Clinical utility of the Bosniak classification version 2019

Arita, Yuki; Yoshida, Soichiro; Kwee, Thomas C.; Edo, Hiromi; Kufukihara, Ryohei; Shigeta, Keisuke; Nagasaka, Misa; Takeshita, Ryo; Okamura, Haruka; Ueda, Ryo

Published in:
European Journal of Radiology

DOI:
10.1016/j.ejrad.2022.110163

Publication date:
2022

European Journal of Radiology 148 (2022) 110163

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Clinical utility of the Bosniak classification version 2019: Diagnostic value of adding magnetic resonance imaging to computed tomography examination

Yuki Arita [a,*], Soichiro Yoshida [b], Thomas C. Kwee [c], Hiromi Edo [d], Ryohei Kufukihara [e], Keisuke Shigeta [e], Misa Nagasaka [a], Ryo Takeshita [a], Haruka Okamura [a], Ryo Ueda [f], Ryota Ishii [g], Shigeo Okuda [a], Yasuhisa Fujii [b]

[a] Department of Radiology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan
[b] Department of Urology, Tokyo Medical and Dental University Graduate School, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8519, Japan
[c] Department of Radiology, Nuclear Medicine, and Molecular Imaging, University Medical Center Groningen, Hanzeplein 1, PO Box 30.001, 9700 RB Groningen, the Netherlands
[d] Department of Radiology, National Defense Medical College, 3-2, Namiki, Tokorozawa, Saitama 359-8513, Japan
[e] Department of Urology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan
[f] Office of Radiation Technology, Keio University Hospital, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan
[g] Department of Biostatistics, Faculty of Medicine, University of Tsukuba, 1-1-1, Tennodai, Tsukuba, Ibaraki 305-8575, Japan

ARTICLE INFO

Keywords:
2019 Bosniak classification
Cystic renal mass
Magnetic resonance imaging
Renal cell carcinoma
Tomography

ABSTRACT

Purpose: To assess the impact of the updated Bosniak classification (BC2019) for cystic renal masses (CRMs) on interobserver agreement between radiologists and urologists and the diagnostic value of adding MRI to CT examination (combined CT/MRI).

Method: This study included 103 CRMs from 83 consecutive patients assessed using contrast-enhanced CT and MRI between 2010 and 2016. Nine readers in three groups (three radiologists, three radiology residents, and three urologists) reviewed CT alone and the combined CT/MRI using BC2019. Bosniak category was determined by consensus in each group for diagnosing malignancy, with a cut-off category of ≧III. Interobserver agreement was assessed using Fleiss' kappa values. The effect of CT or combined CT/MRI on the diagnosis of malignancy was assessed using McNemar's test.

Results: Interobserver agreement of BC2019 for CT alone was substantial for radiologists and residents, moderate for urologists (0.77, 0.63, and 0.58, respectively). Interobserver agreement of BC2019 for combined CT/MRI was substantial for all three groups (radiologists: 0.78; residents: 0.65; and urologists: 0.61). Among residents, the sensitivity/specificity/accuracy rates of combined CT/MRI vs. CT alone were 82.1/74.7/76.7% vs. 75.0/66.7/68.9%, and specificity and accuracy were significantly higher for combined CT/MRI than that for CT alone (p = 0.03 and 0.008, respectively). Similarly, sensitivity/specificity/accuracy values were significantly higher for combined CT/MRI among urologists (78.6/73.3/74.8% vs. 64.3/64.0/64.1%, p = 0.04/0.04/0.008). However, sensitivity/specificity/accuracy did not significantly differ between the two among radiologists (89.3/74.7/78.6% vs. 85.7/73.3/76.7%, p = 0.32/0.56/0.32).

Conclusions: Combined CT/MRI is useful for diagnosing malignancy in patients with CRMs using BC2019, especially for non-expert readers.

Abbreviation: Combined CT/MRI, adding MRI to CT examination; CRM, Cystic renal mass; CT, Computed tomography; MRI, Magnetic resonance imaging; SSFSE, Single-shot fast spin-echo imaging; T1WI, T1-weighted imaging; T2WI, T2-weighted imaging.

* Corresponding author.
E-mail addresses: yarita@rad.med.keio.ac.jp (Y. Arita), 6030uro@tmd.ac.jp (S. Yoshida), thomaskwee@gmail.com (T.C. Kwee), miki3suntree@yahoo.co.jp (H. Edo), kufukihara1988@yahoo.co.jp (R. Kufukihara), keisukeshigeta@yahoo.co.jp (K. Shigeta), misa0419.ok@gmail.com (M. Nagasaka), r-tkst@rad.med.keio.ac.jp (R. Takeshita), hokamura@rad.med.keio.ac.jp (H. Okamura), dyeorua@gmail.com (R. Ueda), rishii@md.tsukuba.ac.jp (R. Ishii), okuda@rad.med.keio.ac.jp (S. Okuda), y-fujii.uro@tmd.ac.jp (Y. Fujii).

https://doi.org/10.1016/j.ejrad.2022.110163
Received 16 December 2021; Received in revised form 11 January 2022; Accepted 13 January 2022
Available online 19 January 2022

1. Introduction

The Bosniak classification system was first described in 1986 for computed tomography (CT)-based risk stratification of cystic renal masses (CRMs) according to the probability of malignancy [1]. This framework was intended to improve the ability of radiologists and other physicians to differentiate between benign and malignant CRMs [1–4] and guide the clinical management by radiologists and urologists [5,6]. After its introduction, the classification system was modified to include category IIF in the 2005 Bosniak classification, thus allowing a subset of CRMs to be followed up during imaging surveillance because of their low probability of malignancy [7–11]. However, the original Bosniak classification needed further improvement as it did not include any criterion for magnetic resonance imaging (MRI) evaluation and exhibited interobserver variability [12–16]. A recent *meta*-analysis by Schoots et al. demonstrated that the interobserver variability of the original Bosniak classification system ranged from 6 to 75%, possibly due to the unclear definitions of the criteria in each category [17].

