BJR

Received: 24 October 2016 | Revised: 16 March 2017 | Accepted: 22 March 2017

https://doi.org/10.1259/bjr.20160826

Cite this article as:
Onal C, Cengiz M, Guler OC, Dolek Y, Ozkok S. The role of delineation education programs for improving interobserver variability in target volume delineation in gastric cancer. *Br J Radiol* 2017; **90**: 20160826.

FULL PAPER

The role of delineation education programs for improving interobserver variability in target volume delineation in gastric cancer

[1]CEM ONAL, MD, [2]MUSTAFA CENGIZ, MD, [1]OZAN C GULER, MD, [1]YEMLIHA DOLEK, MSci and [3]SERDAR OZKOK, MD

[1]Department of Radiation Oncology, Faculty of Medicine, Baskent University, Adana, Turkey
[2]Department of Radiation Oncology, Faculty of Medicine, Hacettepe University, Ankara, Turkey
[3]Department of Radiation Oncology, Faculty of Medicine, Ege University, Izmir, Turkey

Address correspondence to: Prof Cem Onal
E-mail: *hcemonal@hotmail.com*

Objective: To assess whether delineation courses for radiation oncologists improve interobserver variability in target volume delineation for post-operative gastric cancer radiotherapy planning.

Methods: 29 radiation oncologists delineated target volumes in a gastric cancer patient. An experienced radiation oncologist lectured about delineation based on contouring atlas and delineation recommendations. After the course, the radiation oncologists, blinded to the previous delineation, provided delineation for the same patient.

Results: The difference between delineated volumes and reference volumes for pre- and post-course clinical target volume (CTV) were 19.8% (−42.4 to 70.6%) and 12.3% (−12.0 to 27.3%) ($p = 0.26$), respectively. The planning target volume (PTV) differences pre- and post-course according to the reference volume were 20.5% (−40.7 to 93.7%) and 13.1% (−10.6 to 29.5%) ($p = 0.30$), respectively. The concordance volumes between the pre- and post-course CTVs and PTVs were 467.1 ± 89.2 *vs* 597.7 ± 54.6 cm^3 ($p < 0.001$) and 738.6 ± 135.1 *vs* 893.2 ± 144.6 cm^3 ($p < 0.001$), respectively. Minimum and maximum observer variations were seen at the cranial part and splenic hilus and at the caudal part of the CTV. The kappa indices compared with the reference contouring at pre- and post-course delineations were 0.68 and 0.82, respectively.

Conclusion: The delineation course improved interobserver variability for gastric cancer. However, impact of target volume changes on toxicity and local control should be evaluated for further studies.

Advances in knowledge: This study demonstrated that a delineation course based on current recommendations helped physicians delineate smaller and more homogeneous target volumes. Better target volume delineation allows proper target volume irradiation and preventing unnecessary normal tissue irradiation.

INTRODUCTION

For patients with resectable gastric adenocarcinoma, locoregional recurrence is significant and occurs in as many as 80–85% of failures after surgery alone.[1] In 2001, the Gastric Surgical Adjuvant Trial Intergroup 0116 (INT0116) established the benefit of adjuvant chemoradiotherapy (ChRT) in the treatment of high risk, completely resected adenocarcinoma of the stomach and gastro-oesophageal junction.[2] In a meta-analysis by Valentini et al,[3] the authors reported that adjuvant radiotherapy (RT) has a significant impact on survival in resectable gastric cancer patients [HR: 1.31 (95% CI: 1.04–1.66; $p = 0.02$)]. In an evaluation of the role of post-operative ChRT after D2 dissection, patients with pathological lymph node metastasis had superior disease-free survival when they were treated with chemotherapy and RT compared with treatment with chemotherapy alone.[4] These studies strongly support the integration of post-operative ChRT for locally advanced gastric cancer.

A previous INT0116 study conducted to determine the treatment fields in the era of two-dimensional (2D) planning reported higher rates of acute grade 3 (41%) and grade 4 (32%) toxicities, resulting in incomplete treatment in 17% of cases. Recent studies heralded three-dimensional conformal RT (3DCRT) and intensity-modulated RT (IMRT) as potential methods to decrease the observed toxicities of conventional RT.[5,6] Thus, delineation of precise target volumes is essential to utilize 3DCRT or IMRT properly.

In 2002, Smalley et al[7] published a guideline for better defining the essentials of RT application in post-operative gastric cancer cases. Based on the findings of a study that

examined the patterns of recurrence for gastric adenocarcinoma after potentially curative resection, recurrences have been commonly seen at the tumour bed, anastomosis and regional lymphatics.[1,8] For this reason, in order to define target volumes and deliver radiation appropriately to high-risk regions, it is essential to know the location of regional lymphatics and vascular structures. As RT fields become increasingly conformal in an attempt to limit the dose to normal critical structures, accurate identification of treatment volumes on CT-based planning images, including the regional gastric lymph node stations, becomes increasingly important; however, accurate identification of regional gastric lymph node stations is difficult, particularly because post-operative gastric anatomy varies substantially based on the type of surgical resection performed. To reduce contouring variations, strict guidelines coupled with education programs are required. Numerous studies have demonstrated substantial interobserver variability in contouring among radiation oncologists, which can be reduced when providers access contouring reference aids.[9–12]

Thus, the Turkish Society for Radiation Oncology (TSRO) conducted delineation courses for different tumour sites based on the current guidelines and recommendations in order to establish better contouring of the target volume. These courses were held twice per year and targeted most of the radiation oncology residents and specialists nationwide. The aim of this study was to assess whether such delineation courses for radiation oncologists improve interobserver variability in target volume delineation for post-operative gastric cancer patients.

