

Article

Assessment of Liver Metastases Using CT and MRI Scans in Patients with Pancreatic Ductal Adenocarcinoma: Effects of Observer Experience on Diagnostic Accuracy

Masakatsu Tsurusaki¹, Isao Numoto¹, Teruyoshi Oda¹, Miyuki Wakana¹, Ayako Suzuki¹, Yukinobu Yagyu¹, Mitsuru Matsuki¹ and Kazunari Ishii¹

¹ Department of Radiology, Kindai University Faculty of Medicine, Osaka-Sayama, 589-8511, Japan; mtsuru@dk2.so-net.ne.jp (M.T.); numoto.isao@hotmail.co.jp (I.N.); teru.oda0321@gmail.com (T.O.); miyuki0114ball@yahoo.co.jp (M.W.); i_am_sue_19861211_pinky@yahoo.co.jp (A.S.); yagyu@med.kindai.ac.jp (Y.Y.); rad053@poh.osaka-med.ac.jp (M.M.); ishii@med.kindai.ac.jp (K.I.)

* Correspondence: mtsuru@dk2.so-net.ne.jp; Tel.: +81-72-366-0221 (ext. 3133); Fax: +81-72-367-1685

Abstract: To investigate the impact of radiologic experience on the diagnostic accuracy of computed tomography CT vs. magnetic resonance imaging (MRI) reporting for liver metastases of pancreatic ductal adenocarcinoma (LM of PDAC). Intra-individual CT and MRI examinations of 112 patients with clinically proven LM of PDAC were included. Four radiologists with varying years of experience (A > 20, B > 5, C > 1 and D < 1) assessed liver segments affected by LM of PDAC, as well as associated metastases occurring in each patient. Their sensitivity and specificity in evaluating the segments were compared. Cohen's Kappa (κ) for diagnosed liver segments and Intra-class Correlation Coefficients (ICC) for the number of metastatic lesions in each patient were calculated. The radiologists' sensitivity and specificity for the CT vs. MRI were, respectively: Reader A -94.4, 90.3% vs. 96.6, 94.8%; B - 86.7, 79.7% vs. 83.9, 82.0%; C - 78.0, 76.7% vs. 83.3, 78.9% and D - 71.8, 79.2% vs. 64.0, 69.5%. Reviewers A and B achieved greater agreement in assessing results from the MRI (κ = 0.72, p < 0.001; ICC = 0.73, p < 0.001) vs. the CT (κ = 0.58, p < 0.001; ICC = 0.61, p < 0.001), in contrast to readers C and D (MRI: κ = 0.34, p < 0.001; ICC = 0.42, p < 0.001, and CT: κ = 0.48, p < 0.001; ICC = 0.59, p < 0.001). Our results indicate that accurate diagnosis of LM of PDAC depends more on radiologic experience in MRI over CT scans.

Keywords: magnetic resonance imaging; multidetector computed tomography; liver; neoplasm metastasis; gadoxetic acid

1. Introduction

International guidelines advise contrast-enhanced computed tomography (CE-CT) for routine diagnosing and staging of pancreatic cancer, whereas magnetic resonance imaging (MRI) is mostly used for the characterization of indeterminate liver lesions [1,2]. Accurate assessment of liver metastases (LM), both colorectal (CRLM) and non-colorectal LM, is crucial in multidisciplinary oncology [1–4]—especially for patients with pancreatic ductal adenocarcinoma (PDAC). Enhanced detection of LM could reduce futile resection of tumors and markedly increase life expectancy.

At present, CE-CT has been widely used as a standard imaging modality to determine the stage of pancreatic cancer. However, its ability to detect LM less than 1 cm in size is reported to be limited and unsatisfactory, given its rate of accuracy currently stands at just 50% [3,4]. Fortunately, liver-specific magnetic resonance contrast agents like gadoxetate disodium appear to offer great promise because of their ability to provide more precise evaluations of tumor infiltrations [5,6], and are now recommended for the diagnosis and characterization of malignant lesions in non-cirrhotic livers [7].

Moreover, prior studies have demonstrated that multidisciplinary team meetings associate with significant improvements in clinical outcomes, as imaging data is reviewed by all physicians involved in the patient's care, regardless of their radiologic experience [8–11]. It is possible that the widespread availability of medical imaging data may eventually lead to independent read-outs of CT and MRI examinations without consultation of the corresponding radiological report. Yet, since marking images is not a standardized practice amongst radiologists, but depends more on their reporting preferences and level of experience [12], the comprehensibility of oncologic findings for inexperienced readers remains unknown [2].

