CANCER DIAGNOSIS & PROGNOSIS
3: 433-438 (2023) doi: 10.21873/cdp.10236

Interobserver Variability in Contouring Hepatocellular Carcinoma at a Tertiary Center

CAROLINA DE LA PINTA[1], JUAN DAVID GARCÍA[2], DAVID SEVILLANO[2], RAFAEL COLMENARES[2], RAQUEL GARCÍA LATORRE[3], MANUEL GARVÍ[4], VANESA PINO[4], ALFONSO MURIEL[5], MERCEDES MARTÍN[1], EVA FERNÁNDEZ[1], RAUL HERNANZ[1], MARGARITA MARTÍN[1], JOSE ANTONIO DOMÍNGUEZ[1], TERESA MUÑÓZ[1], LUIS CRISTIAN PERNA[6], AGUSTÍN ALBILLOS[7] and SONSOLES SANCHO[1]

[1]Radiation Oncology Department, Ramón y Cajal University Hospital, IRYCIS, Alcalá University, Madrid, Spain;
[2]Medical Physics Department, Ramón y Cajal University Hospital, IRYCIS, Madrid, Spain;
[3]Radiology Department, Ramón y Cajal University Hospital, IRYCIS, Madrid, Spain;
[4]Radiation therapist, Ramón y Cajal University Hospital, IRYCIS, Madrid, Spain;
[5]Clinical Biostatistics Unit, Ramón y Cajal University Hospital, IRYCIS, CIBERESP, Universidad de Alcalá, Madrid, Spain;
[6]Pathology Department, Ramón y Cajal University Hospital, IRYCIS, Madrid, Spain;
[7]Department of Gastroenterology, Ramón y Cajal University Hospital, Universidad de Alcalá, IRYCIS, CIBEREHD, Madrid, Spain

Abstract. *Background/Aim: The optimal imaging test for delineation of the gross tumor volume (GTV) in hepatocellular carcinoma has not been defined. The hypothesis is that magnetic resonance imaging (MRI) allows for better visualization of the extent of tumor and will optimize the accuracy of tumor delineation for liver stereotactic radiotherapy compared with computed tomography (CT) only. We evaluated the interobserver agreement in GTV of hepatocellular carcinoma in a multicenter panel and compared MRI and CT in GTV delineation. Materials and Methods: After the institutional review boards approved the study, we analyzed anonymous CT and MRI obtained from five patients with hepatocellular carcinoma. Eight radiation oncologists at our center used CT and MRI to delineate five GTVs of liver tumors. In both CT and MRI, the GTV volumes were compared. Results: The median GTV volume on MRI was 2.4 cm^3 (range=0.59-15.6 cm^3) compared to 3.5 cm^3 (range=0.52-24.9 cm^3) on CT (p=0.36). The GTV volume as defined on MRI was larger or at least as large as the GTV volume on CT in two cases. Variance and standard deviation between observers in CT and MRI were minor (6 vs. 7.87 cm^3, and 2.5 vs. 2.8 cm^3 respectively). Conclusion: In cases with well-defined tumors, CT is easier and reproducible. In cases with no defined tumor in CT, other tools are needed and MRI can be complementary. The interobserver variability in target delineation of hepatocellular carcinoma in this study is noteworthy.*

Correspondence to: Carolina de la Pinta, Radiation Oncology Department, Ramón y Cajal University Hospital, IRYCIS, Alcalá University, 28034 Madrid, Spain. E-mail: cdelapinta88@gmail.com

Key Words: Contouring, segmentation, variability, liver, hepatocellular carcinoma, SBRT.

The definition of the gross tumor volume (GTV) or macroscopic tumor in hepatocellular carcinoma sometimes requires the use of magnetic resonance imaging (MRI) complementary to planning computed tomography (CT), which can improve the precision in tumor delineation (1). There are publications that try to resolve the most important doubts in the delineation of hepatocellular carcinoma, but there is still no consensus (2, 3). The absence of consensus on the best imaging test to use to define treatment volumes in hepatic tumors requires studies to clarify this issue. Therefore, an observational study was designed to analyze the interobserver variability in CT and MRI of hepatocellular carcinoma.