METHODS AND MATERIALS

In November 2014, 2 contouring courses were conducted in Ankara and Istanbul, Turkey, and were attended by 40 radiation oncologists. Of these, 29 radiation oncologists 18 female; 62% and 11 male; 38% accepted the invitation to participate in this delineation study. One patient with gastric corpus tumour, who had undergone planning CT and had complete pathological findings, surgical reports and pre-operative radiological images, was randomly selected for this study.

The courses were held with the same lecturer with the same clinical case, and all participants attended the course only once. All participants had at least 5 years of experience in radiation oncology (average experience 6.8 years), and all had the opportunity to delineate at least 20–25 gastric cancer patients per year and to perform 3DCRT or IMRT during their routine practice. All participants were in the same classroom, and they were asked to delineate the volumes before and after the course within 1 h of the period. Pre- and post-course delineations had been performed under exactly the same conditions using the same tools and the same clinical information for each course.

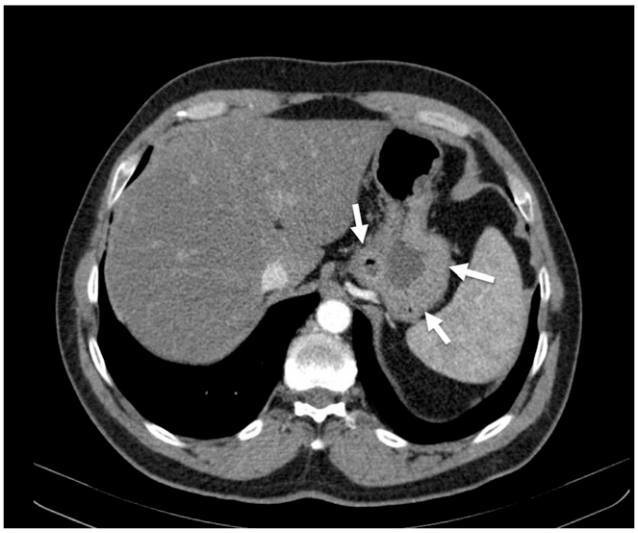
Patient characteristics

The patient was a 48-year-old otherwise healthy male who presented with dyspeptic symptoms. An ulcerovegetan lesion extending from the cardia to incisura angularis was seen during upper endoscopy, and the histopathological finding revealed gastric adenocarcinoma. The pre-operative CT scan demonstrated a thickening of the wall extending from the oesophagogastric junction to the corpus with no lymphadenopathy or distant metastasis (Figure 1). The patient underwent D2 total gastrectomy, and the spleen was preserved. The pathological specimen revealed a stage pT3N1 tumour with metastasis to 2 of the 22 lymph nodes. The metastatic lymph nodes are suprapyloric lymph nodes located at the lesser curvature. The surgical margins were negative, but the distance to the oesophageal surgical margin was 1.5 cm.

Target volume delineation

All participants had a brief education for the MIM Maestro® contouring program because none of them used this contouring program during their routine practice. The planning CT scan was retrieved by all physicians for contouring. In order to better demonstrate vasculature and anastomosis, intravenous and oral contrast was used during planning CT. A brief description about the clinical, radiological and pathological findings of the patient was provided before contouring. Each physician contoured the lymph nodes and tumour bed separately on their computers before the course using the MIM Maestro contouring program. First, the perioesophageal lymphatics, splenic artery and splenic hilus, celiac artery, superior mesenteric artery and portal hilus were delineated. The tumour bed was created by delineating the pancreas corpus and tail and the medial half of the diaphragm. For para-aortic lymphatics, the aorta was contoured from the first lymphatic station below the L3 vertebra, a 2.5–3 cm expansion at the right side and 1.5–2 cm at the left side, 1.5 cm anteriorly and 0.5 cm posteriorly was given. The liver, both the kidneys, spinal cord and intestines were also delineated on the reference CT images for defining organs at risk. The clinical target volumes (CTVs) were created by adding 1 cm to the gastric bed and regional lymphatics, and the planning target volume (PTV) was created by adding 0.5 cm to the CTV. The CTV and PTV were defined by the physicians on the basis of their departmental protocol. None of the observers had knowledge of the volumes outlined by the others.

Figure 1. Pre-operative CT scan demonstrated a thickening of the wall extending from the oesophagogastric junction to the corpus (arrows) with no lymphadenopathy or distant metastasis.