In this study, we aimed to determine the impact of observer experience in CT and MRI examinations on the diagnostic accuracy of (LM of PDAC).

2. Results

2.1. Diagnostic performance

Sensitivity, specificity, PPV and NPV were calculated on a segmental basis. Overall, a trend was found in that the diagnostic performance was proportional to experience for both reporting of CT and MRI examinations in Table 2. In particular, reviewer A reached the highest sensitivity and specificity for both CT (94.4 and 90.3%, respectively) and MRI (96.6 and 94.8%, respectively), compared with reviewer B (CT: 86.7 and 79.7%, MRI: 83.9 and 82.0%, respectively), C (CT: 78.0 and 76.7%, MRI: 83.3 and 78.9%, respectively), and D (CT: 71.8 and 79.2%, MRI: 64.0 and 69.5%, respectively). Furthermore, Reader A achieved the greatest PPV and NPV for CT (88.6 and 95.7%, respectively) and MRI (96.2 and 95.2%, respectively), compared with observer B (CT: 73.3 and 90.7%, MRI: 83.2 and 82.7%, respectively), C (CT: 73.1 and 81.2%, MRI: 79.5 and 82.6%, respectively), and D (CT: 77.8 and 73.3%, MRI: 67.3 and 66.6%, respectively). Comprehensive data regarding diagnostic performance is provided in Table 2. Differences between reviewers A and B, as well as B and D, were significant for CT reporting ($p = 0.001$ for both). Regarding MRI, the following comparisons reached the level of significance, respectively: readers A and C ($p = 0.013$), A and D ($p < 0.001$), B and D ($p = 0.001$) and C and D ($p = 0.014$). Additional data is summarized in Table 3.

Table 1. Mean and standard deviation (SD) of segment-based sensitivity, specificity, PPV and NPV for CT and MRI assessments

	CT				MRI			
	Mean	SD	Mean	SD				
Sensitivity	82.76	8.55	82.15	11.65				
Specificity	81.51	5.21	81.34	9.06				
PPV	78.24	6.30	81.54	10.41				
NPV	85.27	8.62	81.84	10.15				

Table 2. CT and MRI sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) on a segmental base achieved by reviewers A (> 20 y of experience), B (> 10 y), C (> 1 y), and D (< 1 y). Diagnostic performance was mainly proportional to radiological experience.

	CT								MRI								
	Reader A	95%-CI	Reader B	95%-CI	Reader C	95%-CI	Reader D	95%-CI	Reader A	95%-CI	Reader B	95%-CI	Reader C	95%-CI	Reader D	95%-CI	
Sensitivity	94.44	91.7-97.78	86.77	81.9-92.28	81.9-84.34	78.04-84.34	72.9-84.34	71.80-78.58	66.5-78.58	96.63-97.79	92.2-97.79	83.99-87.55	77.5-87.55	83.39-87.10	76.7-87.10	64.02%-69.42	56.2-69.42
Specificity	90.36	86.0-93.73	79.77	74.3-84.50	76.71-81.81	70.9-81.81	79.25-84.50	73.1-84.50	94.85-97.31	91.1-97.31	82.05-86.75	76.5-86.75	78.93-83.89	73.2-83.89	69.53%-75.37	63.1-75.37	
PPV	88.65	85.0-93.26	73.37	67.8-80.24	73.10-79.71	67.8-79.71	77.84-84.19	72.6-84.19	96.25-97.52	91.83-97.52	83.29-86.92	76.81-86.92	79.58-83.61	72.8-83.61	67.04%-72.42	59.1-72.42	
NPV	95.74	92.3-97.94	90.71	86.15-94.16	81.28-86.05	75.69-86.05	73.36-78.97	67.1-78.97	95.26-97.61	91.67-97.61	82.76-87.39	77.2-87.39	82.68-87.33	77.1-87.33	66.67%-72.56	60.36-72.56	

CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value.

Table 3. Statistical comparisons of diagnostic performance were performed using the McNemar test

Reviewers	CT (p-value)	MRI (p-value)
A+B	0.0015	0.0763
C+D	0.0768	0.0138

Significant values (p < 0.05) are given in bold italics.

