Gross tumour delineation on computed tomography and positron emission tomography-computed tomography in oesophageal cancer: a nationwide study
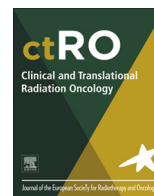
Nowee, M.E.; Voncken, F.E.M.; Kotte, A.N.T.J.; Goense, L.; Rossum, P.S.N. van; Lier, A.L.H.M.W. van; ... ; Dutch Natl Platform Radiotherapy

Citation

Original Research Article

Gross tumour delineation on computed tomography and positron emission tomography-computed tomography in oesophageal cancer: A nationwide study

M.E. Nowee [a,*], F.E.M. Voncken [a], A.N.T.J. Kotte [b], L. Goense [b,c], P.S.N. van Rossum [b], A.L.H.M.W. van Lier [b], S.W. Heijmink [d], B.M.P. Aleman [a], J. Nijkamp [a], G.J. Meijer [b], I.M. Lips [b,2], Dutch National Platform for Radiotherapy of Gastrointestinal Tumours (LPRGE) group [1]

[a] Department of Radiation Oncology, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, the Netherlands
[b] Department of Radiation Oncology, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands
[c] Department of Surgery, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands
[d] Department of Radiology, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, the Netherlands

ABSTRACT

Background and purpose: Accurate delineation of the primary tumour is vital to the success of radiotherapy and even more important for successful boost strategies, aiming for improved local control in oesophageal cancer patients. Therefore, the aim was to assess delineation variability of the gross tumour volume (GTV) between CT and combined PET-CT in oesophageal cancer patients in a multi-institutional study.
Materials and methods: Twenty observers from 14 institutes delineated the primary tumour of 6 cases on CT and PET-CT fusion. The delineated volumes, generalized conformity index (CIgen) and standard deviation (SD) in position of the most cranial/caudal slice over the observers were evaluated. For the central delineated region, perpendicular distance between median surface GTV and each individual GTV was evaluated as in-slice SD.
Results: After addition of PET, mean GTVs were significantly smaller in 3 cases and larger in 1 case. No difference in CIgen was observed (average 0.67 on CT, 0.69 on PET-CT). On CT cranial-caudal delineation variation ranged between 0.2 and 1.5 cm SD versus 0.2 and 1.3 cm SD on PET-CT. After addition of PET, the cranial and caudal variation was significantly reduced in 1 and 2 cases, respectively. The in-slice SD was on average 0.16 cm in both phases.
Conclusion: In some cases considerable GTV delineation variability was observed at the cranial-caudal border. PET significantly influenced the delineated volume in four out of six cases, however its impact on observer variation was limited.

© 2018 The Authors. Published by Elsevier B.V. on behalf of European Society for Radiotherapy and Oncology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Oesophageal cancer is the eighth most common malignancy in the world and the incidence is rising [1]. The addition of neoadjuvant chemoradiation (nCRT) to surgery in the curative treatment of oesophageal cancer patients has resulted in downstaging, more radical resections and a survival benefit [2,3]. Definitive chemoradiation (dCRT) is the treatment of choice for inoperable or irre-

---

sectable oesophageal cancer patients. Loco-regional recurrences are still considerable (30–60%) and mainly occur at the site of the primary tumour [4–9].

However, evidence is rising that a simultaneous-integrated boost to gross primary disease in locally advanced oesophageal cancer may improve local control for patients with unresectable cancer [10]. Therefore, prospective dose-escalation trials are presently enrolling patients for dose intensification to the primary tumour aiming for improved local tumour control [11,12]. The Dutch national trial is randomizing between the standard dose of 50.4 Gy and an experimental dose of 61.6 Gy in 28 fractions to the primary tumour [11].

Accurate delineation of the primary tumour is vital to the success of radiotherapy and even more important for successful boost strategies. Moreover, when the results of dose-escalation studies are interpreted, it is essential to be aware of the interobserver variation of the GTV delineation of the primary tumour.

Furthermore, for patients with a pathologic complete response after nCRT, there is on-going debate if those patients can be treated with organ sparing strategies [13]. Improving GTV definition and thus increasing the chance of a complete response is essential for the success of organ sparing strategies. Therefore, for both dCRT and for nCRT patients, accurate delineation of the gross macroscopic tumour is essential.

For years, computed tomography (CT) has been used for delineation of oesophageal target volumes during radiotherapy treatment planning. However, determination of the proximal and distal extension of the primary tumour is often difficult on CT due to poor soft tissue contrast. Assessment of oesophageal tumour length by CT might overestimate the extension compared to histopathology [14]. Endoscopic Ultrasound (EUS) seems more accurate in assessing the longitudinal oesophageal tumour extension [15]. However, translation of EUS findings into the RT planning process is difficult.

Endoscopy-guided clipping or fiducial insertion at the tumour borders might serve as a valid option for GTV definition, although this adds an additional invasive procedure [16].

