



Bosniak classification version 2019: a prospective comparison of CT and MRI

Yassir Edrees Almalki¹ · Mohammad Abd Alkhalik Basha² · Rania Refaat³ · Sharifa Khalid Alduraibi⁴ · Ahmed A. El-Hamid M. Abdalla² · Hala Y. Yousef² · Mohamed M. A. Zaitoun² · Saeed Bakry Elsayed² · Nader E. M. Mahmoud² · Nader Ali Alayouty² · Susan Adil Ali³ · Ahmad Abdullah Alnaggar² · Sameh Saber² · Ahmed Mohamed El-Maghraby² · Amgad M. Elsheikh² · Mohamed Hesham Saleh Saleh Radwan² · Ahmed Gamil Ibrahim Abdelmegid² · Sameh Abdelaziz Aly⁵ · Waleed S. Abo Shanab⁶ · Ahmed Ali Obaya⁷ · Shaimaa Farouk Abdelhai⁷ · Shereen Elshorbagy⁸ · Yasser M. Haggag⁹ · Hwaida M. Mokhtar¹⁰ · Nesreen M. Sabry¹¹ · Jehan Ibrahim Altohamy¹² · Rasha Taha Abouelkheir¹³ · Tawfik Omran¹⁴ · Ahmed Shalan⁵ · Youssef H. Algazzar¹⁵ · Maha Ibrahim Metwally²

Received: 25 March 2022 / Revised: 13 June 2022 / Accepted: 19 July 2022
© The Author(s), under exclusive licence to European Society of Radiology 2022

Abstract

Objective To assess the diagnostic accuracy and agreement of CT and MRI in terms of the Bosniak classification version 2019 (BCv2019).

Materials and methods A prospective multi-institutional study enrolled 63 patients with 67 complicated cystic renal masses (CRMs) discovered during ultrasound examination. All patients underwent CT and MRI scans and histopathology. Three radiologists independently assessed CRMs using BCv2019 and assigned Bosniak class to each CRM using CT and MRI. The final analysis included 60 histopathologically confirmed CRMs (41 were malignant and 19 were benign).

Results Discordance between CT and MRI findings was noticed in 50% (30/60) CRMs when data were analyzed in terms of the Bosniak classes. Of these, 16 (53.3%) were malignant. Based on consensus reviewing, there was no difference in the sensitivity, specificity, and accuracy of the BCv2019 with MRI and BCv2019 with CT (87.8%; 95% CI = 73.8–95.9% versus 75.6%; 95% CI = 59.7–87.6%; $p = 0.09$, 84.2%; 95% CI = 60.4–96.6% versus 78.9%; 95% CI = 54.4–93.9%; $p = 0.5$, and 86.7%; 95% CI = 64.0–86.6% versus 76.7%; 95% CI = 75.4–94.1%; $p = 0.1$, respectively). The number and thickness of septa and the presence of enhanced nodules accounted for the majority of variations in Bosniak classes between CT and MRI. The inter-reader agreement

✉ Mohammad Abd Alkhalik Basha
Mohammad_basha76@yahoo.com

¹ Division of Radiology, Department of Internal Medicine, Medical College, Najran University, Najran, Kingdom of Saudi Arabia

² Department of Diagnostic Radiology, Faculty of Human Medicine, Zagazig University, Zagazig, Egypt

³ Department of Diagnostic Radiology, Intervention and Molecular Imaging, Faculty of Human Medicine, Ain Shams University, Cairo, Egypt

⁴ Department of Radiology, College of Medicine, Qassim University, Buraidah, Kingdom of Saudi Arabia

⁵ Department of Diagnostic Radiology, Faculty of Human Medicine, Benha University, Benha, Egypt

⁶ Department of Diagnostic Radiology, Faculty of Human Medicine, Port Said University, Port Said, Egypt

⁷ Department of Clinical Oncology, Faculty of Human Medicine, Zagazig University, Zagazig, Egypt

⁸ Department of Medical Oncology, Faculty of Human Medicine, Zagazig University, Zagazig, Egypt

⁹ Department of Urology, Faculty of Human Medicine, Al Azhar University, Cairo, Egypt

¹⁰ Department of Diagnostic Radiology, Faculty of Human Medicine, Tanta University, Tanta, Egypt

