



## Optimal Phase of Dynamic Computed Tomography for Reliable Size Measurement of Metastatic Neuroendocrine Tumors of the Liver: Comparison between Pre- and Post-Contrast Phases

Jimi Huh, MD<sup>1,2</sup>, Jisuk Park, MS<sup>1</sup>, Kyung Won Kim, MD<sup>1</sup>, Hyoung Jung Kim, MD<sup>1</sup>, Jong Seok Lee, MD<sup>1</sup>, Jong Hwa Lee, MD<sup>3</sup>, Yoong Ki Jeong, MD<sup>3</sup>, Atul B. Shinagare, MD<sup>4</sup>, Nikhil H. Ramaiya, MD<sup>5</sup>

<sup>1</sup>Department of Radiology and Research Institute of Radiology, University of Ulsan College of Medicine, Asan Medical Center, Seoul 05505, Korea; <sup>2</sup>Department of Radiology, Ajou University School of Medicine and Graduate School of Medicine, Ajou University Hospital, Suwon 16499, Korea; <sup>3</sup>Department of Radiology, University of Ulsan College of Medicine, Ulsan University Hospital, Ulsan 44033, Korea; <sup>4</sup>Department of Imaging, Dana-Farber Cancer Institute, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02215, USA; <sup>5</sup>Department of Radiology, University Hospitals Cleveland Medical Center, Case Western Reserve University, Cleveland, OH 44106, USA

**Objective:** The reliability of size measurements of liver metastases from neuroendocrine tumors (NETs) on contrast-enhanced computed tomography (CT) phases made by different readers may be hampered due to transient, variable rim enhancement in arterial phase (AP) or portal venous phase (PVP) images. We aimed to assess the reliability of tumor size measurements in pre- and post-contrast scans.

**Materials and Methods:** The study coordinator selected target lesions according to Response Evaluation Criteria in Solid Tumors 1.1 guidelines in 44 consecutive patients with pathologically confirmed NET liver metastases. Two blinded readers measured the longest diameters of target lesions on pre-contrast, AP, and PVP images twice with a 4-week interval. Inter- and intra-observer agreements were evaluated using Bland-Altman plots and 95% limit of agreement (LOA) calculations.

**Results:** Of the 79 target lesions (approximate mean size of 3 cm), 45 showed rim enhancement. Inter-observer agreement assessed based on LOA was highest in pre-contrast CT images (-6.1–5.7 mm), followed by PVP (-7.9–7.1 mm) and AP (-8.5–7.4 mm) images. Intra-observer agreement showed the same trend: -2.8–2.9 mm and -2.9–2.9 mm for readers 1 and 2, respectively, on pre-contrast CT, -2.8–2.9 mm and -3.0–3.2 mm, respectively, on PVP, and -3.2–4.2 mm and -3.4–3.2 mm, respectively, on AP images. Mean tumor diameters differed significantly among the phases in the following increasing order: pre-contrast CT, PVP, and AP images.

**Conclusion:** There was better inter- and intra-observer agreement in size measurements of NET liver metastases on pre-contrast scans than on AP and PVP scans. Pre-contrast CT may be the optimal for measuring NET liver metastases if its accuracy is proven.

**Keywords:** Neuroendocrine tumor; RECIST; Pre-contrast; Unenhanced; Computed tomography; Measurement; Reliability; Repeatability; Reproducibility; Observer agreement

Received October 9, 2017; accepted after revision April 22, 2018.

This study was supported by grants from the Korean Health Technology R&D Project (HI17C1862), the National Research Foundation of Korea (2017R1A2B3011475), and the Asan Medical Center (2017-0467).

**Corresponding author:** Kyung Won Kim, MD, Department of Radiology and Research Institute of Radiology, University of Ulsan College of Medicine, Asan Medical Center, 88 Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Korea.

• Tel: (822) 3010-4377 • Fax: (822) 476-4719 • E-mail: medimash@gmail.com

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## INTRODUCTION

In this era of rapid anticancer drug development, accurate tumor burden assessment is an essential component of clinical trials of both new drugs and routine anticancer treatments. Tumor burden, which is based on tumor size, is generally assessed on contrast-enhanced computed tomography (CT) images according to the Response Evaluation Criteria in Solid Tumors (RECIST) (1, 2). RECIST 1.1 is essentially based on unidimensional assessment of the tumor burden on cross-sectional imaging modalities (2). However, size measurements might be inconsistent and are not always reproducible among different readers and in different measurement sessions (3).

Neuroendocrine tumors (NETs) are primarily hypervascular tumors that mainly metastasize to the liver (4). In several previous studies of NET liver metastasis, the hepatic arterial phase (AP) was considered the most relevant phase to visualize liver metastasis due to its strong arterial enhancement (5, 6). Indeed, one study found that the hepatic AP revealed the greatest number of metastases of NETs (5). However, these liver metastases frequently exhibit peripheral rim enhancement in the AP or portal venous phase (PVP) of dynamic CT (7, 8). This phenomenon might hinder precise and reliable tumor size measurement, and thus affect assessments of treatment response.

In general, pre-contrast CT images are not used for tumor size measurement for RECIST 1.1, mainly because the visual conspicuity of the tumor is lower in pre-contrast CT images when compared to AP or PVP images (5). However, if a tumor is clearly visualized on pre-contrast CT, the tumor size measurement is not affected by contrast enhancement (4). Therefore, we hypothesized that pre-contrast CT might facilitate reliable tumor measurement in patients with NET liver metastases. We aimed to compare pre- and post-contrast CT scans in terms of reliability in tumor size measurement.

## MATERIALS AND METHODS

Our Institutional Review Board approved this retrospective observational study and waived the requirement for informed consent. We followed the Strengthening of the Reporting of Observational Studies in Epidemiology recommendations for the performance and reporting this study (9).

### Patients

We performed a systematic computerized search of our hospital's data warehouse (Asan Biomedical research Environment) from January 2005 to December 2014. We used the pathological diagnosis search terms "neuroendocrine OR endocrine OR carcinoid" and the procedure codes for "liver biopsy" OR "hepatectomy." Through our search strategy, we identified 101 patients with pathologically confirmed liver metastases of NETs. We then selected patients who met the following inclusion criteria: 1) pathologically confirmed liver metastases from NETs; 2) dynamic abdominal CT prior to any kind of treatment for liver metastases, including chemotherapy, surgery, radiofrequency ablation, or transarterial chemoembolization; and 3) the presence of measurable (tumor size > 1 cm in long diameter) metastatic lesions in the liver according to RECIST 1.1. Of these 101 patients, 57 were excluded for the following reasons: 52 did not undergo appropriate dynamic abdominal CT at our institution prior to treatment for liver metastasis, and five had no measurable lesions > 1 cm in diameter. Finally, 44 patients with 79 target lesions were included in this study.