2.2. Inter-observer agreement

Among all the reviewers, inter-reader agreement for the liver segments that were affected by malignancy was higher for MRI ($\kappa = 0.44$, $p < 0.001$) than for CT ($\kappa = 0.43$, $p < 0.001$). In particular, the experienced reviewers, A and B, achieved greater agreement for MRI ($\kappa = 0.72$, $p < 0.001$) than for CT ($\kappa = 0.58$, $p < 0.001$), unlike the less experienced C and D reviewers (MRI: $\kappa = 0.34$, $p < 0.001$ and CT: $\kappa = 0.48$, $p < 0.001$, respectively). A similar trend was found for the inter-observer agreement regarding the number of LM present in each patient. All reviewers, together, achieved greater inter-observer agreement for MRI ($ICC = 0.59$, $p < 0.001$) than for CT ($ICC = 0.53$, $p < 0.001$). The experienced reviewers, A and B, showed higher agreement for MRI ($ICC = 0.73$, $p < 0.001$) than for CT ($ICC = 0.61$, $p < 0.001$), as opposed to the less experienced reviewers, C and D (MRI: $ICC = 0.41$, $p < 0.001$ and CT: $ICC = 0.59$, $p < 0.001$). Complementary data is shown in Table 4.

Table 4. Inter-observer agreement regarding liver segments that were affected by LM of PDAC (κ) and the number of lesions reported per patient (ICC) for CT and MRI analyses. The experienced reviewers showed a significantly higher agreement for MRI than for CT reporting.

Readers	K		ICC	
	CT	MRI	CT	MRI
A+B+C+D	0.43	0.44	0.53	0.59
A+B	0.58	0.72	0.61	0.73
C+D	0.48	0.34	0.59	0.42

3. Discussion

This study aimed to investigate the impact of observer experience on diagnostic performance and inter-observer agreement in reporting LM of PDAC using CT and MRI scans because data from these imaging modalities are commonly reviewed by physicians with varying levels of experience in the clinical and radiological practice of oncology [13].

We found that diagnostic performance was primarily proportional to reviewer experience, with the most experienced reviewer, A, achieving the highest sensitivity, specificity, PPV and NPV [14–16]. Thus, our results indicate that observer experience is essential for exhibiting a high diagnostic accuracy in the aforementioned imaging modalities (Figure 1). More importantly, values indicating diagnostic accuracy differed more distinctly for MRI than for CT, as sensitivity, specificity, PPV and NPV were distributed over a larger range for the MRI analyses [14–16]. Therefore, our data suggest that reviewer experience has a greater impact on MRI than on CT reporting. This hypothesis is further supported by the fact that the more experienced radiologists, reviewers A and B, showed a greater difference in diagnostic performance indices for MRI. This suggests that there is a greater learning curve for MRI interpretation after 5 years of radiologic experience compared with CT reporting. Accordingly, we detected a greater inter-observer agreement among the experienced reviewers for the MRI analyses, both for the amount of LM of PDAC detected per patient and the affected liver segments, as opposed to the less experienced reviewers, who achieved a higher inter-observer agreement using CT images. This may indicate that experienced reviewers are more likely to be consistent in their findings when evaluating MRI scans, unlike less experienced reviewers.

Potential factors that may contribute to the need for more experience in MRI analyses over CT interpretations are the greater number of images that need to be assessed, as well as its challenging physical theory. For instance, chemical shift and diffusion-weighted imaging in MRI may be more difficult to assess for less experienced reviewers than CT series that are based on density values and primarily defined by the time of image acquisition relative to the contrast media administration.

Contrastingly, standard CT scans for liver imaging in patients with LM of PDAC include a reduced amount of series and often consist solely of a non-contrast and portal-venous phase image acquisition [16]. Thus, our data advocate that reporting of CT images may be more comprehensible and intuitive for less experienced reviewers. Our results suggest that marking and explaining findings in MRI reports will make them easier to understand for less experienced readers and clinicians, who evaluate imaging data in the absence of an experienced radiologist. Additionally, these results further justify the educational practice in radiological departments wherein residents undergo CT before MRI training.