Materials and Methods

The study was approved by the center's Ethics Committee (no. 227/17). For this purpose, we compared the tumor size in the volume delineated in each of the imaging tests used, analyzed the usefulness and ease of use of CT and MRI in the delineation, and finally determined the best imaging test for the delineation of tumor volumes.

The patients included in the study were patients over 18 years of age, with a diagnosis of hepatocellular carcinoma, and with available triphasic CT and MRI performed in their diagnostic process within routine clinical practice. The images were obtained by an Aquileon® multislice helical CT (64 slices) (Toshiba, Tokyo, Japan), and with an Achieva® 1.5 Tesla MRI (Philips Healthcare, Amsterdam, the Netherlands). All images were transferred from the Picture Archiving and Communication System to the Pinnacle3 planning system (Philips Healthcare) and anonymized.

Eight radiation oncologists participated in the study. The oncologists were blinded to each other's defined contours. The specialists had the patient's medical history and the corresponding radiological report of each test to assist them in the delineation of the GTVs. The documentation provided included a hepatic scheme by segments indicating the segment where the lesion was located. The CT and MRI were merged in the Pinnacle3 planning system (Philips Healthcare). The window level was not preset, and the oncologist was able to adjust the width and the window level in the planner, being advised to use the abdomen window. Craniocaudal and coronal reconstructions were available in the planner. Delineation of organs at risk was not performed.

Statistical analysis. The GTV volumes (in cm^3) of the delineated contours in each patient in each of the imaging tests were calculated by determining the maximum volume, minimum volume, median volume, variance and standard deviation. The difference between the delineated volumes in each imaging test was analyzed for each of the lesions. The overlap ratio was defined as the proportion of overlapping volume in the delineations made by the different observers. A perfect agreement between the two GTVs is indicated by a value of 1. The ratio was 0 when at least one of the oncologists had contoured at a location with no overlap with the others.

To establish the ideal test in each of the cases, the index of variability obtained in each of the tests was assessed by calculating the variance and standard deviation, the overlap ratio and the correlation coefficient.

A qualitative analysis was included by means of a questionnaire showing the difficulties expressed by the radiation oncologists in tumor delineation. Those statements in which the specialists agreed on ≥80% of the answers were taken into account, and the questions that reached this level were considered to represent consensus.

Results

We included five males with radiological or histological diagnosis of hepatocellular carcinoma. The mean age of the patients was 64.8 years (range=55-78 years). The location of the liver tumors was: segment II in two lesions, segment IV in one, segment V in one, and segment VI lesion in one.

The median volume in CT and MRI of each patient, their difference and the percentage variation between radiation oncologists were analyzed. The median GTV volume on MRI was 2.4 cm^3 (range=0.59-15.6 cm^3), compared to 3.5 cm^3 (range=0.52-24.9 cm^3) on CT (p=0.36). The GTV as defined on CT was larger or at least as large as the GTV on MRI in two cases. Variance and standard deviation between observers in CT and MRI were minor (6 *vs*. 7.87 cm^3, and 2.5 *vs*. 2.8 cm^3, respectively).

Table I. *Median gross tumor volume by computed tomography (CT) and magnetic resonance imaging (MRI) compared.*

		Median volume, cm^3		Median difference	
Patient	Segment	By CT	By MRI	cm^3	%
1	II	0.52	0.59	0.07	11.9
2	V	24.9	15.6	–9.3	–37.3
3	II	3.7	2.4	–1.3	–35.1
4	VIII	0.9	1	0.1	10
5	VI	3.5	5.4	1.9	35.2

Differences between CT and MRI greater than 20% were found in three patients (Table I). The largest difference between median volumes from the different tests was seen in patient 2, where the contours were at disparate sites, with a 37.3% difference between volumes delineated by CT and MRI. In patient 3, there was a difference in volumes of 35.1% between oncologists in CT *versus* MRI (Table II). The lowest variance was observed for patient 1, despite their having a small tumor. Of note, patient 5, with previous transcatheter arterial chemoembolization (TACE) and remaining tumor was well defined by all oncologists, with low variance in CT (Figure 1, showing contouring of tumor in patient 5).