### Imaging Acquisition

All CT examinations were performed using a Somatom Sensation 16 (Siemens Healthineers, Erlangen, Germany), Somatom Definition (Siemens Healthineers), LightSpeed 16 (GE Healthcare; Milwaukee, WI, USA), or LightSpeed VCT (GE Healthcare) scanner. Dynamic CT scans comprising non-enhanced, arterial, and PVP images were obtained for all patients.

The scan parameters for the Somatom Sensation 16 and Definition scanners were as follows: beam collimation of 16 x 0.75 mm, and 64 x 0.6 mm, respectively; beam pitch of 1; gantry rotation time of 0.5 seconds; 120 kV; and automated dose modulation (CareDose 4D, Siemens Healthineers) using the maximum allowable tube current (200 mA). The scan parameters for the LightSpeed 16 and VCT scanners were as follows: beam collimation of 16 x 1.25 mm, and 64 x 0.625 mm, respectively; beam pitch of 0.984; gantry rotation time of 0.5 seconds; 120 kV; and automated dose modulation (automA and smartmA, GE Healthcare) primarily modulated based on a noise index of 12.81. The same scan settings were used for the different scan phases.

For dynamic multiphasic CT, 100–120 mL of iopromide (Ultravist 370 or Ultravist 300; Bayer Schering Pharma; Berlin, Germany) were administered intravenously at a

rate of 3 mL/sec using an automatic power injector. Bolus tracking was used to determine the timing of the arterial and PVP. AP images were obtained with a 10-second delay on the Sensation 16 and LightSpeed 16 scanners and with a 15-second delay on the Somatom Definition and LightSpeed VCT scanners after aortic attenuation at the thoracolumbar junction had reached a pre-defined threshold of an increase of 100 Hounsfield units (HUs) above baseline, i.e., pre-contrast attenuation. PVP images were obtained using a fixed 75-second delay. Axial CT images were reconstructed with a section thickness of 2.5–5 mm.

### Image Analysis

The study coordinator (a cancer imaging and abdominal radiology specialist with 9 years of clinical experience) anonymized the subjects' dynamic abdominal CT scans to ensure that the readers were blinded to the subjects' clinical information. Per RECIST 1.1, the study coordinator selected up to two target liver metastatic lesions and provided the target lesion information as thumbnail pictures to the readers to ensure target lesion consistency. In the target lesion selection process, the study coordinator selected the best-defined and most reproducibly measurable lesions on contrast-enhanced CT according to RECIST 1.1. He selected the target lesions without any preference for rim enhancement while being blinded to the pre-contrast CT findings.

Two readers with considerable experience with RECIST 1.1 in clinical trials (an abdominal radiologist with 8 years of clinical experience and an image analyst with 7 years of experience) participated in the tumor measurement process. On the basis of information regarding the predefined target lesions, both readers measured the unidimensional longest diameters of the target lesions on the pre-contrast phase, AP, and PVP CT images. Tumor sizes were measured using electronic calipers on a picture archiving and communication system (PACS). The readers repeated the tumor measurements twice, with a 4-week interval between the first and second measurement sessions to minimize recall bias. No annotation markers were left in the PACS between measurement sessions.

When there was a peripheral enhancing rim in the tumor, the readers were instructed to measure tumor size by including the peripheral enhancing rim. The peripheral enhancing rim was defined as an enhancing rim-like structure surrounding the tumor that had higher enhancement than the tumor center.

While measuring the tumor diameter in the pre-contrast phase, AP, and PVP, the readers were allowed to refer to the three phase images from the same study subject during the same measurement session. However, they were blinded to the results from the other measurement session and to each other's measurement results.

To evaluate CT image noise, we placed five regions of interest (ROIs) outside the subjects and measured CT values in HU in the ROIs. The mean standard deviation (SD) of the five ROIs was considered the CT image noise (10). The radiation doses of the CT scans were evaluated using the volume of the CT dose index (11).

### Statistical Analysis

Mean and SD values were determined for all measurements. Repeated-measures analyses of variance (RMANOVAs) and post hoc tests with Bonferroni correction were used to compare the mean values of these measurements. For the RMANOVAs, the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity.

Inter-observer agreement and intra-observer agreement were evaluated using Bland-Altman plots with the mean difference and 95% limit of agreement (LOA) (12). Agreement between two measurements was visually assessed using Bland-Altman plots in which the differences were plotted against the average values of the two size measurements. The difference between two measurements was defined as  $M1 - M2$ , where  $M1$  = measurement 1 and  $M2$  = measurement 2, and the mean difference was the average value of the differences between all tumors. The 95% LOA was the range of the mean difference  $\pm 1.96 \times SD$  of the mean difference. The RMANOVA and Bland-Altman plot calculations were performed using GraphPad Prism 7 (GraphPad Software, La Jolla, CA, USA).

The repeatability coefficient (RC) was calculated using a web-based calculator used to compute the RC for two repeated measurements of a continuous parameter (available at <http://datasharing.aim-aicro.com/reliability>), as described elsewhere (13-15).

To identify the effects of peripheral rim enhancement of the target lesions on tumor measurements, we performed a subgroup analysis by dividing subjects into groups with and without peripheral rim enhancement.

## RESULTS

### Patients

The characteristics of the study subjects and the CT parameters are summarized in Table 1. Of the 44 patients, 35 had multiple liver metastases and 9 had single liver metastases, yielding a total of 79 target lesions. Peripheral rim enhancement was independently observed in 45 of the 79 (56.9%) lesions by reader 1 and in 42 of the 79 (53.1%) lesions by reader 2. The study coordinator reviewed the discrepant cases and confirmed that there were 45 (56.9%) lesions with peripheral rim enhancement. In all of these cases, peripheral rim enhancement was more prominent in AP images than in PVP images. Of the 79 target lesions, six

target lesions (7.6%) with a mean tumor size of 28.3 mm on the pre-contrast images were subtle on pre-contrast CT, as assessed by both readers.