Fluorodeoxyglucose (FDG) uptake is seen in 68–100% of oesophageal tumours [17]. FDG-PET is superior to CT in detecting nodal and distant metastases [18], therefore adding value to the (re-) staging of oesophageal tumours [19].

In clinical practice, fused PET-CT images are often used for oesophageal tumour delineation and this seems to improve accuracy in target contouring. However, previous studies are mostly single-center studies or assessed a limited number of observers [20–27].

The aims of this study were to evaluate the interobserver delineation variation of the primary oesophageal tumour in daily clinical practice in The Netherlands, and to study the effects of the addition of FDG-PET to CT images on interobserver variability in a nationwide multi-institutional setting.

2. Materials and methods

2.1. Patients and observers

All 21 radiotherapy institutes in the Netherlands were approached to participate. Imaging of six locally advanced oesophageal cancer patients was selected from a previous study, in which the response of oesophageal tumours to nCRT was evaluated with different imaging modalities. This study was approved by the local medical ethics committee and written informed consent was obtained from all patients (NCT 02125448). Cases were selected based on availability of a planning PET-CT in radiotherapy position, tumour location according to the American Joint Committee on Cancer (AJCC) classification [28], and histology. To represent daily

clinical practice, we included cases equally spread along the oesophagus, and both squamous cell and adenocarcinoma cases (Table 1).

2.2. FDG-PET/CT scans

All patients underwent a planning 18F-FDG-PET/CT scan in radiotherapy treatment position between December 2013 and December 2014. Patients had to stay sober for at least six hours before injection of 18F-FDG, and blood glucose levels were measured for potential hyperglycemia. The administered activity of intravenously administered 18F-FDG was 2.0 MBq/kg. Approximately 60 min after administration of 18F-FDG, PET and CT imaging were performed from neck to abdomen using a 18F-FDG PET/CT system (Siemens, Erlangen, Germany). Before PET acquisition a diagnostic quality Iodine contrast-enhanced CT was performed with the following settings: 120 kV, 20 mA, 0.5 s tube rotation time, pitch of 1.0, and 3.0 mm slice width. PET was performed using 3-dimensional acquisition, an axial field of view of 216 mm, and a scanning time of 3 min/bed position. 18F-FDG PET/CT data were reconstructed using iterative ordered-subsets expectation maximization for 21 subsets and 4 iterations (Gaussian filter). CT images were reconstructed with slice thickness of 3 mm. Administration of oral and intravenous contrast agents was performed in all cases.

2.3. Delineation software and study outline

DICOM files with the co-registered CT and PET data and a digital manual were sent to all observers. Basic clinical information (age, gender, histology and endoscopy-EUS report) and imaging information (CT and PET report) were provided. All observers delineated the gross tumour volumes according to institutional guidelines and used their own delineation software, preventing observer variation by facing new delineation software. GTV definition was not standardized since the aim of this study was to reveal interobserver variability in a multi-institutional setting in current clinical practice. The digital manual stated to delineate all macroscopic primary tumour, except pathological lymph nodes, conform institutional guidelines. To standardize the window-level of FDG-PET scans for tumour delineation, the mean activity in the liver was used as reference value for physiological soft tissue uptake of FDG under fasting conditions [29]. A nuclear medicine physician standardized this pre-set for all cases. In the first phase, observers delineated the GTV of the primary tumour on CT (GTV_CT), based on all available clinical and diagnostic information. Observers were provided with the FDG-PET/CT report, however were not allowed to review the non-co-registered PET/CT on a second screen. In the second phase, this delineation on CT was adjusted after co-registration and visual interpretation of the FDG-PET scan, creating a second GTV (GTV_PET).

2.4. Contour analysis

The mean and standard deviation of the delineated target volumes was calculated per patient and compared subsequently between the first phase and the second phase. To quantify the overlap between CT and FDG-PET/CT based target volumes for each patient, the generalized conformity index (CIgen) was calculated [30]. The CIgen is the sum of the common volumes between observer pairs divided by the sum of the encompassing volumes between each pair of observers. A CIgen of 1 represents perfectly overlapping structures with identical volume, location and shape. A CIgen of 0 means there is no overlap. The cranial and caudal delineation variation was calculated as the standard deviation in the position of the most proximal and distal delineated slice over

Table 1
Baseline characteristics of the six oesophageal cancer cases.