The overlap ratio between the different observers in each imaging test was calculated (Table II). For all patients there was a lower overlap ratio with CT. For patient 3, there was no overlap for either of the imaging modalities used. The overlap ratio was low for patients 1 and 4 in CT. With these results, we analyzed the possible causes of the greater differences between the oncologists. In the CT of patients 2 and 4, poor contrast uptake of the lesion was observed. In case 3, an oncologist identified as tumor an anfractuous and heterogeneous mass in hepatic segment IVb, in close relation with left portal branch, predominantly hypodense in all phases and with peripheral nodular enhancement, compatible with chronic thrombosis of the left portal branch with data for revascularization and portal cavernomatosis (Figure 2).

Quantitative data were compared using variance and overlap ratio and qualitative data using the survey conducted by the oncologists included information related to the grade of difficult of contouring cases (Table III). For the quantitative study, the imaging test with the lowest variance and the test with the best overlap ratio were included. Correlation between quantitative and qualitative analysis was observed for only three patients (patients 1, 4 and 5).

Table II. *Variance, standard deviation and overlap ratio of gross tumor volume determined by five oncologists by computed tomography (CT) and magnetic resonance imaging (MRI).*

Patient	CT			MRI		
	Variance	Standard deviation	Overlap ratio	Variance	Standard deviation	Overlap ratio
1	0.017	0.13	0.17	0.087	0.29	0.16
2	45.46	6.75	0.45	7.87	2.8	0.21
3	927.9	30.46	0	1742	41.7	0
4	6	2.5	0	0.4	0.64	0.5
5	0.16	0.4	0.58	10.29	3.2	0.18

Discussion

Our results indicate considerable variations in the delineation of GTVs among the participating radiation oncologists. The greatest variability between CT and MRI was observed in patient 2 (median CT volume 24.9 cm^3 and median MRI volume 15.6 cm^3). The oncologists consensually chose CT as the best imaging test for tumor delineation for this patient. Analyzing the images in detail, uptake of intravenous contrast by the atypical lesion and large motion artifact on MRI were observed, which would explain these differences and the oncologist's selection of CT as the test of choice.

In case 4, there was not much variability when comparing the median volumes in CT and MRI, however, an adequate overlap ratio was not achieved in either CT or MRI. In CT, this might have once again been due to contrast uptake of the lesion. In MRI, it was probably due to motion artifact because the lesion was located in the hepatic dome, a very mobile area. The delineation of volumes in simple MRI was subjectively considered and MRI was established by the oncologists as the one of the best imaging tests for tumor delineation.

On the other hand, patient 1 presented little variability in CT and MRI (median volume in CT 0.52 cm^3 and median volume in MRI 0.59 cm^3), with similar overlap ratio (median volume in CT 0.17 cm^3 and median volume in MRI 0.16 cm^3), and in the subjective evaluation of the oncologists, CT allowed easier delineation, the lesion being more visible than in MRI. This may be due to the fact that despite being a small lesion, it presented typical contrast uptake, that the imaging test most used and to which radiation oncologists are most accustomed is CT, and that MRI of small lesions or very mobile regions can be complex to interpret. The same occurred with patient 5, whose data presented adequate correlation with typical contrast uptake.

These results are in agreement with available data, where the authors demonstrated that delineation was better and showed a greater consensus in lesions with a typical pattern of contrast uptake (4, 5). We must take into account that hepatic deformations range from 2.8 to 10.7 mm according to studies on rigid registration in CT and MRI. Vásquez Osorio *et al*. studied vessel-guided image registration and concluded that deformable registration is necessary to achieve adequate alignment in abdominal compression or free breathing (6). The use of deformable image registration algorithms can help improve liver alignment from different sets of images when liver deformation occurs (7). Moreover, images should be fused as closely as possible to the target region based on both liver contouring and, potentially, vascular anatomy, in the same way as was done in this study. In our study, the alignment of the entire liver was poor, being optimal in the tumor region, because of the limitations of the registration algorithms. The behavior of the liver is complex and this has an impact on its registration, which is why some authors have advised against the use of MRI (2, 3). Furthermore, it is advisable that the immobilization used to restrict respiratory movement in CT also be used to acquire the MR image, avoiding differences in position and changes due to breathing (8); this was not possible in our study due to the incompatibility of the immobilization system and the MRI. In our study, we used images which were not always acquired for abdominal stereotactic body radiotherapy, making this fusion difficult. Another factor that can influence the measured volume is the window configuration, which is more important for small lesions (2). Some authors have compared the results of MRI and CT with contrast with histopathological data after resection, showing that MRI was better for delimiting tumors <2 cm but not for tumors ≥2 cm, which would indicate that it may play a role in small lesions, as was the case in patients 1 and 4, and may explain the important variability between one test and another in the contouring of the rest of the patients. Other studies show high sensitivity of MRI (8-10) and it can be the test of choice for the delineation of lesions; however, the lack of concurrence between contour volumes in our study suggests that they are complementary tests.