### Tumor Size Measurement

The mean diameters of the target lesions measured by readers 1 and 2 during the first and second measurement sessions, and the resulting inter- and intra-observer mean differences and 95% LOAs are summarized in Tables 2 and 3, respectively. The mean tumor diameters measured by readers 1 and 2 were the largest in the AP (31.3 mm and 31.9 mm, respectively, in session 1; and 30.8 mm and 32.0 mm, respectively, in session 2), followed by the PVP (29.8 mm and 30.2 mm, respectively, in session 1; and 29.8 mm and

**Table 1. Characteristics of Subjects/Lesions and CT Parameters**

Item	Characteristics
Subjects/lesions	
Age (mean ± SD)	59.3 ± 15.1 years (range, 38–79 years)
Sex (%)	Male 19 (43.2), female 25 (56.8)
BMI (mean ± SD)	Male 23.8 ± 4.1 kg/m <sup>2</sup> , female 21.9 ± 4.5 kg/m <sup>2</sup>
Primary n tumor sites (%)	Pancreas (n = 27, 61.4) Colorectum (n = 6, 13.6) Gallbladder (n = 4, 9.1) Unknown primary site (n = 4, 9.1) Lung/bronchus (n = 2, 4.5) Stomach (n = 1, 2.3)
Target lesion location in liver (%)	S1 (n = 0, 0), S2 (n = 6, 7.6), S3 (n = 7, 8.9), S4 (n = 17, 21.5), S5 (n = 11, 13.9), S6 (n = 10, 12.7), S7 (n = 9, 11.4), and S8 (n = 19, 24.0)
CT parameters	
Slice thickness	2.5 mm (n = 5), 3 mm (n = 12), 5 mm (n = 27)
Image noise	Pre-contrast (7.7 ± 4.9 HU) AP (8.5 ± 5.5 HU) PVP (8.2 ± 5.2 HU)
Radiation dose in CTDIvol	Pre-contrast (9.4 ± 2.4 mGy) AP (9.7 ± 3.7 mGy) PVP (10.3 ± 4.6 mGy)

AP = arterial phase, BMI = body mass index, CT = computed tomography, CTDIvol = volume of CT dose index, HU = Hounsfield Unit, mGy = milligray, PVP = portal venous phase, S = segment of liver, SD = standard deviation

**Table 2. Mean Diameters of Target Lesions at Each Measurement Session**

	Pre-Contrast Phase (mm)	AP (mm)	PVP (mm)	P of RMANOVA	P of Post Hoc Tests*
Session 1					
Reader 1	28.9 ± 18.9	31.3 ± 19.9	29.8 ± 20.0	< 0.001	< 0.001, < 0.001, 0.05
Reader 2	29.1 ± 19.2	31.9 ± 20.2	30.2 ± 19.8	< 0.001	< 0.001, < 0.001, 0.0105
Session 2					
Reader 1	28.8 ± 18.9	30.8 ± 19.6	29.8 ± 19.8	< 0.001	0.0025, < 0.001, 0.0467
Reader 2	29.0 ± 18.9	32.0 ± 20.4	30.1 ± 19.4	< 0.001	0.001, < 0.001, 0.0123

Data: mean ± SD. \*p values of post-hoc tests between pre-contrast vs. AP, AP vs. PVP, and PVP vs. pre-contrast. RMANOVA = repeated-measures analysis of variance

**Table 3. Inter-Observer and Intra-Observer Agreement**

	Pre-Contrast Phase (mm)	AP (mm)	PVP (mm)
Inter-observer agreement (session 1)			
Mean difference	-0.2	-0.6	-0.4
95% LOA	-6.1–5.7	-8.5–7.4	-7.9–7.1
RC	5.9	8.0	7.5
Inter-observer agreement (session 2)			
Mean difference	-0.2	-1.2	-0.3
95% LOA	-6.1–5.7	-10.2–7.8	-8.3–7.7
RC	5.9	9.2	7.9
Intra-observer agreement (reader 1)			
Mean difference	0.1	0.5	0.1
95% LOA	-2.8–2.9	-3.2–4.2	-2.8–2.9
RC	2.8	3.8	3.4
Intra-observer agreement (reader 2)			
Mean difference	0.0	-0.1	0.1
95% LOA	-2.9–2.9	-3.4–3.2	-3.0–3.2
RC	2.9	3.3	3.0

LOA = limit of agreement, RC = repeatability coefficient

30.1 mm, respectively, in session 2) and the pre-contrast phase (28.9 mm and 29.1 mm, respectively, in session 1; and 28.8 mm and 29.0 mm, respectively, in session 2).

### Inter-Observer and Intra-Observer Agreement

Table 3 and Figure 1 present Bland-Altman analysis results and plots. RCs are also presented in Table 3. Pre-contrast CT yielded the best inter-observer agreement, as assessed using 95% LOAs, in both the first and second measurement sessions (-6.1–5.7 mm in both sessions). This was followed by the PVP (-7.9–7.1 mm and -8.3–7.7 mm, respectively) and AP (-8.5–7.4 mm and -10.2–7.8 mm, respectively). Pre-contrast CT also yielded the best intra-observer agreement, as assessed using 95% LOAs, for both readers 1 and 2 (-2.8–2.9 mm and -2.9–2.9 mm, respectively). This was followed by the PVP (-2.8–2.9 mm and -3.0–3.2 mm, respectively) and the AP (-3.2–4.2 mm and -3.4–3.2 mm, respectively). In all CT phases, intra-observer agreement was better than inter-observer agreement for both readers.

### Subgroup Analysis

A subgroup analysis based on the presence of peripheral rim enhancement on target lesions revealed that ratings of tumors lacking rim enhancement exhibited better inter-observer and intra-observer agreement when compared to those of tumors with peripheral rim enhancement (Table 4, Fig. 2). This tendency was obvious in the AP and PVP rather than in the pre-contrast phase. These results