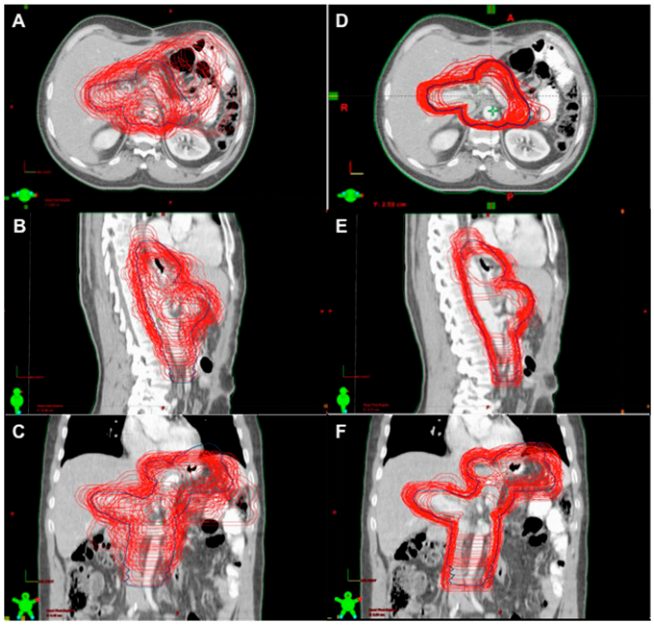


An experienced physician chosen by the TSRO lectured about the anatomy of the stomach and lymphatic drainage and defined some pitfalls about contouring target volumes in post-operative gastric cancer patients, on the basis of recommendations and the delineation atlas.[7,13–16] The CTV encompasses the gastric bed, oesophagojejunal anastomosis and regional lymphatics according to the tumour location. The lecturer was asked to delineate the target volume simultaneously during the course, and the same contours were used for all contouring programs. The contours delineated during the lecture were compared with the reference image. Surgical and pathological reports as well as pre-operative and post-operative diagnostic images were used for delineation. At the end of the delineation course, all observers were asked to delineate the volumes again without any knowledge of pre-course contours or contours delineated by the lecturer.

Comparison of contours

All contours were collected separately by the physician and by contouring time: pre-course *vs* post-course. The pre-course (CTV_{pre}) and post-course (CTV_{post}) CTVs and PTVs were redefined using the commercial software ARTiView™ v. 3.4.1 (AQUILAB, Loos les Lille, France) (Figure 2). For each physician, the CTVs and PTVs were calculated in cubic centimetres. The major variant was spatial volume discrepancy between the reference CTV as defined by the lecturer (CTV_{ref}) and CTV_{pre} and CTV_{post} values defined by the other observers for evaluation of the position and shape. For the first step, each target volume was compared with each other and with the reference volume. In second step, Jaccard index, which was calculated using the mathematical formula $(A \cup B)/(A \cap B)$, was used for evaluating the overlap area between the contoured volumes and reference contours. The concordance of spatial volumes was calculated using the intersection volume formula as $(A \cap B)$, which represents the congruity between volumes delineated by physicians and the reference image. Measures were expressed in terms of absolute difference between the volumes for both the CTV and PTV. Then, the kappa index was calculated for pre- and post-course volumes for analyzing the variance between pre- and post-course delineated volumes. The strength of kappa agreement was defined as: <0 is poor; 0–0.20 is slight; 0.21–0.40 is fair; 0.41–0.60 is moderate; 0.61–0.80 is substantial; and 0.81–1.00 is almost perfect.[17] Additionally, the individual observer variation was determined by measuring the distance from the reference PTV to all individual delineations at four different sites: cranial (paraoesophageal and hemidiaphragm), portal hilus, splenic hilus and caudally lower border of the PTV (para-aortic lymphatics). The comparison of distance variations was performed between the pre- and post-course volumes.

Figure 2. A representative image demonstrating planning target volumes (PTVs) delineated by course attendees and the reference (bold line) at (a) transverse, (b) sagittal and (c) coronal sections before the delineation course. (d–f) The PTVs delineated by attendees after the delineation course.



Statistical analysis

All statistical analyses were performed using SPSS® v. 20.0 (IBM Corp., New York, NY; formerly SPSS Inc., Chicago, IL). Student's *t*-tests were used to compare pre- and post-course volumes. The contour evaluation module in the commercial software ARTiView was used to evaluate the pre- and post-course delineated CTVs and PTVs. The interobserver and intraobserver variability and standard deviations were calculated using the output generated by one-way analysis of variance. Moreover, intraobserver variability for the lecturer was also assessed by a blind repetition of the target volume contouring during the two courses and one contouring before the courses. The concordance and discordance between the pre- and post-course volumes were compared. Another comparison was performed between the pre- and post-course volumes and reference volumes separately. All reported *p*-values were two-sided, and a $p < 0.05$ was considered statistically significant.

RESULTS

Volume comparsions

The reference volumes that were delineated by the lecturer were 724.6 cm^3 for the CTV and 1052.1 cm^3 for the PTV. There were no significant differences in the CTV and PTV between the pre- and post-course volumes (Table 1). The ranges in volumes were larger in pre-course volumes than in post-course volumes for individual observers (Figure 3a,b).