Evidence regarding the influence of observer experience in interpreting oncological CT and MRI examinations is scarce, despite the fact that both clinicians and radiologists with varying levels of experience routinely review such studies and thus may affect therapeutic regimens and patient care. Although few investigations have examined the impact of reader experience on the reproducibility of tumor measurements [12,17], a prior study examined the influence of observer expertise on CT and MRI reporting in patients with CRLM [11]. The authors concluded that MRI analysis of CRLM is more affected by observer experience than CT interpretation [11]. Our results are consistent with this hypothesis and likewise indicate that reviewer experience is a crucial determinant in diagnostic accuracy. More importantly, they also support the assumption that this effect is more distinct within MRI over CT reporting. Therefore, we conclude that experience may have a greater impact on diagnostic accuracy for MRI reporting of LM of varying primary carcinomas compared with CT analyses. In addition to this prior investigation, we found consistent results using CT and MRI examinations on the same individuals, which included a diffusion-weighted MRI series.

This study has limitations that should be mentioned. First, we only included four radiologists with varying levels of experience. Because the influence of experience on diagnostic accuracy is challenging to investigate, results may be more representative following a larger population of reviewers in future studies. Additionally, subtle changes of LM of PDAC that were not indicated by criteria in the image analyses may limit comparability of the intra-individual CT and MRI examinations.

4. Materials and Methods

4.1. Patients

This retrospective study was approved by our local institutional review board (Ethical Committee of Kindai University No. 23-101), and written informed consent was obtained from each patient before undergoing a CT or MRI scan. This study used data from clinical records and images collected from Kindai University Hospital, a high-volume regional referral center. Between January 2009 and December 2017, patients with histologically-confirmed pancreatic cancer who underwent a CE-CT and gadoxetic acid-enhanced MRI were enrolled in the study. CT and MRI examinations of 164 patients with LM of PDAC were retrospectively included. From the initial study group, 52 patients were excluded because their time intervals between the CT and MRI examination were greater than one month, so 112 individuals in total participated (62 males, mean age \pm standard deviation: 62 ± 12.4 y). There were 172 of LM of PDAC and 157 liver segments that were affected by LM of PDAC.

4.2. CT imaging protocol

Intravenous CE-CT imaging was performed using a 64-channel multidetector row scanner (Light Speed VCT Vision, GE Healthcare, Waukesha, WI) with a tube voltage of 100-170 kV, an automatic dose modulation, a pitch of 2.0 and a slice thickness of 1 mm. Axial, coronal and sagittal slices were reconstructed with a section thickness of 5.0 mm and an increment of 3 mm. After the unenhanced images were acquired, 510 mg/kg of iodinated contrast material (Optiray 320, Guerbet Japan, Tokyo, Japan) was administered intravenously into the antecubital vein at a rate of 3–4 mL/s. Scanning was performed at the beginning of the pancreatic parenchymal phase (after 40 s) and the subsequent liver phase was obtained 70 s after intravenous administration of the contrast material.

4.3. MR imaging protocol

Enhanced magnetic resonance imaging (EOB-MRI) was performed with two superconducting magnet systems on a 3.0-T scanner (Magnetom Trio, Siemens Medical Systems, Erlangen, Germany; n = 48; Achieva TX, Philips Healthcare, Best, The Netherlands; n = 64) using a 32-channel phased-array body coil for all patients. The Magnetom Trio scanner was actively shielded with a 45 mT/m gradient field strength and slew rate of 200 T/m/s. The Achieva TX scanner was actively shielded with a 50 mT/m gradient field strength and slew rate of 220 T/m/s. For both scanners, after breath-hold double-echo T1-weighted gradient recalled echo (GRE) images (in-phase and opposed-phase images) and navigator-triggered fat-suppressed T2-weighted turbo spin-echo (TSE) images were obtained, dynamic fat-suppressed T1-weighted images were obtained with a three-dimensional (3D) GRE sequence before (pre-contrast), 14 – 30 s after (arterial phase by means of a bolus-triggered technique), 70 s after and 3 min after intravenous administration of gadoxetic acid (EOB Primovist; Bayer Yakuhin, Osaka, Japan), which was injected as a bolus (2.0 mL/s) at a dose of 0.025 mmol/kg of body weight, followed by a 20 mL saline flush. Hepatocyte-phase images were obtained 20 min after the gadoxetic acid injection.