In addition to the problems with contrast uptake of the lesions and image fusion, in patient 3 there was no overlap

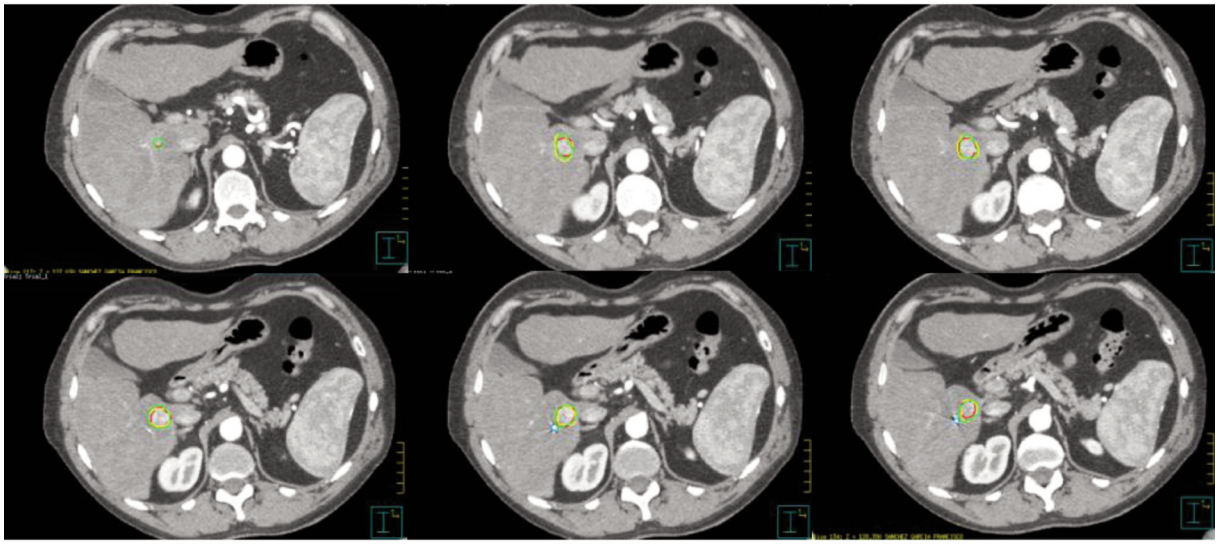

Figure 1. *Delineation of the gross tumor volume in patient with hepatocellular carcinoma previously treated with transcatheter arterial chemoembolization (case number 5). Each color represents the evaluation by one assessor.*

