders in esophageal cancer. Radiat Oncol. 2019;14:1-8.

7. Vesprini D, Ung Y, Dinniwell R, et al. Improving Observer Variability in Target Delineation for Gastro-oesophageal Cancer-the Role of 18Ffluoro-2-deoxy-d-glucose Positron Emission Tomography/Computed Tomography. Clin Oncol. 2008;20:631-638.

8. Kato H, Kuwano H, Nakajima M, et al. Comparison between positron emission tomography and computed tomography in the use of the assessment of esophageal carcinoma. Cancer. 2002;94:921-928.

9. Jameson MG, Holloway LC, Vial PJ, Vinod SK, Metcalfe PE. A review of methods of analysis in contouring studies for radiation oncology. J Med Imaging Radiat Oncol. 2010;54:401-410.

10. Yu J, Li X, Xing L, et al. Comparison of Tumor Volumes as Determined by Pathologic Examination and FDG-PET/CT Images of Non-Small-Cell Lung Cancer: A Pilot Study. Int J Radiat Oncol Biol Phys. 2009;75:1468-1474.

11. Biehl KJ, Kong FM, Dehdashti F, et al. 18F-FDG PET definition of gross tumor volume for radiotherapy of non-small cell lung cancer: Is a single standardized uptake value threshold approach appropriate? J Nucl Med. 2006;47:1808-1812.

12. Konski A, Doss M, Milestone B, et al. The integration of 18-fluoro-deoxy-glucose positron emission tomography and endoscopic ultrasound in the

treatment-planning process for esophageal carcinoma. Int J Radiat Oncol Biol Phys. 2005;61:1123-1128.

13. Im HJ, Bradshaw T, Solaiyappan M, Cho SY. Current Methods to Define Metabolic Tumor Volume in Positron Emission Tomography: Which One is Better? Nucl Med Mol Imaging. 2018;52:5-15.

14. Erdi YE, Mawlawi O, Larson SM, et al. Segmentation of lung lesion volume by adaptive positron emission tomography image thresholding. Cancer. 1997;80:2505-2509.

15. Paulino AC, Johnstone PAS. FDG-PET in radiotherapy treatment planning: Pandora's box? Int J Radiat Oncol Biol Phys. 2004;59:4-5.

16. Fernandez DC, Hoffe SE, Barthel JS, et al. Stability of endoscopic ultrasound-guided fiducial marker placement for esophageal cancer target delineation and image-guided radiation therapy. Pract Radiat Oncol. 2013;3:32-39.

국문 초록

식도암에서의 내시경적으로 결정된 육안적 종양 체적과 대사성 종양 체적의 비교

목적: 본 연구에서는 식도암에서 내시경적으로 정의된 육안적 종양 체적과 양전자 방사 단층촬영에 기반한 대사성 종양 체적의 세로 길이를 비교하고, 육안적 종양 체적을 기준으로 하여 대사성 종양 체적을 결정하는 최적의 방법을 찾아보고자 하였다.

대상환자 및 방법: 식도의 편평세포암에 대해서 내시경적으로 위치표지자를 설치하였던 9명의 환자의 의료기록을 후향적으로 분석하였다. 내시경적 지혈클립을 위치표지자로 사용하여, 내시경적으로 보이는 병변의 위쪽 경계와 아래쪽 경계에 위치시켰다. 육안적 종양 체적은 위치표지자의 위치에만 근거하여 새로 설정하였다. 육안적 종양 체적의 위쪽 및 아래쪽 경계에 해당하는 표준섭취계수의 한계치는 각각의 육안적 종양 체적의 경계에 대사성 종양 체적이 닿도록 하는 가장 높은 한계치로 정의하였다. 고정 상대 한계법과 고정 절대 한계법이 모두 사용되었다. 두 방법으로 구해진 한계치 값에서 계산된 변동계수를 비교하여, 어떠한 방

법이 육안적 종양 체적의 경계에 해당하는 한계치를 결정하는데 있어서 좀 더 일관된 결과를 얻을 수 있는지 결정하였다.

결과: 고정 상대 한계치와 고정 절대 한계치의 중간값은 각각 32%와 3.8이었다. 변동계수는 고정 상대 한계법에 대해서 0.781, 고정 절대 한계법에 대해서 0.400이었으며, 이는 고정 절대 한계법이 더 일관된 결과를 내놓았다는 것을 의미한다. 육안적 종양 체적의 경계는 두 개를 제외하곤 표준섭취계수 한계치 2.5에서의 대사성 종양 체적에 포함되었는데, 이 한계치는 이전 다른 연구에서도 사용된 바가 있다. 최대 표준섭취계수가 20을 넘는 식도 종양은 육안적 종양 체적의 경계에 해당하는 한계치 값이 2.5에 보다 가까운 경향이 있었다 (중간값 2.8 대 3.6, p = 0.069).

결론: 식도 종양에서 대사성 종양 체적의 한계치를 결정하는데 있어서, 고정 절대 하계법이 고정 상대 한계법에 비해서 좀 더 적합한 것으로 결론지었다. 표준섭취계수 2.5는 최대 표준섭취계수가 20을 넘는 식도 종양에서 적합한 것으로 나타났다. 모든 식도 종양에서 있어서 적용 가능한 한계치를 찾는데에 있어서 추가적인 연구가 필요하다.

주요어: 식도암, 양전자 방사 단층촬영, 내시경, 종양 체적

학번: 2018-29944