Case no.	cTNM[†]	AJCC location[*]	Histology	Male/Female	Age (yr)
1	T3N2M0	Upper thoracic	SCC	M	67
2	T3N1-2M0	Middle thoracic	SCC	M	54
3	T2N2M0	Lower thoracic, EGJ[1]	AC	M	71
4	T3N2-3M0	Lower thoracic, EGJ	AC	F	70
5	T2N0-1M0	Lower thoracic	AC	M	64
6	T3N1M0	Upper thoracic	SCC	F	74

Abbreviations: no. = number; yr = years; M = Male; F = Female; SCC = squamous cell carcinoma; AC = adenocarcinoma.
[†] Clinical tumour-node-metastasis (cTNM) stage according to 7th edition TNM classification.
[*] American Joint Committee on Cancer (AJCC) classification 2012.
[1] Lower thoracic tumour including the oesophageal-gastric junction (EGJ).

the observers [31]. For the central delineated region, designated as the GTV by all observers, a surface distance variation was calculated. The interobserver delineation variation was assessed relative to a reference contour for each patient, the median surface GTV, which was computed in three dimensions [32]. The median surface, encompassing 50% coverage of the GTVs of all observers, was subsequently sampled using 8000 equally distributed points. For each point the perpendicular distance to each delineated GTV surface was calculated. The variation of the different observers for each point on the median surface was expressed in a local observer variation (local SD). For the central delineated region, the overall observer variation was calculated for each patient as the quadratic mean of the local SD, representing in-slice SD. Analyses were performed with in-house developed software [31].

2.5. Statistical methods

Mean volumes and generalized conformity indices were compared using a pairwise *t*-test (paired 2-sided Student's *t*-test). To compare standard deviations a 2-sided *F*-test was used (Microsoft Excel). P-values lower than 0.05 were considered to be statistically significant.

3. Results

A total of 20 upper gastro-intestinal dedicated radiation oncologists from 14 different radiotherapy institutes participated in this study. The baseline characteristics of the six patients used for delineation are shown in Table 1. For case no. 6 the GTV_PET target volume was missing from 1 observer, therefore 19 observers were included in the analysis of this case.

3.1. Phase 1 – delineation on CT

Interobserver delineation variation on CT was in some cases substantial, as shown in Fig. 1A. Mean delineated volumes of the 6 individual cases ranged from 13.3 cm^3 to 65.8 cm^3 (Table 2). The CIgen on CT ranged from 0.58 to 0.76 and was on average 0.67 (Table 2).

The local surface distance variation (1SD) of the 20 observers projected on the median surface of each of the 6 cases is shown in Fig. 2. The main variation between observers on CT was found in the definition of the cranial and caudal border of the GTV. In patient 1, for example, cranial border differences up to 3.6 cm were seen between observers, resulting in a cranial delineation variation of 1.5 cm (1SD) (Table 2). Cranial delineation variation ranged between 0.3 and 1.5 cm SD with an average of 0.99 cm SD, while caudal variation ranged from 0.2 to 1.0 cm SD, with an average of 0.75 cm SD. The in-slice variation on CT in the central region, designated as the GTV by all observers, was small. The central overall
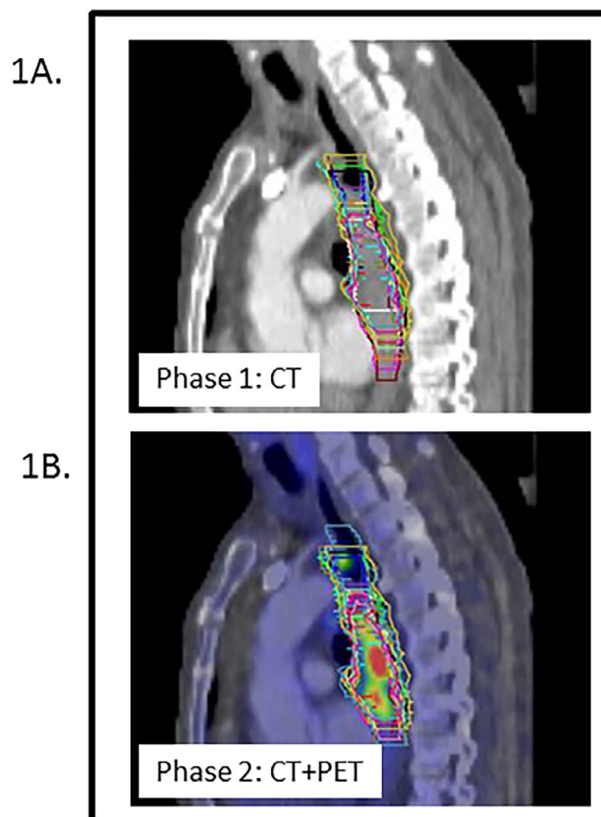
Fig. 1. Delineation variation in case 1 for 20 observers. **1A (phase 1).** The GTV of the primary tumour was delineated by 20 observers on CT, with available clinical and diagnostic information. **1B (phase 2).** After fusion of the FDG-PET scan, the GTV was adjusted based on the visual interpretation of the FDG-PET scan. The colour wash image indicates the intensity of the SUV signal. This case contained satellite lesions proximal of the tumour. Although the presence of satellite lesions was included in the clinical information, these were included in the GTV by only 9 and 11 of the 20 observers in phase 1 and 2 respectively.

SD ranged between 0.10 and 0.25 cm with an average of 0.16 cm SD (Table 2).