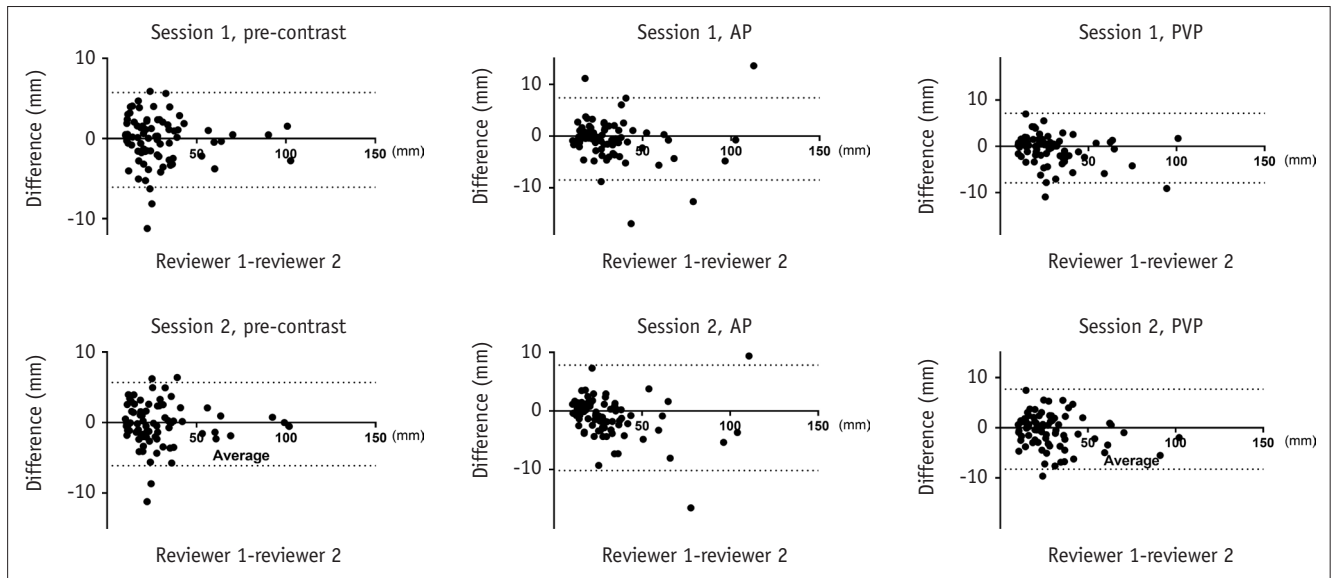
indicate that peripheral rim enhancement in AP and PVP images may frequently obscure the tumor margin and thus increase inter-observer and intra-observer variability (Fig. 3). However, even in tumors without peripheral rim enhancement, inter-observer agreement and intra-observer agreement were better in pre-contrast images than in AP and PVP images (Fig. 4).

## DISCUSSION

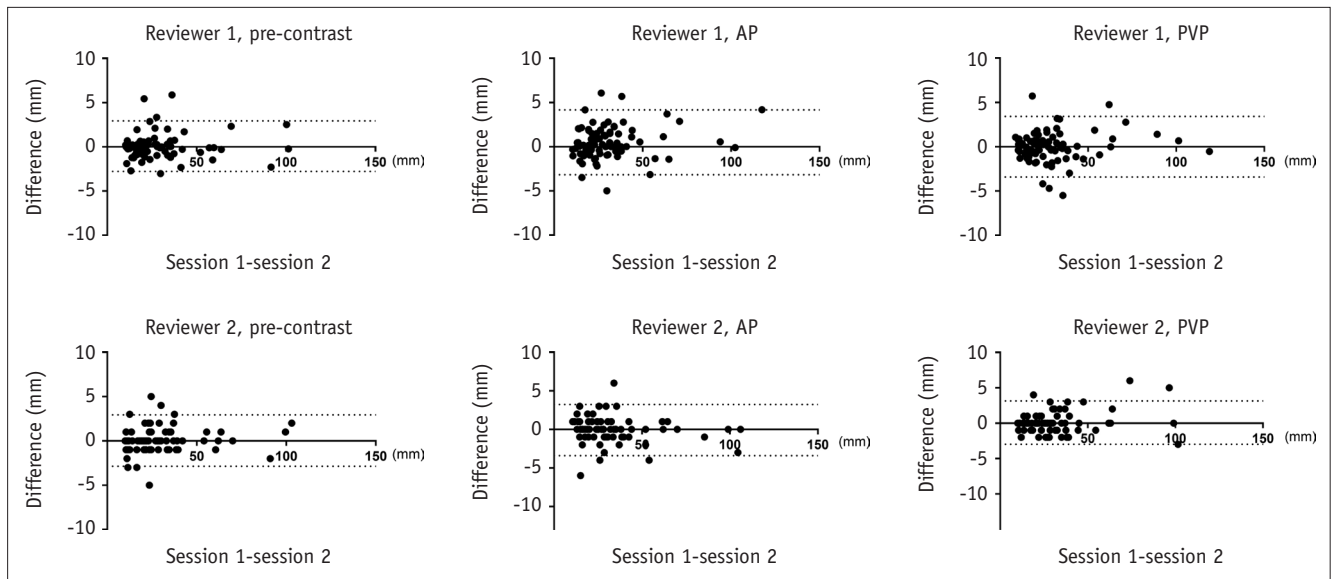
We began our study with the hypothesis that pre-contrast CT might facilitate reliable tumor measurement. Our results demonstrate that the measured tumor size varied with the CT phase and that pre-contrast CT yielded the best inter-observer and intra-observer agreements, thus supporting our hypothesis.

CT is primarily used to detect, characterize, and measure hepatic metastases of NETs, as reported in a recent study (4). When performing tumor response assessments with RECIST 1.1, consistent tumor size measurements may be the most important issue. Our study results indicate that pre-contrast CT may be the most reliable phase for tumor size measurement in lesions with good conspicuity on pre-contrast images.

It is important to note the advantages and drawbacks of pre-contrast imaging of NET liver metastases. We postulated that one advantage of pre-contrast imaging to AP or PVP imaging would be the ability to clearly visualize the



**A**



**B**

**Fig. 1. Bland-Altman plots for assessing inter-observer (A) and intra-observer agreement (B).** AP = arterial phase, PVP = portal venous phase

boundary between a liver metastasis and the adjacent liver parenchyma, wherein the tumor periphery is sometimes obscured by rim enhancement. The readers experienced difficulty drawing clear margins on all tumors with peripheral rim enhancement. Accordingly, our subgroup analysis revealed poorer inter-observer and intra-observer agreement in tumors with peripheral rim enhancement when compared to those without this feature, especially on the AP.