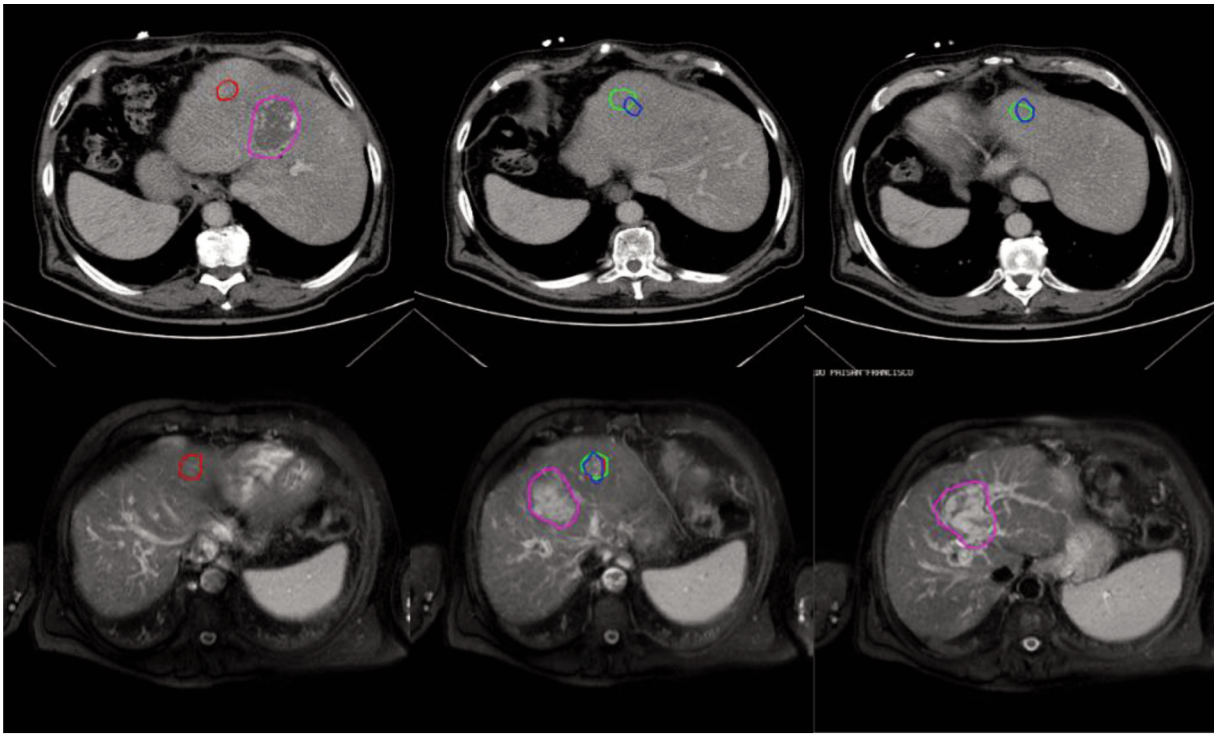

Figure 2. *Patient with hepatocellular carcinoma and chronic portal thrombosis, by computed tomography (upper row) and magnetic resonance (lower row) images (case number 3). Each color represents the evaluation by one assessor.*

in any of the imaging tests, probably due to the confusion of an arterial thrombus with a tumor thrombus. This situation led to the delineation of a non-tumorous alteration as a tumor. In addition, in this patient there were important differences in the median volume between CT and MRI. In spite of this, the oncologists agreed that MRI was the simpler technique. This is in agreement with the data from the study by Hong *et al*., who analyzed interobserver variability in

Table III. *Quantitative and qualitative analysis comparing contouring by imaging modality for each patient according to consensus in oncologists' questionnaire responses.*

Patient	Lower variance	Higher overlap ratio	Quantitative	Qualitative	Concordance
1	CT	CT	CT	CT	Yes
2	MRI	CT	MRI/CT	CT	Yes, MRI support
3	CT	-	CT	MRI	No
4	MRI	MRI	MRI	MRI	Yes
5	CT	CT	CT	CT	Yes

CT: Computed tomography; MRI: magnetic resonance imaging.

hepatocellular carcinoma with or without portal venous thrombosis; fewer false-negative contours of hepatocellular carcinoma were observed in MRI than in multiphase CT (17% *vs*. 30% of the total of 83 lesions, respectively) (2).

One of the factors related to difficulty and variability in contouring is delineation after previous TACE, and there is no consensus on the treatment volume (3). The inclusion of the tumor area previously treated with TACE as a GTV is recommended when the recurrence is marginal; in all other situations, the treatment volume should be assessed in a multidisciplinary manner. Our study showed great concordance probably due to imaging being performed more than 1 month after TACE.

In our series, the factors related to variability in delineation in hepatocellular carcinoma were the poorly defined tumor margin due to contrast uptake or tumor size, the location of the tumor and its impact on image fusion due to movement, and the presence of non-tumor thrombosis. In our study, CT contouring can be considered the standard test, with MRI support being necessary for small lesions, lesions with atypical contrast uptake, or for patients with portal invasion. In these cases, MRI acquisition under the same conditions as the planning CT is recommended.

MRI was shown to underestimate tumor size in a surgical cohort with radiological and pathological correlation (11, 12). In our study, MRI images did not align well with planning CT images due to differences in patient positioning and movement during testing.