The Bosniak classification system was updated in 2019 to address these shortcomings by incorporating the criteria for MRI and detailing subjective evaluation criteria [4,5,18]. Compared to the 2005 version, the latest version has been reported to provide a better interobserver agreement, irrespective of the radiologists' experience [19].

This classification system was originally developed by both radiologists and urologists; however, the impact of the 2019 Bosniak classification on the interobserver agreement between them and the value of adding MRI to CT examination (combined CT/MRI) remain unknown. This study aimed to compare the interobserver agreement between radiologists and urologists and the diagnostic value of CT alone and combined CT/MRI using the 2019 Bosniak classification.

2. Materials and methods

2.1. Patients

This study was approved by the Institutional Review Board. The requirement for written informed consent was waived due to the retrospective nature of the study. According to the standard workup of CRMs in our hospital, both CT and MRI are performed for the baseline evaluation, or MRI is considered for further evaluation after the use of other modalities, such as CT or ultrasonography. We searched the electronic database of the hospital and identified all patients with a CRM characterized using both CT and MRI between August 2010 and January 2016. Eligible patients were identified by searching the CT and MRI reports for the terms "cystic renal mass," "cystic renal neoplasm," or "Bosniak." The exclusion criteria were as follows: examinations obtained without using the renal mass protocol, for example, unenhanced acquisitions or presence of significant MR artifacts [4,5,18], an interval between CT and MRI that exceeded 3 months [20,21], an imaging follow-up period (if surgery was not performed) of<60 months [4,22,23], and a mass with more than 25% of the mass composed of enhancing tissue [4,5]. The information regarding patient/tumor characteristics, including data related to the clinical presentation and pathological findings, was obtained from the medical records of the patients.

2.2. Image acquisition

CT images were obtained using 128-slice thickness multi-detector CT. The triphasic renal mass CT protocol used at our institution consisted of non-contrast-enhanced acquisition of the kidneys followed by corticomedullary phase acquisition of the kidneys and nephrographic phase acquisition of the entire abdomen. The timing of the corticomedullary phase (approximately 30–40 s after the injection) was established using bolus tracking. The timing of the nephrographic phase was approximately 90–100 s after the injection. The images were acquired in the axial plane with a slice thickness of 1.25 mm. Multi-planar reformatted images were available for every patient. Iopamidol (Iopamiron 300 or 370; Bayer HealthCare Pharmaceuticals, Berlin, Germany) was used as a contrast agent to a maximum of 150.0 mL depending on body weight.

MRI examinations were performed using a 1.5-T MRI system with 32-channel phased-array coils. The following sequences were acquired to cover the upper abdomen: axial and coronal T2-weighted imaging (T2WI) with breath-holding, axial in-phase and opposed-phase gradient-echo sequences, axial diffusion-weighted imaging, and axial unenhanced and contrast-enhanced T1-weighted imaging (T1WI) with fat saturation at 30–40, 90–100, 180–190, and 240 s (LAVA-Flex, GE Healthcare). The specific sequence parameters are summarized in Supplementary Table 1. Gadobutrol (Gadavist; Bayer HealthCare Pharmaceuticals) was used as a contrast agent at a dose of 0.1 mL/kg body weight.

2.3. Image analysis

The readers included three board-certified radiologists with 33, 9, and 7 years of experience in abdominal imaging, three radiology residents with 2, 2, and 1 year of experience in abdominal imaging, and three board-certified urologists with 19, 11, and 8 years of clinical experience. The readers independently reviewed CT alone and combined CT/MRI using the 2019 Bosniak classification. For each reader, there was at least a 1-month interval between the reading pairs of images (i.e., CT alone and combined CT/MRI) to avoid recall bias. The critical imaging features assessed included unenhanced and contrast-enhanced attenuation or signal intensity, the appearance of the septa (thickness; smoothness, irregularity, or protrusions; and number), wall (thickness and smoothness, irregularity, or protrusions), and calcifications, according to the 2019 Bosniak classification [4]. All tumors were classified as benign (category I/II), low probability of malignancy (category IIF), or high probability of malignancy (category III/IV) by each reader based on these assessments. For combined CT/MRI assessment, we reviewed first the CT and subsequently the MRI data according to the critical features for each modality. The final combined CT/MRI Bosniak category was then determined according to the 2019 Bosniak classification guidelines (superior contrast for MRI may demonstrate soft tissue features better relative to CT; however, features such as the septa may appear thicker on MRI than on CT due to inherent artifacts) [4,5,18,24]. All nine readers were blinded to the clinical information, including the pathologic findings, follow-up imaging, and original interpretations. Subsequently, an overall Bosniak category and individual critical imaging features were determined for all tumors by consensus in each of the three groups of readers (radiologists, radiology residents, and urologists). We then evaluated the impact of the imaging protocol, that is, CT alone and combined CT/MRI, and the reader's specialty on the rate at which the consensus the Bosniak category was upgraded from category IIF to category III/IV during the follow-up and the histopathological (if surgery was performed) or clinical diagnosis of each Bosniak category.