Only one prior study has investigated observer agreement in patients with NET liver metastases (4). The authors of that study explored the effects of various magnetic

resonance (MR) sequences on the size measurements of NET liver metastases. In that study, contrast-enhanced MR images in the AP obtained 60 and 120 seconds after the injection of Gd-ethoxybenzyl diethylenetriamine pentaacetic acid (Primovist or Eovist, Bayer AG, Berlin, Germany) frequently exhibited blurred tumor margins that hindered accurate tumor size measurements and reduced inter-observer/intra-observer agreement. In contrast, the 20-minute delayed phase (i.e., hepatobiliary phase) displayed clear tumor margins and was the best sequence for reliable tumor size measurement (4). These prior study



**Table 4. Subgroup Analysis according to Peripheral Rim Enhancement**

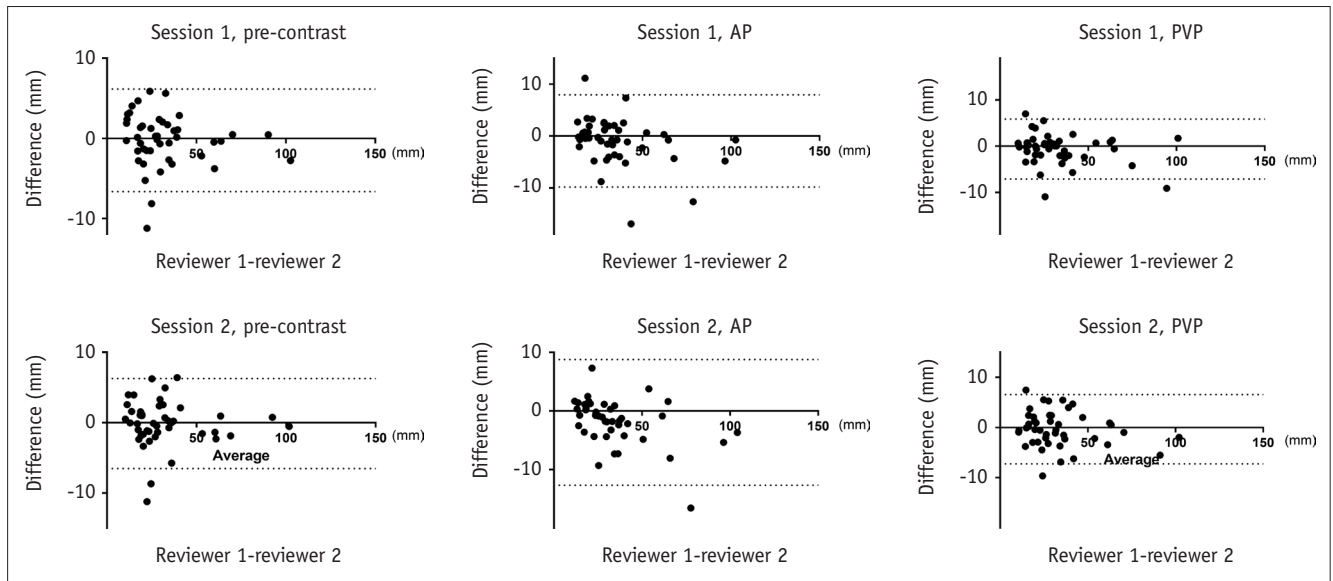
	Pre-Contrast Phase (mm)	AP (mm)	PVP (mm)
Inter-observer agreement (session 1)			
Rim (+)			
Mean difference	-0.24	-0.9	-0.6
95% LOA	-6.6–6.2	-9.9–8.0	-7.1–5.8
RC	6.3	9.0	6.5
Rim (-)			
Mean difference	-0.1	-0.0	-0.0
95% LOA	-5.3–5.2	-6.5–6.5	-8.8–8.8
RC	5.2	6.3	8.6
Inter-observer agreement (session 2)			
Rim (+)			
Mean difference	-0.1	-2.0	-0.4
95% LOA	-6.5–6.3	-12.7–8.8	-7.3–6.5
RC	6.3	11.2	6.9
Rim (-)			
Mean difference	-0.3	-0.1	-0.2
95% LOA	-5.6–4.9	-5.6–5.4	-9.5–9.1
RC	5.2	5.4	9.2
Intra-observer agreement (reader 1)			
Rim (+)			
Mean difference	0.1	0.9	0.1
95% LOA	-3.0–3.1	-2.9–4.6	-3.7–3.8
RC	3.0	4.1	3.7
Rim (-)			
Mean difference	0.1	0.0	-0.1
95% LOA	-2.5–2.8	-3.4–3.5	-3.0–2.9
RC	2.6	3.4	2.9
Intra-observer agreement (reader 2)			
Rim (+)			
Mean difference	0.2	-0.1	0.4
95% LOA	-3.3–3.6	-4.0–3.7	-3.1–3.8
RC	3.4	3.8	3.4
Rim (-)			
Mean difference	-0.1	-0.0	-0.3
95% LOA	-2.1–1.8	-2.5–2.4	-2.6–2.1
RC	1.9	2.4	2.4

results obtained using MR imaging are concordant with those of our study, which were obtained using CT.

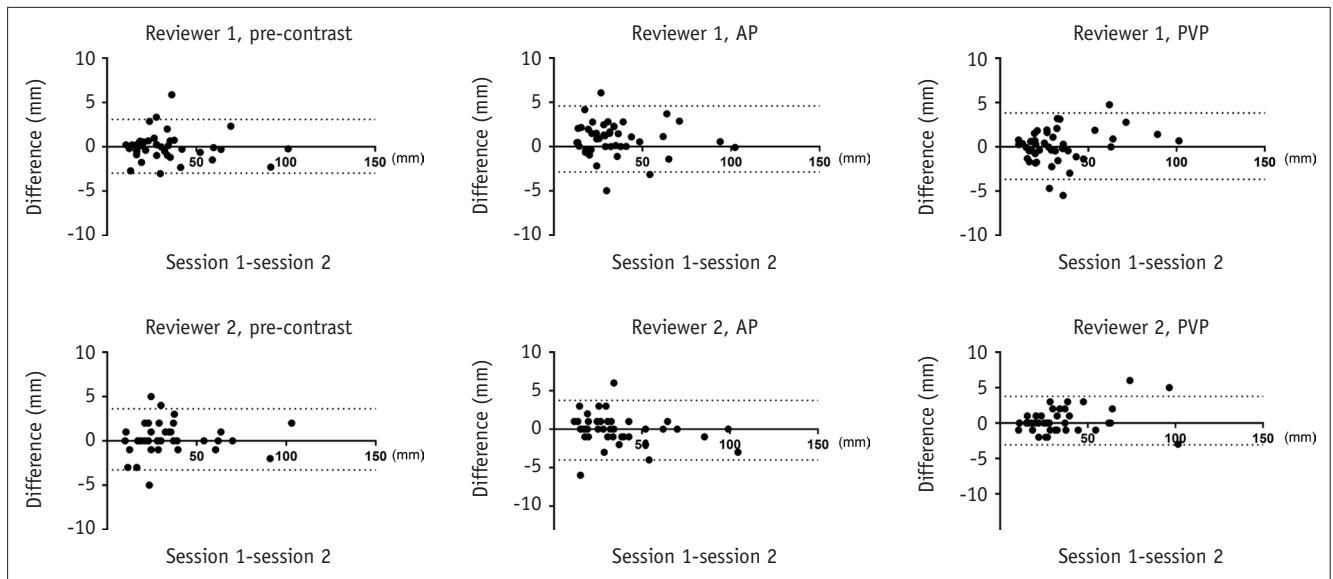
One disadvantage of using pre-contrast CT for tumor size measurement is that the liver metastasis detection rate is generally lower when using pre-contrast CT than when using contrast-enhanced CT (5, 16). In our study, the readers identified target lesions on pre-contrast CT with the aid of AP and PVP CT images. Six of the 79 target lesions (7.6%), which had a mean tumor size of 28.3 mm, were subtle on pre-contrast CT. In such cases, readers should adjust the window level and width setting to identify the