3.2. Phase 2-delineation on PET-CT

Large differences in observer delineations were also present in phase 2, as shown in Fig. 1B. Mean delineated volumes of the 6 individual cases ranged from 12.4 cm^3 to 76.7 cm^3 (Table 2). CIgen's of the 6 cases ranged from 0.56 to 0.77 with an average of 0.69 for FDG-PET/CT based delineations.

The local surface distance variation (1SD) of the 20 observers projected on the median surface of each of the 6 cases after addi-

Table 2
Volumetric analysis and interobserver variation in CT and FDG-PET/CT based GTV delineations of six oesophageal cancer patients for 20 observers.

Case	Mean volume in cm³		Generalized Conformity Index (CIgen)		Cranial delineation variation 1SD (cm)			Caudal delineation variation 1SD (cm)			Central overall SD (RMS, cm)		
	CT	FDG-PET/CT	p[*]	CT	FDG-PET/CT	CT	FDG-PET/CT	p[**]	CT	FDG-PET/CT	p[**]	CT	FDG-PET/CT
1	51.0	53.1	0.362	0.67	0.69	1.5	1.3	0.597	0.9	0.4	0.002	0.17	0.17
2	54.2	49.2	0.001	0.76	0.77	0.3	0.4	0.295	1.0	0.5	0.005	0.17	0.15
3	65.8	76.7	0.002	0.66	0.69	1.1	0.4	<0.001	0.7	0.8	0.526	0.12	0.13
4	52.0	45.0	0.005	0.62	0.65	1.0	0.9	0.594	0.8	0.7	0.671	0.25	0.25
5	28.3	26.8	0.402	0.58	0.56	1.1	1.0	0.792	0.6	0.8	0.165	0.10	0.09
6	13.3	12.4	0.020	0.75	0.76	0.3	0.3	0.556	0.2	0.2	0.978	0.11	0.10

Abbreviations: CT = computed tomography; FDG-PET = fluorodeoxyglucose-positron emission tomography; SD = standard deviation; CIgen = generalized conformity index (sum of the common divided by the sum of the encompassing volumes between each pair of observers). Cranial/caudal delineation variation = SD in the position of the most cranial/caudal delineated slice over the observers. Central overall SD = quadratic mean of the local SD weighted for surface (RMS = Root Mean Square) of the central part of the median surface where 100% of the observers agreed to delineate.
  [*] p < 0.05 considered statistically significant, one-sample T-test (2-tailed).
  [**] p < 0.05 considered statistically significant, 2-sided F-test.
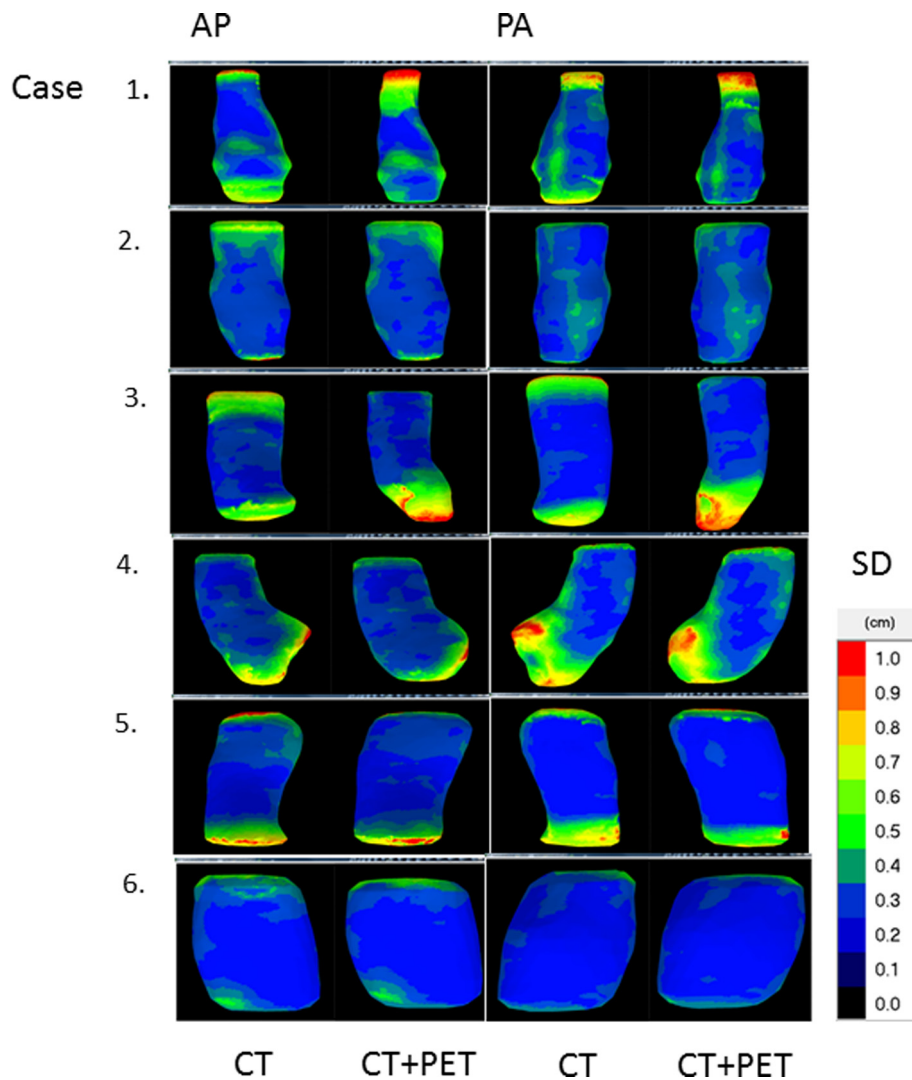
Fig. 2. Local surface distance variation (1SD) over the 20 observers projected on the median surface of the 6 patients. AP = anterior-posterior view; PA = posterior-anterior view.