The CTV and PTV delineated before and after courses were significantly larger than the reference volumes ($p < 0.05$). In addition, 8 of 29 (28%) observers delineated CTV values and 7 of 29 (24%) observers delineated PTV values smaller than the reference CTV and PTV, respectively, before the course. Only 3 of 29 (10%) observers delineated CTV and PTV values smaller than the corresponding reference CTV and PTV after the contouring course. The difference between the delineated volumes and reference volume for pre- and post-course CTV were 19.8% (−42.4 to 70.6%) and 12.3% (−12.0 to 27.3%), respectively (Figure 4a). The PTV differences for pre- and post-course delineation were 20.5% (−40.7 to 93.7%) and 13.1% (−10.6 to 29.5%), respectively (Figure 4b).

Overlap measures and statistical measures

The discrepancy in CTVs between the observers' delineated volumes and the reference volume was significantly higher before the contouring course than after the contouring course (665.0 ± 146.5 *vs* $358.1 \pm 117.2\,cm^3$; $p < 0.001$). Similarly, the PTV_{pre} discrepancy volume was significantly higher than the PTV_{post} discrepancy volume (797.5 ± 181.6 *vs* $432.1 \pm 104.9\,cm^3$; $p < 0.001$); however, the CTV_{pre} concordance volume was significantly lower than the CTV_{post} concordance volume (467.1 ± 89.2 *vs* $597.7 \pm 54.6\,cm^3$; $p < 0.001$). The PTV_{pre} concordance volume with the reference PTV was significantly lower than the PTV_{post} concordance volume with reference PTV (738.6 ± 135.1 *vs* $893.2 \pm 144.6\,cm^3$; $p < 0.001$).

The kappa indices compared with the reference delineation were substantial [0.68 (range 0.40–0.84)] at pre-course contouring and almost perfect [0.82 (0.63–0.93)] at post-course delineation, and the difference between pre- and post-course kappa indices was statistically significant ($p < 0.001$). The observer variations at the cranial part, portal hilus, splenic hilus and caudal part were significantly less at post-course delineation than at pre-course delineation (Table 2). The minimum observer variation was found at the cranial part and splenic hilus (Figure 5). The maximum observer variation was found at the caudal part of the target volume.

DISCUSSION

In this study, we found that the target volumes were larger with very high discrepancies in volumes between observers before course than after course. After attending a contouring course that was mainly based on contouring atlases and guidelines, the target volumes became smaller with less associated discrepancy in target volumes between the observers. Additionally, a brief improvement according to the reference contour was observed at post-course volumes compared with pre-course volumes, especially at the caudal part of the target volume. To our knowledge, this is the first study to date to attempt to establish the importance of education for identification of target volumes based on gastric lymph node stations, particularly in the post-operative setting.

The multimodality treatment strategy has emerged as a viable option for the treatment of localized, resectable gastric cancer.[2] The INT0116 study established post-operative ChRT as an effective adjuvant therapy approach. With a median follow-up of 5 years, adjuvant ChRT improved the 3-year overall survival rates compared with surgery alone (41% *vs* 50%, $p < 0.001$);[2] however, despite these promising results, high rates of acute toxicity (grade 3+, 41%; grade 4+, 32%) have been observed, necessitating early treatment termination in 17% of patients, and such toxicity was the main drawback of the study. After plan revisions in the INT0116 study, the reported plan error was 35%, and 6.5% of the treatment plans required major revisions.[2] The published consensus guidelines for defining microscopic spreads are mainly based on conventional techniques using bony landmarks. Moreover, in the INT0116 study, patients were mostly treated with 2D planning techniques.[7] With implementation of conformal RT techniques for treating gastric cancer, target volume delineation becomes critical. With improved planning techniques using CT-based planning to spare normal tissue, fewer patients should require prolonged treatment breaks or discontinuation of RT.[18] Goodman et al[19] reported that of patients who received 3DCRT, >70% were treated in a timely fashion and only 9% did not complete therapy, an indicator that the quality of care is improving with the incorporation of modern treatment techniques.

The delineation of target volumes is cumbersome in post-operative gastric cancer patients, due, in part, to the complete distortion of anatomy. Other challenges to defining the target volumes in gastric cancer patients include lymphatic drainage variability based on tumour location, bowel movements and the guidelines for post-operative gastric cancer target volume delineation that were published before 3DCRT or IMRT.[7,16] Although the lymph node contouring atlas and guidelines help to define the target volumes for post-operative gastric cancer, these factors remain complex for implementation into routine 3DCRT or IMRT practice.[13,14] Yoon et al[14] analyzed regional lymphatic recurrences in 91 gastric cancer patients treated with D2 dissection. The authors suggested vessel-based contouring to potentially minimize interobserver variability in CTV delineation and decrease geographic misses. Wo et al[13] published a gastric lymph node atlas for three gastric cancer patients treated using different surgical procedures and one patient with intact gastric. The authors planned to define the location of lymphatics for aiding the determination of lymphatic CTV during gastric irradiation. During the delineation courses in our study, we defined the target volumes with vessel-based contouring. Also, we focused on the general guidelines that have been proposed to aid in definition of the CTV for adjuvant radiation treatment fields based on the location, T stage of the primary tumour and N stage.[16] Our CTV definition encompassed the tumour bed, anastomosis and nodal drainage regions.