tumor margin. Without referring to contrast-enhanced CT images, however, the diagnostic confidence to distinguish the tumor on pre-contrast CT alone might be low. Therefore, pre-contrast CT alone cannot be used for tumor size measurements or treatment response assessments.

Currently, the RECIST 1.1 criteria emphasize the importance of intravenous contrast injection and define contrast injection as a mandatory process (2, 17). RECIST 1.1 recommends using multiphasic (at least triphasic) CT scans, especially for the assessment of NETs. We generally agree with these recommendations. However, it should be



**A**



**B**

**Fig. 2. Bland-Altman plots of subgroup analyses.**

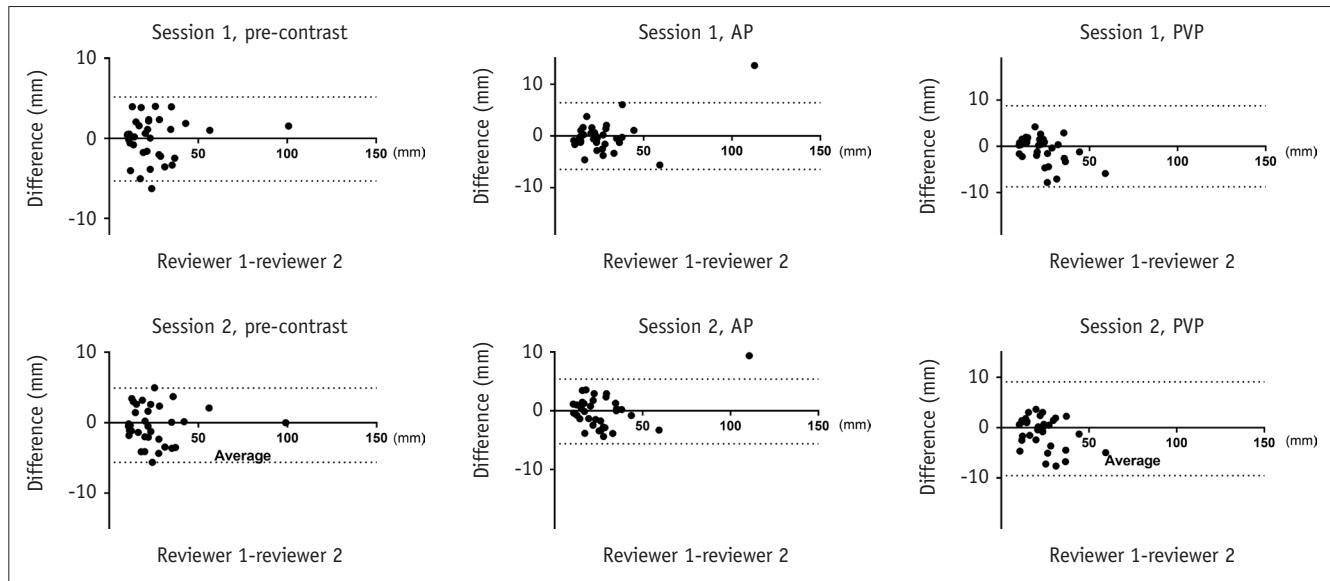
**A.** Inter-observer agreement for tumors with peripheral enhancing rims. **B.** Intra-observer agreement for tumors with peripheral enhancing rims.

noted that the enhancement patterns and morphologies of hypervascular liver metastases from NETs may vary based on the contrast injection protocol and scan timing (4, 6, 8, 18). The original description of RECIST 1.1 did not describe the manner in which peripheral enhancing rims should be measured. Although review papers generally suggest that peripheral enhancing rims should be included in the tumor size measurements (8, 17), our study results indicate that peripheral rim enhancement is a source of inter-observer and intra-observer variability in patients with NET liver metastases. In our study, the average tumor size