tion of PET data to the CT data is shown in Fig. 2. Similar to phase 1, delineation differences between observers on CT + PET were mainly located at the cranial and caudal border. In patient 1, cranial border differences of up to 4.8 cm were observed between observers, resulting in a cranial delineation variation of 1.3 cm (1SD) (Table 2). In phase 2, cranial and caudal delineation variation ranged between 0.3–1.3 cm SD and 0.2–0.8 cm SD, with an average of 0.81 cm SD and 0.61 cm SD, respectively (Table 2). The in-slice

SD (central overall SD) was small and ranged between 0.09 and 0.25 cm SD with an average of 0.16 cm SD (Table 2).

3.3. Comparison of phase 1 and 2: delineation on CT versus on PET-CT

After addition of FDG-PET to CT, the mean GTV significantly changed in 4 out of 6 cases (Table 2). In 3 cases the GTV was reduced after addition of PET, whereas in the other case a volume increase was observed. CIgen's of the 6 cases were comparable between phase 1 and 2, as shown in Table 2 ($p = 0.14$).

The addition of FDG-PET reduced the cranial delineation variation significantly in one case from 1.1 to 0.4 cm SD (case no. 3; $p < 0.001$). In two cases the caudal delineation variation was significantly reduced with CT + PET from 0.9 to 0.4 cm SD and from 1.0 to 0.5 cm SD, respectively (case no. 1, $p = 0.002$ and case no. 2, $p = 0.005$; Table 2). Even after addition of FDG-PET cranial or caudal delineation variation remained more than 0.6 cm SD in 4 out of 6 cases, including all lower thoracic/oesophageal-gastric junction tumours (case 3, 4, 5).

For the central region, FDG-PET had no added value in-slice and the overall variation was on average 0.16 cm SD for phase 1 and phase 2 (Table 2).

4. Discussion

In this nationwide study, we demonstrated that delineation variation of the primary tumour GTV can be considerable both on CT and on PET-CT fusion, and is mainly located at the cranial and caudal border. Although the addition of FDG-PET to CT significantly impacted the delineated volume in two-third of the cases, PET did not translate into reduced observer variation at the cranial/caudal border in 50% of the cases.

Combined PET-CT imaging is often used for oesophageal tumour delineation. Previous studies showed that PET-CT imaging might impact the GTV delineations in up to 84% of patients [22,24–27]. Comparably, we found a significant change of the mean volume in 4/6 of cases, meaning that the addition of PET drives clinicians to alter their delineations substantially. In studies with surgical pathology confirmation, metabolic tumour length on PET-CT of surgically treated oesophageal carcinoma patients correlated well with histopathology [33–35], while oesophageal tumour length assessed using CT did not reflect pathological tumour length [14,15,36]. Observers have a tendency to overestimate tumour length on CT with marked inter- and intra-observer variability. In the largest study with 56 cases tumour lengths on CT were in 68% longer compared to pathology, with a mean difference of 1.67 cm [14]. In our study, surgical pathology conformation was not possible due to the small number of patients and the use of neoadjuvant chemoradiation, but mean volume on CT was significantly reduced by addition of PET in 3 cases and increased in 1 case, suggesting a similar effect.

Interobserver variability can be used as a surrogate to test the complementary value of PET to CT-delineation, assuming lower interobserver variability represents more accurate delineation. Data are scarce and conflicting. Vesprini et al. found improvement in delineation variability with the addition of PET imaging information. The SD for both GTV length and volume was taken as interobserver variability [20]. Scheurs et al. found no significant effect on interobserver variability by the addition of PET, computing an interobserver concordance index (dividing the volume intersections of one observer pair by the volume unions of that observer pair) [21]. In our study the conformity index of the 6 cases was comparable between phase 1 and 2. However, we found that the addition of PET reduced the cranial/caudal delineation variation in 3 out of 6 cases.