Interobserver variation during target volume delineation may potentially increase tumour recurrence, which may be due to incomplete target coverage. Additionally, unnecessarily large target volume delineation may increase the surrounding organ dose, which, in turn, induces toxicity. Previously, variations

Table 1. The mean pre- and post-course volumes delineated by the observers

Volume	Pre-course (cm^3)		Post-course (cm^3)		*p*-value
	Mean ± SD	Range	Mean ± SD	Range	
CTV	880.3 ± 254.7	427.6–1491.6	833.5 ± 101.6	576.4–1049.8	0.38
PTV	1267.7 ± 305.3	623.8–2038.3	1216 ± 140.2	940.2–1523.8	0.42

CTV, clinical target volume; PTV, planning target volume; SD, standard deviation.

Figure 3. The (a) clinical target volumes and (b) planning target volumes delineated before and after the delineation course. The dotted black line indicates the reference volume.

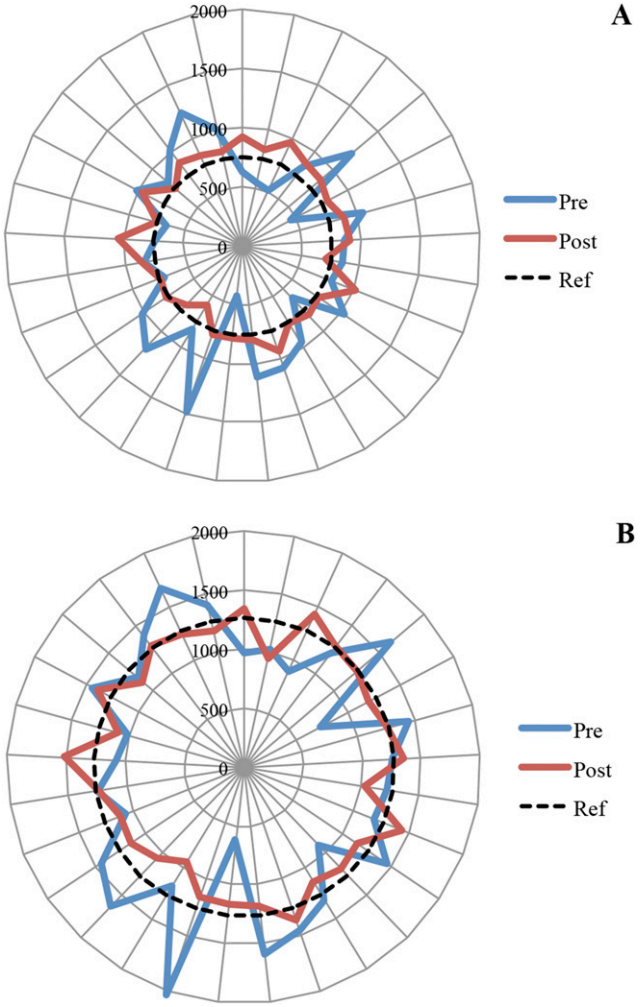
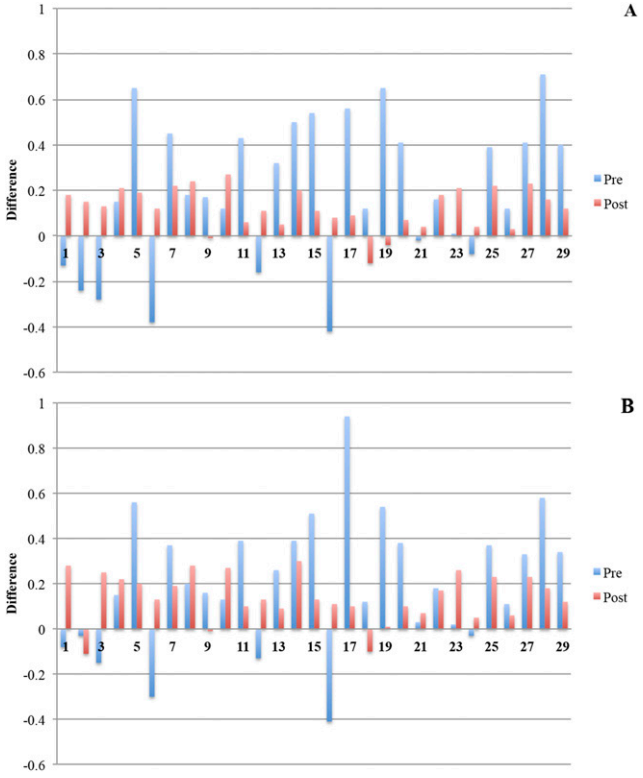


Figure 4. The (a) clinical target volume and (b) planning target volume changes before and after the delineation course according to the reference contouring.