2.4. Reference standard

For patients who underwent radical or partial nephrectomy, the final diagnosis was established based on the histopathological findings. For all other patients, all clinical and imaging follow-up data were used as the reference standard [4,22,23].

2.5. Statistical analysis

The interobserver agreement for the 2019 Bosniak classification for CT alone and combined CT/MRI among all nine readers and each of the three groups of readers (radiologists, radiology residents, and urologists) was assessed based on Fleiss' kappa analysis.

The sensitivity, specificity, and accuracy for the diagnosis of

malignancy according to threshold category III/IV were calculated. The impact of different imaging modalities and readers' specialties on diagnostic performance was assessed using McNemar's test.

The proportion of CRMs whose evaluations were upgraded from category IIF to category III/IV during follow-up and the malignancy rate were calculated. The impact of different imaging modalities and specialties of the readers on the diagnostic performance was assessed using weighted generalized score statistic. Moreover, the impact of the different imaging modalities and specialties of the readers on the detection of individual critical imaging features was assessed using Wilcoxon signed-rank tests. The corresponding interobserver agreement among all three groups of readers was assessed via Fleiss' kappa analysis.

The interobserver agreement was categorized as poor (<0.20), fair (0.20–0.39), moderate (0.40–0.59), substantial (0.60–0.79), almost perfect (0.80–0.99), or perfect (1.0) [20,21].

All p-values were two-sided, and p-values < 0.05 were considered statistically significant. Statistical analyses were performed using SAS software v. 9.0 (SAS Institute Inc., Cary, NC, USA.).

3. Results

3.1. Patient and tumor characteristics

This retrospective study assessed the data of 107 patients with CRMs characterized in both CT and MRI between August 2010 and January 2016. Data were obtained from the electronic database of the hospital. After excluding patients without images using the renal mass protocol (n = 5), patients with an interval exceeding 3 months (n = 5) between CT and MRI, patients with an imaging follow-up period of (if no surgery was performed) <60 months (n = 11), and patients with a mass with more than 25% of the mass composed of enhancing tissue (n = 3), we included 83 consecutive patients (27 women and 56 men; 103 CRMs) in the final analysis (Fig. 1). The median interval between CT and MRI was 4 weeks (range, 1–12 weeks). Of the 103 CRMs, 35 were surgically resected (radical nephrectomy: n = 14; partial nephrectomy: n = 21). The median time interval between the initial imaging (CT or MRI) and surgery was 7 months (range, 1–15 months). Histopathologically, 26 (74.3%) and nine (25.7%) masses were diagnosed as malignant and benign tumors, respectively. Of the remaining 68 masses, two (2.9%) were clinically diagnosed as malignant (solid compartment enlargement on subsequent imaging at follow-up), and 66 (97.1%) were clinically diagnosed as benign by confirming stability during at least 60 months of follow-up (median, 78 months; range, 60–125 months) based on all clinical and imaging follow-up data [4,22,23]. The baseline characteristics of the patients and renal masses are summarized in Table 1.

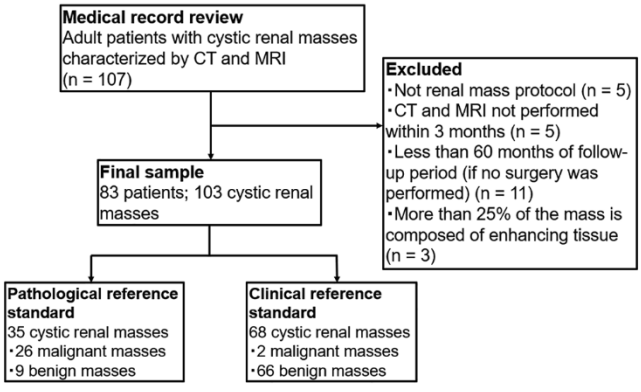

Fig. 1. Flowchart of patient selection. CT: Computed tomography; MRI: Magnetic resonance imaging.

Table 1
The baseline characteristics of the patients (n = 83) and renal masses (n = 103).

Characteristic	Value
Age (y)	
Median (range)	59 (27–85)
Sex	
Male (%)	56 (67.5)
Female (%)	27 (32.5)
Interval between CT and MRI (weeks)	
Median (range)	4 (1–12)
Size (mm)	
Median (range)	25.4 (9–160)
Renal mass follow-up	
Pathologic reference standard - n (%)	35/103 (34.0)
Radical nephrectomy	21
Partial nephrectomy	14
Clinical reference standard - n (%)	68/103 (66.0)
Follow-up period (months)	
Median (range)	78 (60–125)
Histologic results – n	35
Renal cell carcinoma	
Clear cell	19
Papillary	5
Clear cell papillary	1
MCNLMP	1
Benign	
Angiomyolipoma	4
Xanthogranulomatous reaction	3
MEST	2
Clinical results – n	
Malignant	2
Benign	66

CT: Computed tomography; MRI: Magnetic resonance imaging; MEST: Mixed epithelial and stromal tumor; MCNLMP: Multilocular cystic renal neoplasm of low malignant potential.

3.2. Comparison of interobserver agreement among different specialties

The distribution of the Bosniak categories determined by the different readers is summarized in Supplementary Table 2. In the assessment of CT alone, interobserver agreement was substantial for radiologists and residents, and moderate for urologists (Fleiss' κ values: 0.77, 0.63, and 0.58, respectively). For combined CT/MRI, interobserver agreement was substantial for radiologists, residents, and urologists (Fleiss' κ values: 0.78, 0.65, and 0.61, respectively).