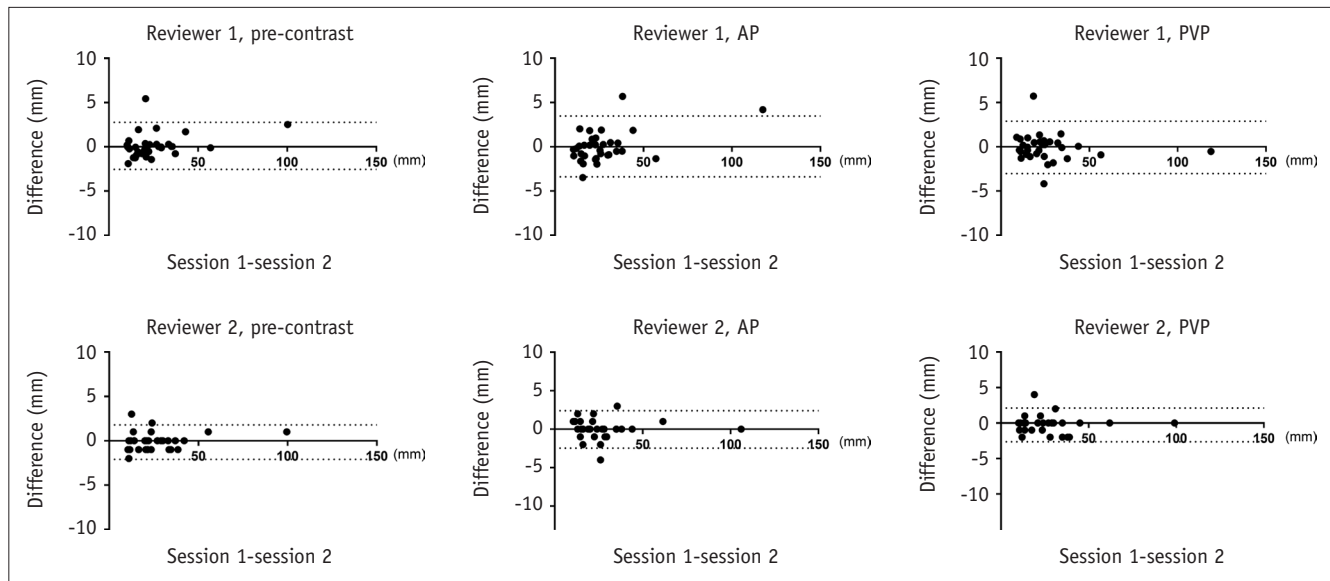
was approximately 3 cm. In this setting, if there is only one target lesion, an increase in size of only 6 mm would lead to exceeding the 20% cut-off for tumor progression. Considering the RCs of inter-observer agreement (5.9 mm for pre-contrast CT, 8.0–9.2 mm for AP, and 7.5–7.9 mm for PVP), the measurement variability derived from the peripheral enhancing rim may change the RECIST response criteria.

On the basis of our observations, we propose that pre-contrast CT might be a good option to decrease variability of tumor size measurement. However, it might be helpful





**C**



**D**

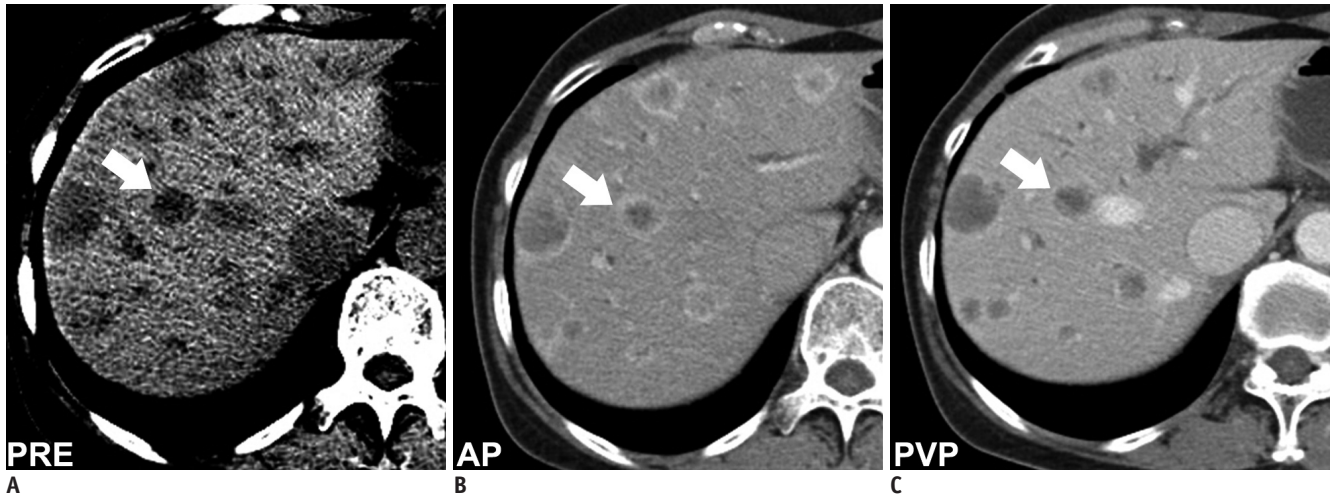
**Fig. 2. Bland-Altman plots of subgroup analyses.**

**C.** Inter-observer agreement for tumors without peripheral enhancing rims. **D.** Intra-observer agreement for tumors without peripheral enhancing rims.

only when the lesion has good conspicuity on pre-contrast images. Further validation to incorporate tumor size measurement on pre-contrast images into the RECIST 1.1 guidelines is required.

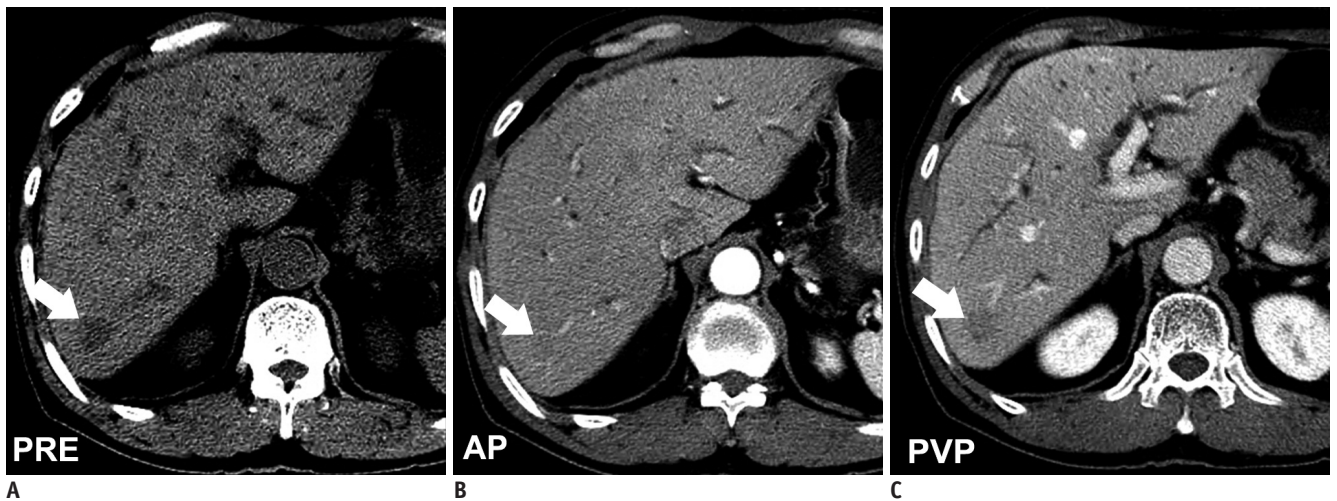
Our study had limitations. One of the limitations of our study was its retrospective nature. We controlled the target lesions used for tumor size measurements in our study, which may have led to overestimation of inter-observer/intra-observer agreement. However, as our study aimed to evaluate the influence of peripheral rim enhancement on tumor size measurement, we were required to predefine the

target lesions. In this study, the target lesions were selected by a study coordinator according to RECIST 1.1 criteria without any preference to peripheral rim enhancement. In real practice, observer variability might derive from differences in selected measurable target lesions or image slices, as well as differences in caliper placement on the target lesion. Second, two readers measured the tumor size twice 4 weeks apart without randomization of case order. This might have introduced bias due to the readers' memory. Third, we used various CT scanners with different CT parameters. This might have influenced lesion conspicuity.