The Siemens scanner acquired breath-hold T1-weighted GRE images (repetition time/echo time [TR/TE] of 120/2.46 ms, flip angle of 66°, matrix size of 320 × 180, bandwidth of 434.3 Hz/pixel, one signal acquired, section thickness of 7 mm, intersectional gap of 1.4 mm and acquisition time of 28 s), navigator-triggered fat-suppressed T2-weighted TSE images using PACE (TR/TE of 3865–5534/71 ms, flip angle of 120°, matrix size of 384 × 202, bandwidth of 506.4 Hz/pixel, one signal acquired, section thickness of 7 mm, intersectional gap of 1.4 mm and acquisition time of approximately 90 s). For the EOB-MRI, fat-suppressed T1-weighted GRE images were acquired with a 3D-VIBE sequence (TR/TE of 3.68/1.22 ms, flip angle of 10°, matrix size of 256 × 192, one signal acquired, section thickness of 3 mm, intersectional gap of 0.6 mm and acquisition time of 21 s).

The Philips scanner included a respiration-triggered T1-weighted turbo field-echo in-phase sequence (TR/TE of 10/2.3 ms, flip angle of 15°, matrix size of 288 × 230, bandwidth of 434.3 Hz/pixel, one signal acquired, section thickness of 7 mm, intersectional gap of 1.2 mm and acquisition time of 26 s) and a respiration-triggered fat-suppressed T2-weighted sequence with a reduction factor of 2 or 4 (TR/TE of 1342/80 ms, flip angle of 90°, matrix size of 320 × 256, bandwidth of 506.4 Hz/pixel, one signal acquired, section thickness of 7 mm, intersectional gap of 1.2 mm and acquisition time of approximately 100 s). For the EOB-MRI, a fat-suppressed T1-weighted 3D turbo field-echo sequence was performed (T1 high-resolution isotropic volume examination [THRIVE], Philips Healthcare) (TR/TE of 3.4/1.8 ms, flip angle of 10°, matrix size of 336 × 206, bandwidth of 995.7 Hz/pixel) with a 3 mm section thickness and a field-of-view of 32–38 cm.

4.4. Image analysis

All CT and MRI series were reviewed by four reviewers with varying levels of experience in oncologic radiology (Reviewer A > 20 y, B > 15 y, C > 1 y, D < 1 y) according to Albrecht's methodology [20]. They were aware that this study focused on the detection of LM, though were not told that LM of PDAC was present in each patient and were blinded to patients' ages, primary cancers, clinical course and previous treatments. Images were assessed in a randomized order, with a mandatory time interval between intra-individual read-outs of CT and MRI datasets of one week, so as to reduce potential recall biases. Preset window settings could be freely adjusted. Reviewers reported the amount of metastases they detected per patient, as well as the liver segments that were affected by LM of PDAC, according to the Couinaud classification of hepatic anatomy. Segment 4 was evaluated as a cranial (4a) and caudal portion (4b). Thus, each liver segment was either rated as affected (positive) or not affected (negative) regarding malignant infiltration by LM of PDAC.

The criteria for radiological diagnosis of LM on the CE-CT were that they had to be ill-defined, heterogeneous nodules with a higher attenuation than bile and having some degree of enhancement. The criteria for radiological diagnosis of LM on the EOB-MRI were that they had to be focal, discrete nodular lesions that showed a high signal intensity relative to the liver parenchyma on T2-weighted FSE images (with a lower signal intensity than the gallbladder or cerebrospinal fluid), a low signal

intensity relative to the liver parenchyma on T1-weighted GRE images obtained at 70 s and 3 min after the gadoxetic acid injection, and were more conspicuous on the hepatocyte-phase images. The diagnosis of LM was more definite when perilesional enhancement was detected on the T1-weighted GRE images obtained 30 s after the gadoxetic acid injection (Figure 1).