The number of cases changing from category I/II on CT alone to category IIF on combined CT/MRI was 8 (7.8%) for radiologists, 12 (11.7%) for residents, and 13 (12.6%) for urologists. Furthermore, the numbers of cases changing from category IIF on CT alone to category III/IV on combined CT/MRI were 4 (3.9%) for radiologists, 6 (5.8%) for residents, and 8 (7.8%) for urologists. A representative case of discordance in the categorization between CT and MRI is shown in Fig. 2.

3.3. Diagnostic performance of the overall Bosniak categorization based on each specialty's consensus

The diagnostic performances of each group of readers (radiologists, residents, and urologists) for detecting malignancy are shown in Table 2.

Among residents, the sensitivity/specificity/accuracy rates of combined CT/MRI vs. CT alone were 82.1/74.7/76.7% vs. 75.0/66.7/68.9%, respectively, for diagnosing malignancy by threshold category III/IV. The specificity and accuracy were significantly higher for combined CT/MRI than that for CT alone ($p = 0.03$ and 0.008, respectively), whereas the sensitivity did not significantly differ ($p = 0.10$). Among the urologists, the sensitivity/specificity/accuracy values were significantly higher for combined CT/MRI than that for CT alone (78.6/73.3/74.8% vs. 64.3/64.0/64.1%, $p = 0.04/0.04/0.008$). However, among the radiologists, sensitivity/specificity/accuracy for detecting malignancy did not significantly differ between combined CT/MRI and CT alone (89.3/

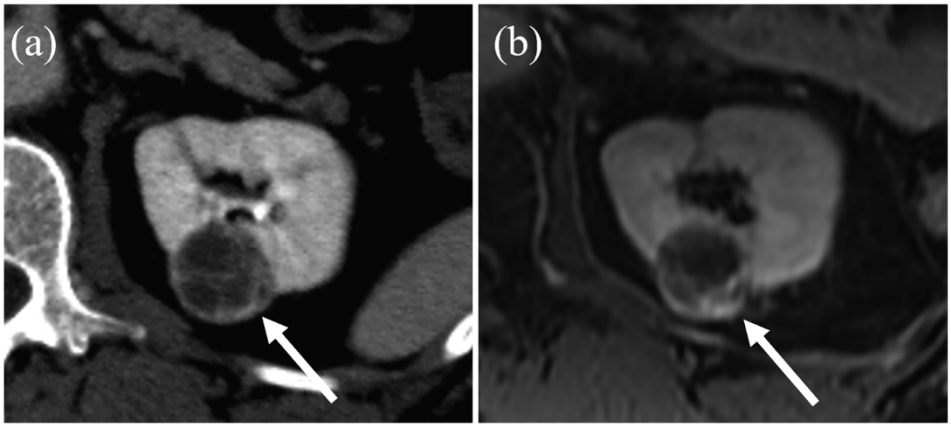

Fig. 2. A representative case of discordance in Bosniak categorization between CT and MRI. A 46-year-old woman with a complex left renal cyst: (a) Axial post-contrast CT scan in the nephrographic phase showing cystic masses with a smooth, minimally thickened (≤3 mm width) enhancing wall, with several (≥4) smooth thin (≤2 mm width) enhancing septa (arrow). (b) Axial post-contrast fat-suppressed T1WI scan in the nephrographic phase showing an enhancing thick wall (4 mm width), with several (≥4) smooth thin (≤2 mm width) enhancing septa (arrow). The lesion was classified as category IIF by CT and category III by MRI by all three reader groups (i.e., radiologists, residents, and urologists). According to the pathological specimens obtained by radical nephrectomy, the lesion was diagnosed as clear-cell renal cell carcinoma. CT underestimated the width of the enhancing wall thickness (3 mm width) on the axial post-contrast scan. CT: Computed tomography; MRI: Magnetic resonance imaging; T2WI: T2-weighted imaging; T1WI: T1-weighted imaging.

Table 2
Diagnostic performance for the high-malignancy (III, IV) and low-malignancy (IIF) 2019 Bosniak classification categories.

	Radiologists		Residents		Urologists	
Characteristic	CT	Combined CT/MRI	CT	Combined CT/MRI	CT	Combined CT/MRI
Bosniak III/IV lesions						
Sensitivity	24/28 (85.7)	25/28 (89.3)	21/28 (75.0)	23/28 (82.1)	18/28 (64.3)	22/28 (78.6)
p-value	Ref	0.32	Ref	0.10	Ref	0.04*
Specificity	55/75 (73.3)	56/75 (74.7)	50/75 (66.7)	56/75 (74.7)	48/75 (64.0)	55/75 (73.3)
p-value	Ref	0.56	Ref	0.03*	Ref	0.04*
Accuracy	79/103 (76.7)	81/103 (78.6)	71/103 (68.9)	79/103 (76.7)	66/103 (64.1)	77/103 (74.8)
p-value	Ref	0.32	Ref	0.008*	Ref	0.008*
Bosniak IIF lesions						
Progressed to category III/IV at follow-up	4/34 (11.8)	4/37 (10.8)	3/33 (9.1)	4/40 (10.0)	3/36 (8.3)	4/41 (9.8)
p-value	Ref	0.32	Ref	0.75	Ref	0.29
Malignancy rate	3/34 (8.8)	3/37 (8.1)	2/33 (6.1)	3/40 (7.5)	2/36 (5.6)	3/41 (7.3)
p-value	Ref	0.32	Ref	0.60	Ref	0.29

CT: Computed tomography; MRI: Magnetic resonance imaging; Combined CT/MRI: Adding MRI to CT examination.
Data are presented as numbers (percentages).
* $p < 0.05$ is statistically significant.