between clinicians in target volume description have been attributed to intraphysician variability due to the incapability to reconstruct the same target volumes on the representative scans, problems in defining the gastric bed and differences in treatment philosophy.[20] Some earlier studies assessed the interobserver variability in delineating gastric cancer volumes.[21–23] Chung et al[21] found significant variations in RT field areas in 2D planning, although no significant change was observed in CTV volumes. Jansen et al[22] assessed interobserver variability between six physicians who delineated the target volumes of a patient with distal gastric cancer with the help of a delineation guide. These investigators found large observer variations in CTV (240–821 cm^3) and PTV (634 and 1677 cm^3), and the CTV and PTV overlaps between observers were 72% and 78%, respectively. In another study, Moretones et al[23] analyzed the interobserver variability by four physicians who delineated nine gastric cancer patients. These authors demonstrated broader average differences between physicians with discrepancies that ranged from 146.90 to 551.80 cm^3. Furthermore, the localization of higher variability observed during contouring is also an important problem to be solved. Jansen et al[22] found that a large variation during contouring was observed at the cranial part of the CTV, mainly during delineation of part of the diaphragm and perioesophageal nodes. In another study, the dome of the diaphragm, anterior abdominal wall, duodenal stump and porta hepatis were delineated by 20 radiation oncology residents before and after training courses.[24] The greater delineation variations were observed at the dome of the diaphragm and duodenal stump. Weiss et al[25] pointed out the importance of education programs, image optimization and closer partnership with radiologists and surgeons for selected cases, in order to minimize the delineation variabilities. In this current study, we observed larger interobserver variation at the caudal part of the CTV before course, which was particularly improved after course. This large variation could be explained as a result of different interpretation of some guidelines or with unfamiliarity of anatomical locations of the lymph nodes.

Our study was designed to evaluate the importance of education, on the basis of recommendations and guidelines, on interobserver variations in target volume delineation. The initial comparison of the delineated target volumes revealed that the primary tumour volume was defined as significantly larger before the education course than after the course. The large variations in target volumes suggest that our observers may not have ability to exactly delineate CTV on axial CT images. Instead, bony landmarks have been used to outline CTV during 2D planning. Importantly, the volume ranges decreased following the educational course, and a significant increase in concordance volume and a significant decrease in discordance volume were observed, indicating

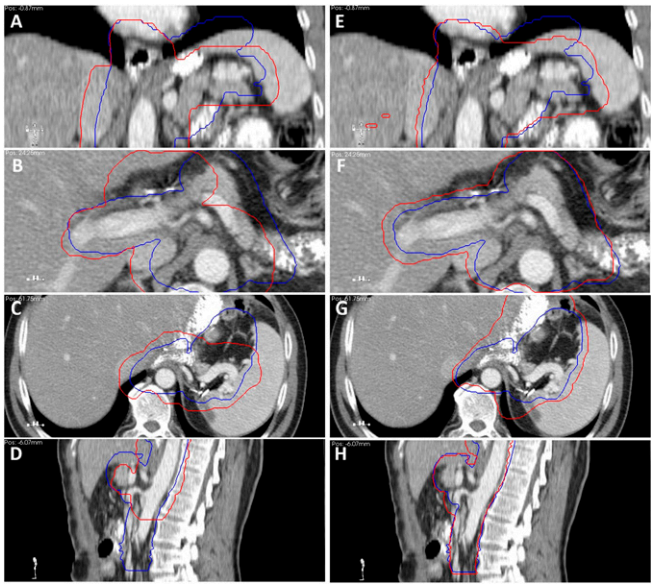
Table 2. Mean pre- and post-course variations calculated at the cranial part, portal hilus, splenic hilus and caudal part of the target volume delineated by the observers according to the reference image

Distance	Pre-course (mm)	Post-course (cm^3)	*p*-value
	Mean (range)	Mean (range)	
Cranial	−13.1 (−30.5 to 16.2)	−7.7 (−20.1 to 4.7)	0.001
Portal	−12.8 (−63.4 to 17.2)	4.7 (−10.7 to 12.1)	<0.001
Splenic	1.0 (−35.9 to 13.0)	4.9 (−9.0 to 14.2)	0.03
Caudal	−56.7 (−30.5 to 16.2)	−12.0 (−39.0 to 34.1)	<0.001

that the delineated target volumes became more standardized following the educational contouring courses. Vinod et al[26] reported that interobserver variability in volume delineation could be reduced with the use of guidelines, provision of autocontours and teaching courses. The authors reported that the guidelines significantly reduced interobserver variability in 7/9 studies, teaching interventions reduced interobserver variability in 8/9 studies and autocontour improved consistency of contouring in 6/7 studies. However, in the case of absence of these education programs, contouring atlases and interactive web courses may help to define target volumes better and diminish interobserver variability.[27,28]

This study has some limitations. First, the study cohort is quite heterogeneous as the physicians work at different clinics with different treatment policies, and these differences may have contributed to the large variations in the initial target volumes. Second, the study was based on only one patient, with a corpus tumour with total gastrectomy. Although another course program employing another patient with a different surgical procedure and a different tumour location may be more feasible, such an inclusion is technically difficult and time intensive. The course conducted by the TSRO was programmed to encompass all tumour sites, although the educational course designed for only one site took approximately 3–4 h. Thus, rather than include different tumour sites, we preferred to include the most difficult case that required larger volumes. Lastly, what remains largely unknown is the long-term effect of the different educational methods. Although we only focused on the importance of such educational programs on interobserver variability for very difficult disease sites, such evaluation with further delineation 5–6 months after the course may be the subject of another study.