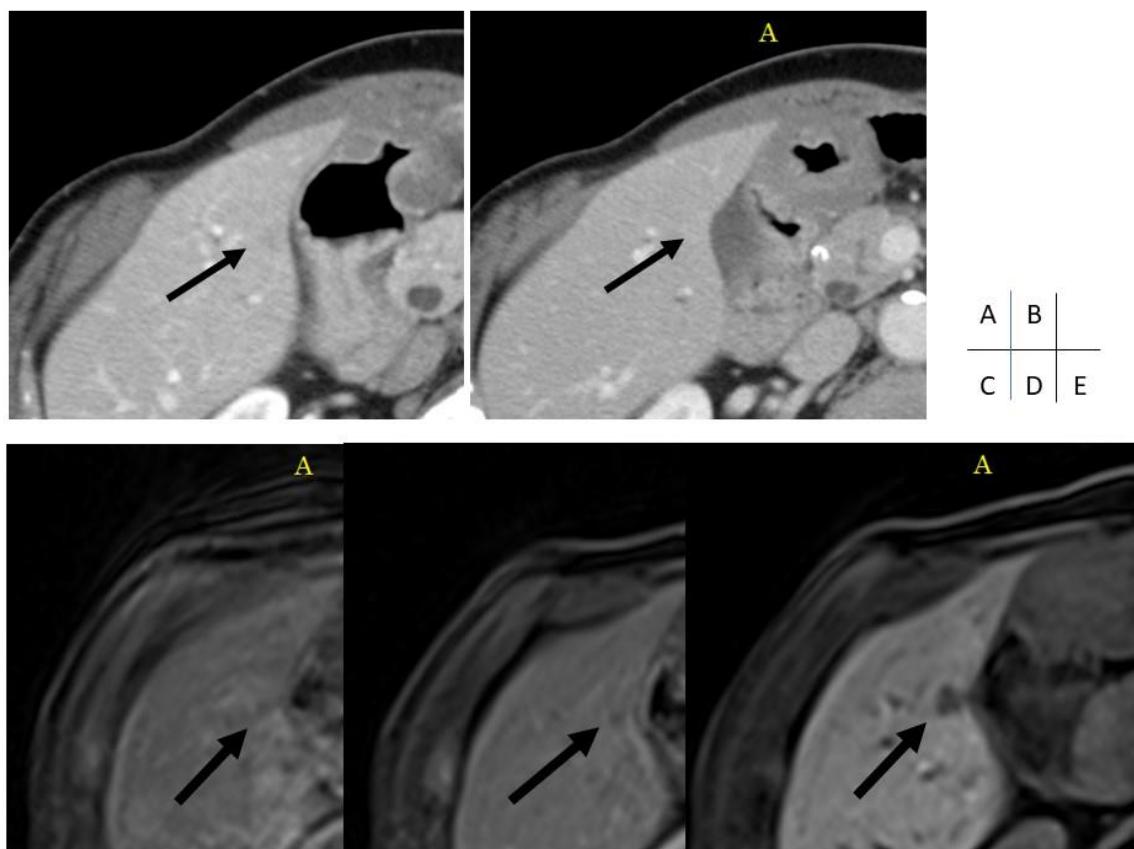


Figure 1. Images of CT and MRI examinations of a 74-year-old female patient with small metastasis of PDAC (A, arterial-phase CT; B, portal-venous phase CT; C, arterial-phase MRI; D, portal-venous phase MRI; E, hepatobiliary phase MRI). A small lesion in segment 5 (arrow) was only detected by the experienced readers A and B using MRI (only depicted by hepatobiliary phase MRI), but the lesion was missed by all readers using CT.

4.5. Standard of reference

In patients who underwent endoscopic ultrasonography fine-needle aspiration (EUS-FNA) for liver tumors, the diagnosis of LM was made based on a combination of histopathological findings of EUS-FNA samples and follow-up imaging examinations. In cases without histopathological diagnosis, the final diagnosis was confirmed by combining all available imaging examinations performed 2 to 3 m after the initial CT and MRI. If possible, observation with these imaging examinations of the lesions was followed up every 2 mos. In cases lacking a histopathological diagnosis, the final diagnosis was made when a significant change in the tumor's size was confirmed using these imaging examinations. Most of tumor (159 tumors) were increased in the interval. A few tumors were confirmed to have shrunk by chemotherapy and were not abscesses by combining all available imaging examinations. In this study, LM were considered to pre-exist at the time of the initial examination, despite no detection by any imaging modalities (metastatic lesions were identified by follow-up imaging within 3 m of the initial examination).