74.7/78.6% vs. 85.7/73.3/76.7%, $p = 0.32/0.56/0.32$).

The proportion of CRMs whose evaluations were upgraded from category IIF to category III/IV during the follow-up did not significantly differ between combined CT/MRI and CT alone according to the evaluations of all three reader groups (radiologists: 10.8% vs. 11.8%, $p = 0.32$; residents: 10.0% vs. 9.1%, $p = 0.75$; urologists: 9.8% vs. 8.3%, $p = 0.29$).

Table 3
Distribution of the readers' assessments of septa and protrusions on CT and MRI (n = 103).

Characteristic	Radiologists		Residents		Urologists		Fleiss' κ values	
	CT	Combined CT/MRI	CT	Combined CT/MRI	CT	Combined CT/MRI	CT	Combined CT/MRI
Septa								
Number							0.66	0.68
0	41	29	38	25	37	24		
1–3	45	52	46	51	46	50		
≥ 4	17	22	19	27	20	29		
p-value	Ref	0.04*	Ref	0.002*	Ref	0.009*		
Maximal thickness (mm)							0.73	0.67
≤ 2	52	47	55	49	56	47		
3	25	28	24	27	25	29		
≥ 4	26	28	24	27	22	27		
p-value	Ref	0.29	Ref	0.35	Ref	0.15		
Protrusions							0.67	0.69
None	89	88	91	89	91	88		
Present	14	15	12	14	12	15		
p-value	Ref	0.75	Ref	0.70	Ref	0.47		

CT: Computed tomography; MRI: Magnetic resonance imaging; Combined CT/MRI: Adding MRI to CT examination.
Data are presented as numbers.
* $p < 0.05$ is statistically significant.

3.4. Impact of the different modalities and readers' specialties on the detection of critical imaging features

Individual critical imaging features and the corresponding interobserver agreement values are summarized in Table 3.

All three reader groups (radiologists, residents, and urologists) identified a significantly higher number of septa on combined CT/MRI than that on CT alone (radiologists: $p = 0.04$; residents: $p = 0.02$; urologists: $p = 0.009$). The corresponding interobserver agreements for detecting the number of septa among all three reader groups were substantial for CT alone (Fleiss' $\kappa = 0.66$) and combined CT/MRI (Fleiss' $\kappa = 0.68$). A representative case in which more septa were visible on MRI than that on CT is shown in Fig. 3.

Regarding the measurement of the maximal wall or septa thickness, no significant difference was observed between combined CT/MRI and CT alone in any of the reader groups ($p = 0.29$ for radiologists, $p = 0.35$ for residents, $p = 0.15$ for urologists). The corresponding interobserver agreement among all three reader groups was substantial for CT alone (Fleiss' $\kappa = 0.73$) and combined CT/MRI (Fleiss' $\kappa = 0.67$). Regarding the detection number of protrusions, no significant difference was observed between combined CT/MRI and CT alone in any of the reader groups ($p = 0.75$ for radiologists, $p = 0.70$ for residents, $p = 0.47$ for urologists). The corresponding interobserver agreement among all three reader groups was substantial for CT alone (Fleiss' $\kappa = 0.67$) and for combined CT/MRI (Fleiss' $\kappa = 0.69$).

4. Discussion

In this study, we compared the interobserver agreement for CT alone and combined CT/MRI among three groups of readers using the 2019 Bosniak classification to determine the malignancy of CRMs. Combined CT/MRI yielded similarly high interobserver agreements among radiology residents, board-certified urologists, and board-certified radiologists. All three groups detected significantly more septa on combined CT/MRI than that on CT alone, and the interobserver agreement among the three groups was substantially high when assessing important imaging features, such as the septa, wall thickness, and protrusions, using combined CT/MRI. Notably, radiology residents and urologists achieved significantly higher overall accuracies for malignancies of categories III/IV with combined CT/MRI than with CT alone. Therefore, combined CT/MRI scans may improve the diagnostic utility of the 2019 Bosniak classification, especially in non-expert readers such as residents and urologists.

A recent study by Pacheco et al. reported that interobserver agreement for the 2019 Bosniak classification using a single modality, such as CT alone or MRI alone, was better among radiology fellows than among radiology residents [25], which is consistent with the results of our study. Chan et al. recently reported that the sensitivity, specificity, and accuracy for CT alone were 86.3–100%, 40.9–73.0%, and 71.0–87.0%, respectively, among radiologists [23]. Lucocq et al. reported that the malignancy rates of Bosniak III and IV cysts evaluated by radiologists were 79.3% and 84.5%, respectively [26]. Moreover, according to a recent report by Hindman et al., the rate of progression to malignancy for Bosniak IIF cysts evaluated by radiologists was 10.9% [22]. Our results for non-expert reader groups (radiology residents and urologists) approached these previous findings when combined CT/MRI was used to determine the category as per the Bosniak classification. Meanwhile, our results for radiologists were comparable to the results obtained in the previous studies even when CT alone was used to determine the category according to the Bosniak classification.