Figure 5. The target volume delineated by a representative observer (light line) compared with the reference contour (dark line). Moderate variation at the cranial part, portal hilus and splenic hilus (a–c) and maximum variation at the caudal part (d) were observed before the course. An evident improvement was seen at post-course contouring at every borders of target volume (e–h).



However, this study is important in several ways. Firstly, this study demonstrated the importance of educational programs on target volume delineation, especially in one of the most difficult tumour groups. With this contouring course, the observers were able to learn the pitfalls of contouring of post-operative gastric cancer patients and to evaluate their knowledge by comparing the target volumes that they delineated before and after the contouring course. Second, we analyzed interobserver variability with descriptive statistics, Jaggard index and kappa statistics for making the analysis properly. Additionally, a detailed observation of contour variations was performed at four different parts of the CTV for both pre- and post-course contours. Although there is still no clear guideline for defining interobserver variability in target volume delineation in gastric cancer patients, the combination of descriptive statistics, overlap measure and statistical measure of agreement are used to assess the delineation variability.[29]

CONCLUSION

To our knowledge, our study is the first to demonstrate the importance of educational contouring courses based on guidelines and a contouring atlas in post-operative gastric cancer patients. With the aid of educational programs, the target volumes become smaller, and interobserver variability is minimized for gastric cancer, which may potentially increase local control with less toxicity.

ACKNOWLEDGMENTS

This study was conducted on behalf of the Turkish Society for Radiation Oncology. This work was accepted as a poster presentation at the 57th Annual Meeting of the American Society for Radiation Oncology; 14–17 September 2015; San Antonio, TX.