4.6. Statistical analysis

Statistical evaluation was performed using dedicated software (MedCalc, v. 12.7.2 and BiAS, v. 10.11); p-value < 0.05 was considered to indicate a statistically significant difference. Segment-based sensitivity, specificity, positive predictive values (PPV) and negative predictive values (NPV) were compared using the McNemar test. Confidence intervals (CI) of 95% were calculated. For the qualitative lesion- and segment-based analysis, Cohen's kappa (κ) statistics were performed. Values were interpreted as follows: $\kappa < 0.40$, poor agreement; $\kappa = 0.40 - 0.75$, fair to good agreement; $\kappa > 0.75$, excellent agreement [18]. The quantitative analysis for the number of metastases detected per patient was conducted using Intraclass Correlation Coefficients (ICC). ICC values were classified as follows: $ICC < 0.40$, poor agreement; $ICC = 0.40-0.60$, moderate agreement; $ICC = 0.60 - 0.80$, substantial agreement; $ICC > 0.80$, good agreement [19].

5. Conclusions

Our results indicate that accurate diagnosis of LM of PDAC depends more on radiologic experience in MRI over CT scans. Although observer experience is crucial for a high diagnostic accuracy regarding both CT and MRI analysis in patients with LM of PDAC, this effect seems to be more pronounced in MRI reporting.

Author Contributions: Conceptualization, M.T.; methodology, M.T.; formal analysis, M.T.; investigation, Y.Y.; resources, M.T.; data curation, I.N., T.O., M.W., A.S., and Y.Y.; methodology and software, M.T.; writing—original draft preparation, M.T.; writing—review and editing, M.T.; supervision, M.M., and K.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Acknowledgments: The authors would like to thank Enago (www.enago.jp) for the English language editing.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results

References

1. Ducreux, M.; Caramella, C.; Hollebecque, A.; Burtin, P.; Goéré, D.; Seufferlein, T.; Haustermans, K.; Van Laethem, J.L.; Conroy, T.; Arnold, D. Cancer of the pancreas: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* **2015**, *26*, v56–v68.
2. Kouri, B.E.; Abrams, R.A.; Al-Refaie, W.B.; Azad, N.; Farrell, J.; Gaba, R.C.; Gervais, D.A.; Gipson, M.G.; Kolbeck, K.J.; Marshalleck, F.E.; et al. ACR appropriateness criteria radiologic management of hepatic malignancy. *J Am Coll Radiol* **2016**, *13*, 265–273. doi: 10.1016/j.jacr.2015.12.001.
3. Kuszyk, B.S.; Bluemke D.A.; Urban B.A.; Choti M.A.; Hruban R.H.; Sitzmann J.V.; Fishman E.K. Portal-phase contrast-enhanced helical CT for the detection of malignant hepatic tumors: sensitivity based on comparison with intraoperative and pathologic findings. *AJR Am J Roentgenol* **1996**, *166*, 91–95.
4. Niekel, M.C.; Bipat, S.; Stoker, J. Diagnostic imaging of colorectal liver metastases with CT, MR imaging, FDG PET, and/or FDG PET/CT: a meta-analysis of prospective studies including patients who have not previously undergone treatment. *Radiology* **2010**, *257*, 674–684.
5. Motosugi, U.; Ichikawa, T.; Morisaka, H.; Sou, H.; Muhi, A.; Kimura, K.; Sano, K.; Araki, T. Detection of pancreatic carcinoma and liver metastases with gadoxetic acid enhanced MR imaging: comparison with contrast-enhanced multi-detector row CT. *Radiology* **2011**, *260*, 446–453.
6. Ito, T.; Sugiura, T.; Okamura, Y.; Yamamoto, Y.; Ashida, R.; Aramaki, T.; Endo, M.; Uesaka, K. The diagnostic advantage of EOB-MR imaging over CT in the detection of liver metastasis in patients with potentially resectable pancreatic cancer. *Pancreatology* **2017**, *17*, 451–456.
7. Neri, E.; Bali, M.A.; Ba-Ssalamah, A.; Boraschi, P.; Brancatelli, G.; Alves, F.C.; Grazioli, L.; Helmberger, T.; Lee, J.M.; Manfredi, R.; et al. ESGAR consensus statement on liver MR imaging and clinical use of liver-specific contrast agents. *Eur Radiol* **2016**, *26*, 921–931.
8. Adam, R.; Chiche, L.; Aloia, T.; Elias, D.; Salmon, R.; Rivoire, M.; Jaeck, D.; Saric, J.; Le Treut, Y.P.; Belghiti, J.; et al. Hepatic resection for noncolorectal nonendocrine liver metastases: analysis of 1,452 patients and development of a prognostic model. *Ann Surg* **2006**, *244*, 524–535. doi: 10.1097/01.sla.0000239036.46827.5f