From a cost perspective, performing both CT and MRI for the evaluation of all CRM cases may not be recommended in daily clinical practice. As mentioned above, however, applying the 2019 Bosniak classification using CT alone may include a risk of misdiagnosis by non-specialists, such as radiology residents and urologists. Therefore, the detection, characterization, and classification of CRMs based on CT alone may be considered an exclusive task for radiologists who regularly evaluate images of the kidneys. On the other hand, additional MRI had diagnostic value in patients with thick or nodular calcification at CT (which obscured visualization and characterization of enhancing components) and in some cases of papillary renal cell carcinoma which exhibited equivocal (10–20 HU) or absent (<10 HU) enhancement at CT in our study. This finding is consistent with previous works by Krishna, et al. and Dilauro, et al. [18,27]. However, further investigations with a larger cohort are warranted to confirm that additional MRI would be appropriate in the above-mentioned subgroup of patients in clinical routine.

Critical imaging features, such as thin (versus thick) walls or septa, few (versus numerous) septa, and nodular (versus irregularly thickened) walls or septa, have been explicitly defined in the 2019 Bosniak classification [4]. In the present study, all three reader groups (radiologists, residents, and urologists) detected a significantly higher number of septa on combined CT/MRI than on CT alone, which resulted in an upgrade of the final category in several cases. As described in the literature regarding the 2019 Bosniak classification criteria, MRI is an important tool for the evaluation of cystic masses, in addition to solid renal masses, and is particularly valuable in characterizing cystic masses that are indeterminate on CT and ultrasound images because of the superior contrast resolution and sensitivity of MRI for enhancement [4,5,18,28].

The present study found a high level of interobserver agreement among different readers (radiologists, residents, and urologists) for the assessment of critical imaging features on both CT alone and combined CT/MRI. However, certain features of CRMs may remain difficult to reproducibly quantify or qualify using MRI, as reported by Edney et al. [24], who indicated that pitfalls in interpretation might occur due to differences between T2W images and post-contrast T1W images [24].

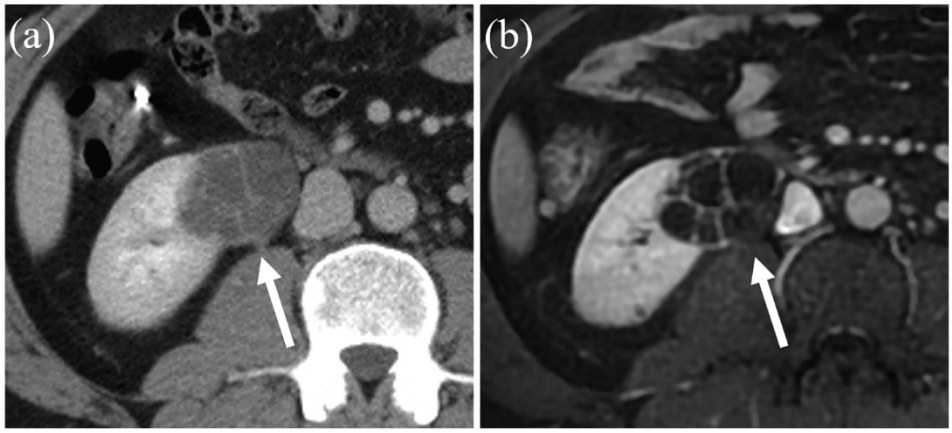

Fig. 3. A representative case in which more septa were visible on MRI than on CT. A 35-year-old-man with a complex right renal cyst: (a) The axial post-contrast CT shows cystic masses with thin (≤2 mm width) and few [1–3] enhancing septa (arrow). (b) The axial post-contrast fat-suppressed T1WI scan in the nephrographic phase shows cystic masses with several (≥4) smooth thin (≤2 mm width) enhancing septa (arrow). All three reader groups (i.e., radiologists, residents, and urologists) identified more septa on MRI than on CT. The lesion was classified as category II on CT but as category IIF on MRI by all three reader groups.CT: Computed tomography; MRI: Magnetic resonance imaging; T1WI: T1-weighted imaging.

Most renal mass MRI protocols use two-dimensional single-shot fast spin-echo imaging (SSFSE) due to its insensitivity to motion and its excellent contrast between fluid and soft tissues, such as septa and nodules [29,30]. However, three-dimensional T1W spoiled gradient-echo acquisitions (i.e., LAVA) are more sensitive to the patients' respiratory motion and have lower in-plane resolution than T2W SSFSE, both of which may contribute to image deterioration [31]. These inherent differences may result in variations in the assessment of the small internal architecture of CRMs. Therefore, specifying or prioritizing the sequence (specifically T2WI or post-contrast fat-suppressed T1WI) for evaluating each critical imaging feature may further improve the interobserver agreement for MRI assessment. However, further studies are required to evaluate this hypothesis.

5. Limitations

Our study had several limitations. First, the readers only assessed the baseline (initial) CT and MRI scans, whereas radiologists often compare subsequent imaging examinations to evaluate the temporal stability or growth in clinical practice, which may also provide information regarding the malignant behavior [20,21]. Second, several masses were categorized as malignancies based on clinical reference standard and not based on histopathological findings following surgical intervention [4,22,23]. Third, according to the standard workup of CRMs in our hospital, both CT and MRI were performed during baseline evaluation. However, in some cases, MRI was performed for further evaluation of CRMs after evaluation by other modalities, such as CT or ultrasonography, which may have introduced some selection bias [4,5,18,28]. Lastly, this retrospective study included a relatively small sample size. Further large-scale validation studies are warranted to confirm our findings.

6. Conclusion

Combined CT/MRI resulted in substantially higher interobserver agreement among radiologists, residents, and urologists. Moreover, the diagnostic performance for category III/IV malignancy improved significantly for residents and urologists with combined CT/MRI compared to evaluations based on CT alone. Thus, combined CT/MRI is useful for diagnosing malignancy in CRMs using the 2019 Bosniak classification, especially for non-expert readers.