Some uncertainty in the delineation of GTV is unavoidable. Some studies have suggested compensating for this uncertainty by using the clinical target volume to address the possible microscopic extent of the tumor on the static planning image. The magnitude of the expansion needed to create the clinical target volume might vary depending on the oncologist's uncertainty in identifying the GTV, although these studies do not recommend its routine use, this being very controversial (11, 13).

Limitations. One of the most important problems described and observed in our study is the image registration system, for which one solution would be the use of CT- and MRI-compatible fiducials to aid in registration. However, MRI is often prior to consultation in radiation oncological, or may be artifactual with the use of fiducials. In addition, our radiation oncologists were not experts in abdominal stereotactic body radiotherapy, therefore it would be interesting to identify a series in which the tumor is best visualized and consult with a hepatobiliary radiologist, promoting contouring education with discussion in consensus meetings (4, 14). Other solutions to improve and reduce interobserver variability in contouring are the use of simulation CT with slices every 3 mm, especially for small tumors; the use of a respiratory motion restriction method, and 4D CT for patients with large respiratory motion (more than 1-1.5 cm); the use of multiphase CT; the use of fusion MRI using vascular structures as landmarks; the use of an abdominal window in CT and reviewing the sagittal and coronal planes; and avoiding delimiting perfusion abnormalities.

Conclusion

In contouring in hepatocellular carcinoma, in patients with well-defined tumors, CT is easier and reproducible. In cases with no defined tumor in CT, other tools are needed and MRI can be complementary. The interobserver variability in target delineation of hepatocellular carcinoma in this study is noteworthy. This study identifies interobserver variability as a relevant limitation in delineating hepatocellular carcinoma lesions when SBRT is planned by less-experienced oncologists.

Conflicts of Interest

The Authors declare no conflicts of interest.

Authors' Contributions

Data curation: Juan David García, David Sevillano, Rafael Colmenares, Raquel García Latorre, Manuel Garví, Vanesa Pino, Mercedes Martín, Eva Fernández, Raul Hernanz, Margarita Martín, Jose Antonio Domínguez and Teresa Muñóz. Formal analysis and methodology: Alfonso Muriel. Conceptualization and supervision:

Luis Cristian Perna, Agustín Albillos and Sonsoles Sancho. Conceptualization, data curation, formal analysis, methodology, supervision and writing-original draft, writing-review and editing: Carolina de la Pinta.

References

1 Fernández-Letón P, Baños C, Bea J, Delgado Rodríguez JM, De Blas Piñol R, Martínez Ortega J, Pérez Moreno JM and Simón de Blas R: Recomendaciones de la Sociedad Española de Física Médica (SEFM) sobre implementación y uso clínico de radioterapia estereotáxica extracraneal (SBRT). SEFM Rev Fis Med 18(2): 77-142, 2017. Available at: https://revistadefisicamedica.es/index.php/rfm/article/view/247/253 [Last accessed on April 10, 2023]

2 Hong TS, Bosch WR, Krishnan S, Kim TK, Mamon HJ, Shyn P, Ben-Josef E, Seong J, Haddock MG, Cheng JC, Feng MU, Stephans KL, Roberge D, Crane C and Dawson LA: 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 89(4): 804-813, 2014. PMID: 24969794. DOI: 10.1016/j.ijrobp.2014.03.041

3 Kim YS, Kim JW, Yoon WS, Kang MK, Lee IJ, Kim TH, Kim JH, Lee HS, Park HC, Jang HS, Kay CS, Yoon SM, Kim MS and Seong J: Interobserver variability in gross tumor volume delineation for hepatocellular carcinoma: Results of Korean Radiation Oncology Group 1207 study. Strahlenther Onkol 192(10): 714-721, 2016. PMID: 27538775. DOI: 10.1007/s00066-016-1028-2

4 Nestle U, Rischke HC, Eschmann SM, Holl G, Tosch M, Miederer M, Plotkin M, Essler M, Puskas C, Schimek-Jasch T, Duncker-Rohr V, Rühl F, Leifert A, Mix M, Grosu AL, König J and Vach W: Improved inter-observer agreement of an expert review panel in an oncology treatment trial--Insights from a structured interventional process. Eur J Cancer 51(17): 2525-2533, 2015. PMID: 26277100. DOI: 10.1016/j.ejca.2015.07.036