REFERENCES

1. Gunderson LL, Sosin H. Adenocarcinoma of the stomach: areas of failure in a re-operation series (second or symptomatic look) clinicopathologic correlation and implications for adjuvant therapy. *Int J Radiat Oncol Biol Phys* 1982; **8**: 1–11.
2. Macdonald JS, Smalley SR, Benedetti J, Hundahl SA, Estes NC, Stemmermann GN, et al. Chemoradiotherapy after surgery compared with surgery alone for adenocarcinoma of the stomach or gastroesophageal junction. *N Engl J Med* 2001; **345**: 725–30. doi: https://doi.org/10.1056/nejmoa010187
3. Valentini V, Cellini F, Minsky BD, Mattiucci GC, Balducci M, D'Agostino G, et al. Survival after radiotherapy in gastric cancer: systematic review and meta-analysis. *Radiother Oncol* 2009; **92**: 176–83.
4. Lee J, Lim DH, Kim S, Park SH, Park JO, Park YS, et al. Phase III trial comparing capecitabine plus cisplatin *versus* capecitabine plus cisplatin with concurrent capecitabine radiotherapy in completely resected gastric cancer with D2 lymph node dissection: the ARTIST trial. *J Clin Oncol* 2012; **30**: 268–73.
5. Wieland P, Dobler B, Mai S, Hermann B, Tiefenbacher U, Steil V, et al. IMRT for postoperative treatment of gastric cancer: covering large target volumes in the upper abdomen: a comparison of a step-and-shoot and an arc therapy approach. *Int J Radiat Oncol Biol Phys* 2004; **59**: 1236–44.
6. Ringash J, Perkins G, Brierley J, Lockwood G, Islam M, Catton P, et al. IMRT for adjuvant radiation in gastric cancer: a preferred plan? *Int J Radiat Oncol Biol Phys* 2005; **63**: 732–8. doi: https://doi.org/10.1016/j.ijrobp.2005.03.013
7. Smalley SR, Gunderson L, Tepper J, Martenson JA Jr, Minsky B, Willett C, et al. Gastric surgical adjuvant radiotherapy consensus report: rationale and treatment implementation. *Int J Radiat Oncol Biol Phys* 2002; **52**: 283–93.
8. Zhu WG, Xua DF, Pu J, Zong CD, Li T, Tao GZ, et al. A randomized, controlled, multicenter study comparing intensity-modulated radiotherapy plus concurrent chemotherapy with chemotherapy alone in gastric cancer patients with D2 resection. *Radiother Oncol* 2012; **104**: 361–6.
9. Fiorino C, Reni M, Bolognesi A, Cattaneo GM, Calandrino R. Intra- and inter-observer variability in contouring prostate and seminal vesicles: implications for conformal treatment planning. *Radiother Oncol* 1998; **47**: 285–92.
10. Yamazaki H, Shiomi H, Tsubokura T, Kodani N, Nishimura T, Aibe N, et al. Quantitative assessment of inter-observer variability in target volume delineation on stereotactic radiotherapy treatment for pituitary adenoma and meningioma near optic tract. *Radiat Oncol* 2011; **6**: 10.
11. Hong TS, Bosch WR, Krishnan S, Kim TK, Mamon HJ, Shyn P, et al. Interobserver variability in target definition for hepatocellular carcinoma with and without portal vein thrombus: radiation therapy oncology group consensus guidelines. *Int J Radiat Oncol Biol Phys* 2014; **89**: 804–13. doi: https://doi.org/10.1016/j.ijrobp.2014.03.041
12. Cui Y, Chen W, Kong FM, Olsen LA, Beatty RE, Maxim PG, et al. Contouring variations and the role of atlas in non-small cell lung cancer radiation therapy: analysis of a multi-institutional preclinical trial planning study. *Pract Radiat Oncol* 2015; **5**: e67–75.
13. Wo JY, Yoon SS, Guimaraes AR, Wolfgang J, Mamon HJ, Hong TS. Gastric lymph node contouring atlas: a tool to aid in clinical target volume definition in 3-dimensional treatment planning for gastric cancer. *Pract Radiat Oncol* 2013; **3**: e11–19.
14. Yoon HI, Chang JS, Lim JS, Noh SH, Hyung WJ, An JY, et al. Defining the target volume for post-operative radiotherapy after D2 dissection in gastric cancer by CT-based vessel-guided delineation. *Radiother Oncol* 2013; **108**: 72–7.
15. Matzinger O, Gerber E, Bernstein Z, Maingon P, Haustermans K, Bosset JF, et al. EORTC-ROG expert opinion: radiotherapy volume and treatment guidelines for neoadjuvant radiation of adenocarcinomas of the gastro-esophageal junction and the stomach. *Radiother Oncol* 2009; **92**: 164–75.
16. Tepper JE, Gunderson LL. Radiation treatment parameters in the adjuvant postoperative therapy of gastric cancer. *Semin Radiat Oncol* 2002; **12**: 187–95.
17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–74.
18. Milano MT, Garofalo MC, Chmura SJ, Farrey K, Rash C, Heimann R, et al. Intensity-modulated radiation therapy in the treatment of gastric cancer: early clinical outcome and dosimetric comparison with conventional techniques. *Br J Radiol* 2006; **79**: 497–503.
19. Goodman KA, Khalid N, Kachnic LA, Minsky BD, Crozier C, Owen JB, et al. Quality research in radiation oncology analysis of clinical performance measures in the management of gastric cancer. *Int J Radiat Oncol Biol Phys* 2013; **85**: 355–62.
20. Cattaneo GM, Reni M, Rizzo G, Castellone P, Ceresoli GL, Cozzarini C, et al. Target delineation in post-operative radiotherapy of brain gliomas: interobserver variability and impact of image registration of MR (pre-operative) images on treatment planning CT scans. *Radiother Oncol* 2005; **75**: 217–23.
21. Chung HT, Shakespeare TP, Wynne CJ, Lu JJ, Mukherjee RK, Back MF. Evaluation of a radiotherapy protocol based on INT0116 for completely resected gastric adenocarcinoma. *Int J Radiat Oncol Biol Phys* 2004; **59**: 1446–53.
22. Jansen EP, Nijkamp J, Gubanski M, Lind PA, Verheij M. Interobserver variation of clinical target volume delineation in gastric cancer. *Int J Radiat Oncol Biol Phys* 2010; **77**: 1166–70. doi: https://doi.org/10.1016/j.ijrobp.2009.06.023
23. Moretones C, Leon D, Navarro A, Santacruz O, Boladeras AM, Macia M, et al. Interobserver variability in target volume delineation in postoperative radiochemotherapy for gastric cancer. A pilot prospective study. *Clin Transl Oncol* 2012; **14**: 132–7.
24. Socha J, Wolakiewicz G, Wasilewska-Tesluk E, Janiga P, Kondraciuk T, Majewska A, et al. Clinical target volume in postoperative radiotherapy for gastric cancer: identification of major difficulties and controversies. *Clin Transl Oncol* 2016; **18**: 480–8. doi: https://doi.org/10.1007/s12094-015-1396-6
25. 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**: 21–30. doi: https://doi.org/10.1007/s00066-003-0976-5
26. Vinod SK, Min M, Jameson MG, Holloway LC. A review of interventions to reduce inter-observer variability in volume delineation in radiation oncology. *J Med Imaging Radiat Oncol* 2016; **60**: 393–406. doi: https://doi.org/10.1111/1754-9485.12462
27. Fuller CD, Nijkamp J, Duppen JC, Rasch CR, Thomas CR Jr, Wang SJ, et al. Prospective randomized double-blind pilot study of site-specific consensus atlas implementation for rectal cancer target volume delineation in the cooperative group setting. *Int J Radiat Oncol Biol Phys* 2011; **79**: 481–9.
28. Allozi R, Li XA, White J, Apte A, Tai A, Michalski JM, et al. Tools for consensus analysis of experts' contours for radiotherapy structure definitions. *Radiother Oncol* 2010; **97**: 572–8.
29. Fotina I, Lutgendorf-Caucig C, Stock M, Potter R, Georg D. Critical discussion of evaluation parameters for inter-observer variability in target definition for radiation therapy. *Strahlenther Onkol* 2012; **188**: 160–7. doi: https://doi.org/10.1007/s00066-011-0027-6