CRediT authorship contribution statement

Yuki Arita: Investigation, Writing – original draft. **Soichiro Yoshida:** Writing – review & editing. **Thomas C. Kwee:** Writing – review & editing. **Hiromi Edo:** Resources, Data curation. **Ryohei Kufukihara:** Resources, Data curation. **Keisuke Shigeta:** Resources, Data curation. **Misa Nagasaka:** Resources, Data curation. **Ryo Takeshita:** Resources, Data curation. **Haruka Okamura:** Resources, Data curation. **Ryo Ueda R.T.:** Resources, Data curation. **Ryota Ishii:** Resources, Data curation. **Shigeo Okuda:** Resources, Data curation. **Yasuhisa Fujii:** Supervision, Project administration.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejrad.2022.110163.

References

[1] M.A. Bosniak, The current radiological approach to renal cysts, Radiology. 158 (1) (1986) 1–10, https://doi.org/10.1148/radiology.158.1.3510019.

[2] M.A. Bosniak, The use of the Bosniak classification system for renal cysts and cystic tumors, J. Urol. 157 (5) (1997) 1852–1853, https://doi.org/10.1016/S0022-5347(01)64883-3.

[3] G.M. Israel, M.A. Bosniak, An update of the Bosniak renal cyst classification system, Urology. 66 (3) (2005) 484–488, https://doi.org/10.1016/j.urology.2005.04.003.

[4] S.G. Silverman, I. Pedrosa, J.H. Ellis, N.M. Hindman, N. Schieda, A.D. Smith, E. M. Remer, A.B. Shinagare, N.E. Curci, S.S. Raman, S.A. Wells, S.D. Kaffenberger, Z. J. Wang, H. Chandarana, M.S. Davenport, Bosniak Classification of Cystic Renal Masses, Version 2019: An Update Proposal and Needs Assessment, Radiology, version 292 (2) (2019) 475–488, https://doi.org/10.1148/radiol.2019182646.

[5] N. Schieda, M.S. Davenport, S. Krishna, E.A. Edney, I. Pedrosa, N. Hindman, R. H. Baroni, N.E. Curci, A. Shinagare, S.G. Silverman, Bosniak Classification of Cystic Renal Masses, Version 2019: A Pictorial Guide to Clinical Use, RadioGraphics, version 41 (3) (2021) 814–828, https://doi.org/10.1148/rg.2021200160.

[6] M.S. Davenport, E.M. Hu, A.D. Smith, H. Chandarana, K. Hafez, G.S. Palapattu, J. Stuart Wolf, S.G. Silverman, Reporting standards for the imaging-based diagnosis of renal masses on CT and MRI: a national survey of academic abdominal radiologists and urologists, Abdom. Radiol. (NY) 42 (4) (2017) 1229–1240, https://doi.org/10.1007/s00261-016-0962-x.

[7] M.A. Bosniak, The small (less than or equal to 3.0 cm) renal parenchymal tumor: detection, diagnosis, and controversies, Radiology. 179 (2) (1991) 307–317, https://doi.org/10.1148/radiology.179.2.2014269.

[8] M.A. Bosniak, Problems in the radiologic diagnosis of renal parenchymal tumors, Urol. Clin. North Am. 20 (2) (1993) 217–230, https://doi.org/10.1016/S0094-0143(21)00482-1.

[9] T.E. Wilson, E.A. Doelle, R.H. Cohan, K. Wojno, M. Korobkin, Cystic renal masses: A reevaluation of the usefulness of the Bosniak classification system, Acad. Radiol. 3 (7) (1996) 564–570, https://doi.org/10.1016/s1076-6332(96)80221-2.

[10] M.A. Bosniak, Diagnosis and management of patients with complicated cystic lesions of the kidney, AJR Am. J. Roentgenol. 169 (3) (1997) 819–821, https://doi.org/10.2214/ajr.169.3.9275903.

[11] G.M. Israel, M.A. Bosniak, Follow-up CT of moderately complex cystic lesions of the kidney (Bosniak category IIF), AJR Am. J. Roentgenol. 181 (3) (2003) 627–633, https://doi.org/10.2214/ajr.181.3.1810627.

[12] G.M. Israel, N. Hindman, M.A. Bosniak, Evaluation of cystic renal masses: comparison of CT and MR imaging by using the Bosniak classification system, Radiology. 231 (2) (2004) 365–371, https://doi.org/10.1148/radiol.2312031025.

[13] A.M. Ferreira, R.B. Reis, P.P. Kajiwara, G.E.B. Silva, J. Elias, V.F. Muglia, MRI evaluation of complex renal cysts using the Bosniak classification: a comparison to CT, Abdom. Radiol. (NY) 41 (10) (2016) 2011–2019, https://doi.org/10.1007/s00261-016-0797-5.

[14] J. Zhong, F. Cao, X. Guan, J. Chen, Z. Ding, M. Zhang, Renal cyst masses (Bosniak category II–III) may be over evaluated by the Bosniak criteria based on MR findings, Med. (Baltim.) 96 (2017), e9361, https://doi.org/10.1097/MD.0000000000009361.

[15] W.B. Kim, S.W. Lee, S.W. Doo, W.J. Yang, Y.S. Song, J.S. Jeon, I.H. Choi, S.Y. Jin, Category migration of renal cystic masses with use of gadolinium-enhanced magnetic resonance imaging, Korean, J. Urol. 53 (8) (2012) 573, https://doi.org/10.4111/kju.2012.53.8.573.