**Fig. 3.** 45-year-old male patient with pancreatic NET and liver metastases.

Metastatic nodule in hepatic segment VIII (arrows) appears as low-attenuated nodule with clear margin on pre-contrast image (A), nodule with thick enhancing rim in AP (B), and nodule with thin rim in PVP (C). Inter-observer and intra-observer variabilities were lowest with pre-contrast CT. CT = computed tomography, NET = neuroendocrine tumor



**Fig. 4.** 50-year-old male patient with pancreatic NET and liver metastases.

Metastatic nodule in hepatic segment VI (arrows) appears as low-attenuated nodule with clear margin on pre-contrast image (A) and indistinct nodule without enhancing rim in AP (B) and PVP (C). Inter-observer and intra-observer variabilities were lowest with pre-contrast CT.

Fourth, this study was conducted in a retrospective manner in a relatively small number of patients. A large-scale prospective validation study is thus warranted.

In conclusion, inter-observer agreement and intra-observer agreement for size measurements of NET liver metastases on CT were better on pre-contrast scans than on AP or PVP scans. Pre-contrast CT may be the optimal phase for measuring NET liver metastases with good conspicuity if the accuracy of this approach is proven.

## REFERENCES

1. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst* 2000;92:205-216
2. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009;45:228-247
3. Thiesse P, Ollivier L, Di Stefano-Louineau D, Négrier S, Savary J, Pignard K, et al. Response rate accuracy in oncology trials: reasons for interobserver variability. Groupe Français d'Immunothérapie of the Fédération Nationale des Centres de Lutte Contre le Cancer. *J Clin Oncol* 1997;15:3507-3514
4. Luersen GF, Wei W, Tamm EP, Bhosale PR, Szklaruk J.

- Evaluation of magnetic resonance (MR) biomarkers for assessment of response with response evaluation criteria in solid tumors: comparison of the measurements of neuroendocrine tumor liver metastases (NETLM) with various MR sequences and at multiple phases of contrast administration. *J Comput Assist Tomogr* 2016;40:717-722
5. Patten RM, Byun JY, Freeny PC. CT of hypervascular hepatic tumors: are unenhanced scans necessary for diagnosis? *AJR Am J Roentgenol* 1993;161:979-984
  6. Soyer P, Gueye C, Somveille E, Laissy JP, Scherrer A. MR diagnosis of hepatic metastases from neuroendocrine tumors versus hemangiomas: relative merits of dynamic gadolinium chelate-enhanced gradient-recalled echo and unenhanced spin-echo images. *AJR Am J Roentgenol* 1995;165:1407-1413
  7. Soyer P, Tidjani K, Laissy JP, Sibert A, Menu Y. Dynamic Gd-DOTA-enhanced MR imaging of hepatic metastases from pancreatic neuroendocrine tumors. *Eur J Radiol* 1994;18:180-184
  8. Chalian H, Töre HG, Horowitz JM, Salem R, Miller FH, Yaghmai V. Radiologic assessment of response to therapy: comparison of RECIST Versions 1.1 and 1.0. *Radiographics* 2011;31:2093-2105
  9. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; Iniciativa STROBE. [The Strengthening the Reporting of Observational Studies in Epidemiology [STROBE] statement: guidelines for reporting observational studies]. *Gac Sanit* 2008;22:144-150
  10. Sprawls P. AAPM tutorial. CT image detail and noise. *Radiographics* 1992;12:1041-1046
  11. Ha HI, Hong SS, Kim MJ, Lee K. 100 kVp low-tube voltage abdominal CT in adults: radiation dose reduction and image quality comparison of 120 kVp abdominal CT. *J Korean Soc Radiol* 2016;75:285-295
  12. Krajewski KM, Nishino M, Franchetti Y, Ramaiya NH, Van den Abbeele AD, Choueiri TK. Intraobserver and interobserver variability in computed tomography size and attenuation measurements in patients with renal cell carcinoma receiving antiangiogenic therapy: implications for alternative response criteria. *Cancer* 2014;120:711-721
  13. Raunig DL, McShane LM, Pennello G, Gatsonis C, Carson PL, Voyvodic JT, et al.; QIBA Technical Performance Working Group. Quantitative imaging biomarkers: a review of statistical methods for technical performance assessment. *Stat Methods Med Res* 2015;24:27-67
  14. Barnhart HX, Barboriak DP. Applications of the repeatability of quantitative imaging biomarkers: a review of statistical analysis of repeat data sets. *Transl Oncol* 2009;2:231-235
  15. Park JE, Han K, Sung YS, Chung MS, Koo HJ, Yoon HM, et al. Selection and reporting of statistical methods to assess reliability of a diagnostic test: conformity to recommended methods in a peer-reviewed journal. *Korean J Radiol* 2017;18:888-897
  16. Jee HB, Park MJ, Lee HS, Park MS, Kim MJ, Chung YE. Is non-contrast CT adequate for the evaluation of hepatic metastasis in patients who cannot receive iodinated contrast media? *PLoS One* 2015;10:e0134133
  17. van Persijn van Meerten EL, Gelderblom H, Bloem JL. RECIST revised: implications for the radiologist. A review article on the modified RECIST guideline. *Eur Radiol* 2010;20:1456-1467
  18. Sahu S, Scherthaner R, Ardon R, Chapiro J, Zhao Y, Sohn JH, et al. Imaging biomarkers of tumor response in neuroendocrine liver metastases treated with transarterial chemoembolization: can enhancing tumor burden of the whole liver help predict patient survival? *Radiology* 2017;283:883-894