5 Schimek-Jasch T, Troost EG, Rücker G, Prokic V, Avlar M, Duncker-Rohr V, Mix M, Doll C, Grosu AL and Nestle U: A teaching intervention in a contouring dummy run improved target volume delineation in locally advanced non-small cell lung cancer: Reducing the interobserver variability in multicentre clinical studies. Strahlenther Onkol 191(6): 525-533, 2015. PMID: 25665799. DOI: 10.1007/s00066-015-0812-8

6 Vasquez Osorio EM, Hoogeman MS, Mendez Romero A, Wielopolski P, Zolnay A and Heijmen BJ: Accurate CT/MR vessel-guided nonrigid registration of largely deformed livers. Med Phys 39(5): 2463-2477, 2012. PMID: 22559617. DOI: 10.1118/1.3701779

7 Xu H, Gong G, Wei H, Chen L, Chen J, Lu J, Liu T, Zhu J and Yin Y: Feasibility and potential benefits of defining the internal gross tumor volume of hepatocellular carcinoma using contrast-enhanced 4D CT images obtained by deformable registration. Radiat Oncol 9: 221, 2014. PMID: 25319176. DOI: 10.1186/s13014-014-0221-7

8 Chou R, Cuevas C, Fu R, Devine B, Wasson N, Ginsburg A, Zakher B, Pappas M, Graham E and Sullivan SD: Imaging techniques for the diagnosis of hepatocellular carcinoma: a systematic review and meta-analysis. Ann Intern Med 162(10): 697-711, 2015. PMID: 25984845. DOI: 10.7326/M14-2509

9 Tsurusaki M, Sofue K, Isoda H, Okada M, Kitajima K and Murakami T: Comparison of gadoxetic acid-enhanced magnetic resonance imaging and contrast-enhanced computed tomography with histopathological examinations for the identification of hepatocellular carcinoma: a multicenter phase III study. J Gastroenterol 51(1): 71-79, 2016. PMID: 26130441. DOI: 10.1007/s00535-015-1097-5

10 Kanematsu M, Semelka RC, Leonardou P, Mastropasqua M and Lee JK: Hepatocellular carcinoma of diffuse type: MR imaging findings and clinical manifestations. J Magn Reson Imaging 18(2): 189-195, 2003. PMID: 12884331. DOI: 10.1002/jmri.10336

11 Ghandi D, Mitchell KA, Miick R and Chang BW: Clinicopathologic analysis of gross tumor size and microscopic extension in liver metastases from colorectal cancer (CRC): Implications for stereotactic body radiotherapy (SBRT). Proceedings of the American Society for Radiation Oncology 51st Annual Meeting 01-05 November 2009. Int J Radiat Oncol Biol Phys 75(3): S29, 2009. DOI: 10.1016/j.ijrobp.2009.07.089

12 Tanenbaum D, Hall W, Mittal P, Colbert L, Kim S, Squires M, Fisher S, Kooby D, Maithel S, Staley C, Yu D, Curran W and Landry J: Radiographic to pathologic correlation of gross tumor size of hepatic metastases from colorectal cancer. Int J Radiat Oncol Biol Phys 93(3): S205, 2018. DOI: 10.1016/j.ijrobp.2015.07.491

13 Méndez Romero A, Verheij J, Dwarkasing RS, Seppenwoolde Y, Redekop WK, Zondervan PE, Nowak PJ, Ijzermans JN, Levendag PC, Heijmen BJ and Verhoef C: Comparison of macroscopic pathology measurements with magnetic resonance imaging and assessment of microscopic pathology extension for colorectal liver metastases. Int J Radiat Oncol Biol Phys 82(1): 159-166, 2012. PMID: 21183292. DOI: 10.1016/j.ijrobp.2010.10.032

14 Eriksen JG, Salembier C, Rivera S, De Bari B, Berger D, Mantello G, Müller AC, Martin AN, Pasini D, Tanderup K, Palmu M, Verfaillie C, Pötter R, Valentini V and ESTRO: Four years with FALCON - an ESTRO educational project: achievements and perspectives. Radiother Oncol 112(1): 145-149, 2014. PMID: 25070586. DOI: 10.1016/j.radonc.2014.06.017

Received March 13, 2023
Revised April 7, 2023
Accepted April 13, 2023