[16] O. Graumann, S.S. Osther, J. Karstoft, A. Hørlyck, P.JS. Osther, Bosniak classification system: inter-observer and intra-observer agreement among experienced uroradiologists, Acta Radiol. 56 (3) (2015) 374–383, https://doi.org/10.1177/0284185114529562.

[17] I.G. Schoots, K. Zaccai, M.G. Hunink, P.C.M.S. Verhagen, Bosniak classification for complex renal cysts reevaluated: a systematic review, J. Urol. 198 (1) (2017) 12–21, https://doi.org/10.1016/j.juro.2016.09.160.

[18] S. Krishna, N. Schieda, I. Pedrosa, et al., Update on MRI of cystic renal masses including Bosniak version 2019, J. Magn. Reson. Imaging. 54 (2021) 341–356, https://doi.org/10.1002/jmri.27364.

[19] X.u. Bai, S.-M. Sun, W. Xu, H.-H. Kang, L. Li, Y.-Q. Jin, Q.-G.-L. Gong, G.-C. Liang, H.-Y. Liu, L.-L. Liu, S.-L. Chen, Q.-R. Wang, P. Wu, A.-T. Guo, Q.-B. Huang, X.-J. Zhang, H.-Y. Ye, H.-Y. Wang, MRI-based Bosniak Classification of Cystic Renal Masses, Version 2019: Interobserver Agreement, Impact of Readers' Experience, and Diagnostic Performance, Radiology, version 297 (3) (2020) 597–605, https://doi.org/10.1148/radiol.2020200478.

[20] J.R. Tse, J. Shen, L. Yoon, A. Kamaya, Bosniak classification version 2019 of cystic renal masses assessed with MRI, AJR Am. J. Roentgenol. 215 (2020) 413–419. https://doi.org/10.2214/AJR.19.22740.

[21] J.R. Tse, J. Shen, L. Shen, L. Yoon, A. Kamaya, Bosniak classification of cystic renal masses version 2019: comparison of categorization using CT and MRI, AJR Am. J. Roentgenol. 216 (2021) 412–420. https://doi.org/10.2214/AJR.20.23656.

[22] N.M. Hindman, E.M. Hecht, M.A. Bosniak, Follow-up for Bosniak category 2F cystic renal lesions, Radiology. 272 (3) (2014) 757–766, https://doi.org/10.1148/radiol.14122908.

[23] J. Chan, J.H. Yan, J. Munir, H. Osman, S. Alrasheed, T. McGrath, T. Flood, N. Schieda, assessed by CT and MRI, Abdom. Radiol. (NY) 46 (11) (2021) 5268–5276, https://doi.org/10.1007/s00261-021-03236-z.

[24] E. Edney, M.S. Davenport, N. Curci, N. Schieda, S. Krishna, N. Hindman, S. G. Silverman, I. Pedrosa, Bosniak classification of cystic renal masses, version 2019: interpretation pitfalls and recommendations to avoid misclassification, Abdom. Radiol. (NY), version 46 (6) (2021) 2699–2711, https://doi.org/10.1007/s00261-020-02906-8.

[25] E.O. Pacheco, U.S. Torres, A.M.A. Alves, D. Bekhor, G. D'Ippolito, Bosniak classification of cystic renal masses version 2019 does not increase the interobserver agreement or the proportion of masses categorized into lower Bosniak classes for non-subspecialized readers on CT or MR, Eur. J. Radiol. 131 (2020) 109270, https://doi.org/10.1016/j.ejrad.2020.109270.

[26] J. Lucocq, S. Pillai, R. Oparka, G. Nabi, Complex renal cysts (Bosniak ≥IIF): interobserver agreement, progression and malignancy rates, Eur. Radiol. 31 (2) (2021) 901–908, https://doi.org/10.1007/s00330-020-07186-w.

[27] M. Dilauro, M. Quon, M.D.F. McInnes, M. Vakili, A. Chung, T.A. Flood, N. Schieda, Comparison of Contrast-Enhanced Multiphase Renal Protocol CT Versus MRI for Diagnosis of Papillary Renal Cell Carcinoma, American Journal of Roentgenology 206 (2) (2016) 319–325, https://doi.org/10.2214/AJR.15.14932.

[28] S.E. Zeng, M.Y. Du, Y. Yu, et al, Ultrasound, CT, and MR imaging for evaluation of cystic renal masses, J. Ultrasound Med. (2021); Epub ahead of print. https://doi.org/10.1002/jum.15762.

[29] Y. Yamashita, Y.i. Tang, Y. Abe, K. Mitsuzaki, M. Takahashi, Comparison of ultrafast half-Fourier single-shot turbo spin-echo sequence with turbo spin-echo sequences for T2-weighted imaging of the female pelvis, J. Magn. Reson. Imaging. 8 (6) (1998) 1207–1212, https://doi.org/10.1002/jmri.1880080605.

[30] R.C. Semelka, N.L. Kelekis, D. Thomasson, M.A. Brown, G.A. Laub, Haste MR imaging: description of technique and preliminary results in the abdomen, J. Magn. Reson. Imaging. 6 (4) (1996) 698–699, https://doi.org/10.1002/jmri.1880060420.

[31] G.E. Wile, J.R. Leyendecker, Magnetic Resonance Imaging of the Liver: Sequence Optimization and Artifacts, Magnetic Resonance Imaging Clinics of North America 18 (3) (2010) 525–547, https://doi.org/10.1016/j.mric.2010.07.010.