Remarkably, in some cases still large observer variation was present after PET-CT fusion. Cranial delineation variation was most pronounced in case 1, a patient in which gastroscopy showed satellite lesions proximal of the tumour. Although the presence of satellite lesions was included in the clinical information, these were included in the GTV by only 9 and 11 of the 20 observers in phase 1 and 2 respectively.

Instructions for delineation in this study were to delineate GTV volumes according to institutional guidelines. The aim was to clarify the amount of variation in current clinical practice. Most guidelines recommend to contour the entire circumference when contouring the GTV [37], however, individual interpretation of those guidelines might vary as shown in case 1. With regard to boost targets in dose escalation studies, GTV definition should be evident for valuable interpretation of results. Clear GTV definition may have a big effect on target volume delineation.

The two cases including the oesophageal-gastric junction (EGJ; case 3 and 4), showed substantial observer variation at the caudal border even after the addition of PET, showing lack of consensus when delineating the ingrowth of the cardia. However, PET did show an impact on the delineated volume in both cases.

A possible explanation for these large variations could be a difference when interpreting clinical and diagnostic information for delineation and also institutional differences in target definition. Agreement on the GTV definition and international GTV guidelines are warranted to overcome these large variations. With better agreement by delineation guidelines and by training of radiation-oncologists, variations in target volume dimensions could further be reduced [38].

Furthermore, we expected that PET would facilitate demarcating the caudal border of lower thoracic and EGJ tumours, where visualization on CT is poor [39]. PET might help to determine the distal borders, unfortunately gastritis can also cause FDG uptake. In our study, the caudal variation was not significantly reduced after PET addition in lower thoracic and EGJ tumours. Endoscopy-guided clipping or fiducial insertion at the distal tumour border might overcome this difficulty [16].

Additionally, co-registering PET-CT images for oesophageal target volume definition has some limitations. Firstly, not all oesophageal carcinomas are FDG avid [17]. Secondly, FDG-PET images do not provide accurate information on the external and internal contour of the tumour because of the limited spatial resolution [40]. Thirdly, target volume extension can depend on the threshold chosen and enlarged or decreased depending on the windowing. And finally, PET image acquisition covers all phases of the respiratory cycle and thus oesophageal tumour motion blurs the tumour borders [41]. Literature on 4 Dimensional (4D) CT shows main primary tumour motion for distal tumours in cranial-caudal direction [42] and the effects of tumour motion will especially play a role in PET visualization of lower tumours [43]. The use of 4D CT and 4D PET-CT might be worthwhile exploring in future research.

Because of these limitations it has been suggested that increased uptake should be used only for tumour localization and not to define precise boundaries. Threshold based delineation might overcome these shortcomings in the future. Different threshold methods have been evaluated, e.g. an isocontour of the standardized uptake volume (SUV) of 2.5, a predefined percentage of the maximum SUV, or a gradient based method [44], although no consensus seems to emerge [45,46].

Moreover, magnetic resonance imaging (MRI) provides better soft tissue contrast and diffusion-weighted MRI has the advantage of additional functional information. For many tumour sites MRI has additional value for delineation and the role of MRI for oesophageal GTV delineation might be worthwhile exploring [47].

To our knowledge, this is the first multi-institutional, nationwide delineation study for oesophageal cancer patients to date rep-

resenting observer variation of the primary tumour GTV in daily clinical practice.

Solutions to reduce delineation variation might have great impact on daily clinical practice. We warrant international consensus guidelines for delineation of the primary tumour GTV and research to define the distal tumour extent more clearly is encouraged especially for lower thoracic/EGJ tumours.

5. Conclusion

Inter-observer variability in primary tumour delineation was in some cases substantial and mainly located at the cranial and caudal border. Fusion of CT with PET influenced the delineated volume in two-thirds of the patients, however its impact on observer variation was limited.

Declarations of interest

None.

Sources of support

None.

Collaborators

The following members of the Dutch National Platform for Radiotherapy of Gastrointestinal Tumours group were collaborators in the study:

Braam PM, Buijsen J, Ceha HM, Dewit L, Franssen JH, van Gestel K, Grootenboers DARH, Intven M, Jansen EPM, Kerkmeijer LGW, Mul VE, Muller K, Neelis KJ, Oppedijk V, Rozema T, Spruit PH.

References

[1] Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer 2013;49:1374–403.

[2] Shapiro J, van Lanschot JJB, Hulshof M, 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;16:1090–8.

[3] 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;12:681–92.

[4] 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;21:306–14.

[5] Bedenne L, Michel P, Bouche O, Milan C, Mariette C, Conroy T, et al. Chemoradiation followed by surgery compared with chemoradiation alone in squamous cancer of the esophagus: FFCD 9102. J Clin Oncol 2007;25:1160–8.

[6] Crosby TD, Brewster AE, Borley A, Perschky L, Kehagioglou P, Court J, et al. Definitive chemoradiation in patients with inoperable oesophageal carcinoma. Br J Cancer 2004;90:70–5.