9. O'Rourke, T.R.; Tekkis, P.; Yeung, S.; Fawcett, J.; Lynch, S.; Strong, R.; Wall, D.; John, T.G.; Welsh, F.; Rees, M. Long-term results of liver resection for non- colorectal, non-neuroendocrine metastases. *Ann Surg Oncol* **2008**, *15*, 207–218. doi: 10.1245/s10434-007-9649-4.
10. Tan, M.C.B.; Jarnagin W.R. Surgical management of non-colorectal hepatic metastasis. *J Surg Oncol* **2014**, *109*, 8–13. doi: 10.1002/jso.23462.
11. Albrecht, M.H.; Wichmann, J.L.; Müller, C.; Schreckenbach, T.; Sakthibalan, S.; Hammerstingl, R.; Bechstein, W.O.; Zangos, S.; Ackermann, H.; Vogl, T.J. Assessment of colorectal liver metastases using MRI and CT: impact of observer experience on diagnostic performance and inter-observer reproducibility with histopathological correlation. *Eur J Radiol* **2014**, *83*, 1752–1758. doi: 10.1016/j.ejrad.2014.07.005.
12. McErlean, A.; Panicek, D.M.; Zabor, E.C.; Moskowitz, C.S.; Bitar, R.; Motzer, R.J.; Hricak, H.; Ginsberg, M.S. Intra- and interobserver variability in CT measurements in oncology. *Radiology* **2013**, *269*, 451–459. doi: 10.1148/radiol.13122665.
13. Nishino, M.; Jagannathan, J.P.; Ramaiya, N.H.; Van den Abbeele, A.D. Revised RECIST guideline version 1.1: What oncologists want to know and what radiologists need to know. *Am J Roentgenol* **2010**, *195*, 281–289. doi: 10.2214/AJR.09.4110.
14. Holzapfel, K.; Reiser-Erkan, C.; Fingerle, A.A.; Erkan, M.; Eiber, M.J.; Rummeny, E.J.; Friess, H.; Kleeff, J.; Gaa, J. Comparison of diffusion-weighted MR imaging and multidetector-row CT in the detection of liver metastases in patients operated for pancreatic cancer. *Abdom Imaging* **2011**, *36*, 179–184.
15. Vilgrain, V.; Esvan, M.; Ronot, M.; Caumont-Prim, A.; Aubé, C.; Chatellier, G. A meta-analysis of diffusion-weighted and gadoxetic acid-enhanced MR imaging for the detection of liver metastases. *Eur Radiol* **2016**, *26*, 4595–4615.
16. Marion-Audibert, A.M.; Vullierme, M.P.; Ronot, M.; Mabrut, J.Y.; Sauvanet, A.; Zins M.; Cuilleron, M.; Sa-Cunha, A.; Lévy, P.; Rode, A. Routine MRI with DWI sequences to detect liver metastases in patients with potentially resectable pancreatic ductal carcinoma and normal liver CT: A prospective multicenter study. *AJR Am J Roentgenol* **2018**, *211*, W217–W225. doi: 10.2214/AJR.18.19640.
17. Hopper, K.D.; Kasales, C.J.; Van Slyke, M.A.; Schwartz, T.A.; TenHave, T.R.; Jozefiak, J.A. Analysis of interobserver and intraobserver variability in CT tumor measurements. *Am J Roentgenol* **1996**, *167*, 851–854. doi: 10.2214/ajr.167.4.8819370.
18. Cohen, J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* **1960**, *20*, 37–46. doi: 10.1177/001316446002000104.
19. Bartko, J.J. The intraclass correlation coefficient as a measure of reliability. *Psychol Rep* **1966**, *19*, 3–11. doi: 10.2466/pr0.1966.19.1.3.
20. M.H. Albrecht, JL Wichmann, C Müller, T Schreckenbach, S Sakthibalan, R Hammerstingl, WO Bechstein, S Zangos, S Ackermann, TJ Vogl. Assessment of Colorectal Liver Metastases Using MRI and CT: Impact of Observer Experience on Diagnostic Performance and Inter-Observer Reproducibility With Histopathological Correlation. *Eur J Radiol*. 2014;83(10):1752-8. doi: 10.1016/j.ejrad.2014.07.005.