[7] Ohtsu A, Boku N, Muro K, Chin K, Muto M, Yoshida S, et al. Definitive chemoradiotherapy for T4 and/or M1 lymph node squamous cell carcinoma of the esophagus. J Clin Oncol 1999;17:2915–21.

[8] Stahl M, Stuschke M, Lehmann N, Meyer HJ, Walz MK, Seeber S, et al. Chemoradiation with and without surgery in patients with locally advanced squamous cell carcinoma of the esophagus. J Clin Oncol 2005;23:2310–7.

[9] 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;118:2632–40.

[10] Welsh JW, Seyedin SN, Allen PK, Hofstetter WL, Ajani JA, Chang JY, et al. Local control and toxicity of a simultaneous integrated boost for dose escalation in locally advanced esophageal cancer: interim results from a prospective phase I/II Trial. J Thorac Oncol 2017;12:375–82.

[11] Nederlands Trial Register [Internet]. Amsterdam: Academic Medical Center (The Netherlands); 2004 Oct 26. Identifier NTR3532 Artodeidcfp.

[12] ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000 Feb 29. Identifier NCT01348217, Radiochemotherapy with and without dose escalation in patients presenting locally advanced or inoperable carcinoma of the oesophagus (CONCORDE); 2011 May 5; [about 5 screens]. Available from: https://clinicaltrials.gov/ct2/show/NCT01348217?term=dose+escalation%2C+concorde&cond=Oesophageal+Cancer&rank=1.

[13] Noordman BJ, Wijnhoven BPL, Lagarde SM, Biermann K, van der Gaast A, Spaander MCW, et al. Active surveillance in clinically complete responders after neoadjuvant chemoradiotherapy for esophageal or junctional cancer. Dis Esophagus 2017;30:1–8.

[14] Sillah K, Williams LR, Laasch HU, Saleem A, Watkins G, Pritchard SA, et al. Computed tomography overestimation of esophageal tumor length: implications for radiotherapy planning. World J Gastrointest Oncol 2010;2:197–204.

[15] 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.

[16] Thomas L, Lapa C, Bundschuh RA, Polat B, Sonke JJ, Guckenberger M. Tumour 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.

[17] Muijs CT, Beukema JC, Pruim J, Mul VE, Groen H, Plukker JT, et al. A systematic review on the role of FDG-PET/CT in tumour delineation and radiotherapy planning in patients with esophageal cancer. Radiother Oncol 2010;97:165–71.

[18] Westerterp M, Van Westreenen HL, Sloof GW, Plukker JT, Van Lanschot JJ. Role of positron emission tomography in the (re-)staging of oesophageal cancer. Scand J Gastroenterol Suppl 2006;116–22.

[19] Stiekema J, Vermeulen D, Vegt E, Voncken FE, Aleman BM, Sanders J, et al. Detecting interval metastases and response assessment using 18F-FDG PET/CT after neoadjuvant chemoradiotherapy for esophageal cancer. Clin Nucl Med 2014;39:862–7.

[20] Vesprini D, Ung Y, Dinniwell 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-ᴅ-glucose positron emission tomography/computed tomography. Clin Oncol (R Coll Radiol) 2008;20:631–8.

[21] 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: PET/CT improves esophageal target definition. Dis Esophagus 2010;23:493–501.

[22] Moureau-Zabotto L, Touboul E, Lerouge D, Deniaud-Alexandre E, Grahek D, Foulquier JN, et al. Impact of CT and 18F-deoxyglucose positron emission tomography image fusion for conformal radiotherapy in esophageal carcinoma. Int J Radiat Oncol Biol Phys 2005;63:340–5.

[23] Konski A, Doss M, Milestone B, Haluszka O, Hanlon A, Freedman G, 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;61:1123–8.

[24] Leong T, Everitt C, Yuen K, Condron S, Hui A, Ngan SY, et al. A prospective study to evaluate the impact of FDG-PET on CT-based radiotherapy treatment planning for oesophageal cancer. Radiother Oncol 2006;78:254–61.

[25] Gondi V, Bradley K, Mehta M, Howard A, Khuntia D, Ritter M, et al. Impact of hybrid fluorodeoxyglucose positron-emission tomography/computed tomography on radiotherapy planning in esophageal and non-small-cell lung cancer. Int J Radiat Oncol Biol Phys 2007;67:187–95.

[26] Hong TS, Killoran JH, Mamede M, Mamon HJ. Impact of manual and automated interpretation of fused PET/CT data on esophageal target definitions in radiation planning. Int J Radiat Oncol Biol Phys 2008;72:1612–8.

[27] Muijs CT, Schreurs LM, Busz DM, Beukema JC, van der Borden AJ, Pruim J, et al. Consequences of additional use of PET information for target volume delineation and radiotherapy dose distribution for esophageal cancer. Radiother Oncol 2009;93:447–53.

[28] Li Z, Rice TW. Diagnosis and staging of cancer of the esophagus and esophagogastric junction. Surg Clin North Am 2012;92:1105–26.

[29] Nestle U, Kremp S, Schaefer-Schuler A, Sebastian-Welsch C, Hellwig D, Rube C, et al. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. J Nucl Med 2005;46:1342–8.

[30] Kouwenhoven E, Giezen M, Struikmans H. Measuring the similarity of target volume delineations independent of the number of observers. Phys Med Biol 2009;54:2863–73.

[31] Nijkamp J, de Haas-Kock DF, Beukema JC, Neelis KJ, Woutersen D, Ceha H, et al. Target volume delineation variation in radiotherapy for early stage rectal cancer in the Netherlands. Radiother Oncol 2012;102:14–21.

[32] Deurloo KE, Steenbakkers RJ, Zijp LJ, de Bois JA, Nowak PJ, Rasch CR, et al. Quantification of shape variation of prostate and seminal vesicles during external beam radiotherapy. Int J Radiat Oncol Biol Phys 2005;61:228–38.

[33] Mamede M, El Fakhri G, Abreu-e-Lima P, Gandler W, Nose V, Gerbaudo VH. Pre-operative estimation of esophageal tumor metabolic length in FDG-PET images with surgical pathology confirmation. Ann Nucl Med 2007;21:553–62.

[34] Zhong X, Yu J, Zhang B, Mu D, Zhang W, Li D, et al. Using 18F-fluorodeoxyglucose positron emission tomography to estimate the length of gross tumor in patients with squamous cell carcinoma of the esophagus. Int J Radiat Oncol Biol Phys 2009;73:136–41.

[35] Han D, Yu J, Yu Y, Zhang G, Zhong X, Lu J, et al. Comparison of (18)F-fluorothymidine and (18)F-fluorodeoxyglucose PET/CT in delineating gross

tumor volume by optimal threshold in patients with squamous cell carcinoma of thoracic esophagus. Int J Radiat Oncol Biol Phys 2010;76:1235–41.

[36] 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;103:344–52.

[37] Wu AJ, Bosch WR, Chang DT, Hong TS, Jabbour SK, Kleinberg LR, et al. Expert consensus contouring guidelines for intensity modulated radiation therapy in esophageal and gastroesophageal junction cancer. Int J Radiat Oncol Biol Phys 2015;92:911–20.

[38] Tai P, Van Dyk J, Yu E, Battista J, Stitt L, Coad T. Variability of target volume delineation in cervical esophageal cancer. Int J Radiat Oncol Biol Phys 1998;42:277–88.

[39] Kienle P, Buhl K, Kuntz C, Dux M, Hartmann C, Axel B, et al. Prospective comparison of endoscopy, endosonography and computed tomography for staging of tumours of the oesophagus and gastric cardia. Digestion 2002;66:230–6.

[40] Caldwell CB, Mah K, Skinner M, Danjoux CE. Can PET provide the 3D extent of tumor motion for individualized internal target volumes? A phantom study of the limitations of CT and the promise of PET. Int J Radiat Oncol Biol Phys 2003;55:1381–93.

[41] Kruis MF, van de Kamer JB, Vogel WV, Belderbos JS, Sonke JJ, van Herk M. Clinical evaluation of respiration-induced attenuation uncertainties in pulmonary 3D PET/CT. EJNMMI Phys 2015;2:4.

[42] Patel AA, Wolfgang JA, Niemierko A, Hong TS, Yock T, Choi NC. Implications of respiratory motion as measured by four-dimensional computed tomography for radiation treatment planning of esophageal cancer. Int J Radiat Oncol Biol Phys 2009;74:290–6.

[43] Lever FM, Lips IM, Crijns SP, Reerink O, van Lier AL, Moerland MA, et al. Quantification of esophageal tumor motion on cine-magnetic resonance imaging. Int J Radiat Oncol Biol Phys 2014;88:419–24.

[44] Werner-Wasik M, Nelson AD, Choi W, Arai Y, Faulhaber PF, Kang P, et al. What is the best way to contour lung tumors on PET scans? Multiobserver validation of a gradient-based method using a NSCLC digital PET phantom. Int J Radiat Oncol Biol Phys 2012;82:1164–71.

[45] Zaidi H, El Naqa I. PET-guided delineation of radiation therapy treatment volumes: a survey of image segmentation techniques. Eur J Nucl Med Mol Imaging 2010;37:2165–87.

[46] Lee JA. Segmentation of positron emission tomography images: some recommendations for target delineation in radiation oncology. Radiother Oncol 2010;96:302–7.

[47] van Rossum PS, van Hillegersberg R, Lever FM, Lips IM, van Lier AL, Meijer GJ, et al. Imaging strategies in the management of oesophageal cancer: what's the role of MRI? Eur Radiol 2013;23:1753–65.