¹¹ Department of Clinical Oncology, Faculty of Human Medicine, Tanta University, Tanta, Egypt

¹² Department of Diagnostic Radiology, National Institute of Urology and Nephrology, Cairo, Egypt

¹³ Department of Diagnostic Radiology, Urology and Nephrology Center, Mansoura University, Mansoura, Egypt

¹⁴ Department of Diagnostic Radiology, Faculty of Human Medicine, Helwan University, Cairo, Egypt

¹⁵ Department of Internal Medicine, Katai Gabor Hospital, Karcag, Hungary

(IRA) was substantial for determining the Bosniak class in CT and MRI ($k = 0.66$; 95% CI = 0.54–0.76, $k = 0.62$; 95% CI = 0.50–0.73, respectively). The inter-modality agreement of the BCv2019 between CT and MRI was moderate ($\kappa = 0.58$).

Conclusion In terms of BCv2019, CT and MRI are comparable in the classification of CRMs with no significant difference in diagnostic accuracy and reliability.

Key Points

- *There is no significant difference in the sensitivity, specificity, and accuracy of the BCv2019 with MRI and BCv2019 with CT.*
- *The number of septa and their thickness and the presence of enhanced nodules accounted for the majority of variations in Bosniak classes between CT and MRI.*
- *The inter-reader agreement was substantial for determining the Bosniak class in CT and MRI and the inter-modality agreement of the BCv2019 between CT and MRI was moderate.*

Keywords Classification · Kidney, cystic · Magnetic resonance imaging · Tomography, X-ray computed · Prospective study

Abbreviations

BC	Bosniak classification
BCv2019	Bosniak classification version 2019
CRMs	Cystic renal masses
CI	Confidence interval
CT	Computed tomography
ICC	Intraclass correlation coefficient
IRA	Inter-reader agreement
MRI	Magnetic resonance imaging
RCC	Renal cell carcinoma

Introduction

Cystic renal masses (CRMs) are prevalent in radiological practice, being discovered in up to 40% of patients on computed tomography (CT) [1, 2]. Although the vast majority of CRMs are simple benign cysts [3], up to 8–15% may have a complex appearance [4]. Imaging plays an essential role in the identification and characterization of CRMs as well as in grading and follow-up of patients with renal cell carcinoma (RCC) [5]. Ultrasound is frequently used as the first-line imaging evaluation of the abdomen and kidney and can easily detect simple, fluid-filled renal cysts [6]. Complex CRMs cannot be reliably described at the ultrasound and often need contrast-enhanced CT or magnetic resonance imaging (MRI) [6, 7].

Cross-sectional imaging reported an excellent sensitivity and specificity in the diagnosis of CRMs, but if a radiology report misses the diagnosis, there may be potential implications for the treating urologist [8]. In an attempt to standardize the description and management of CRMs, Morton Bosniak in 1986 designed an organized classification system called the Bosniak classification (BC) [9]. Although the original BC was initially based on CT findings, numerous researches published in the last two decades have shown that MRI may potentially be valuable [10–12]. Most recently, an update to the classification system (Bosniak classification version 2019) (BCv2019) has been released, which suggests several changes to the original classification [13]. The BCv2019 provides a

detailed algorithm that categorizes CRMs by their likelihood of being definitely benign (I), reliably benign (II), likely benign, but warranting follow-up (IIF), potentially malignant (III), and highly likely malignant (IV) [14]. Although BCv2019 now officially integrates MRI into the classification system and adds new modalities-specific criteria, it is unclear whether and how the technical variations between CT and MRI will affect CRM imaging and the final BCv2019 class [15]. The number of septa, the thickness of the wall and septa, and the irregularity/protrusion of the wall and septa may vary between CT and MR. Hence, the selection of approach may impact BCv2019 and the management of CRMs.

In the clinical setting of CRM diagnostic algorithms, there is an apparent information gap about the comparative accuracies of the various imaging approaches (researches are limited, single-institution, retrospective, lacked a strong reference standard, and lacked direct comparisons). Few studies have compared CT and MRI in BC using the old version (v2005) [10] and the recently updated version (v2019) [15, 16]. Owing to the difference between CT- and MRI-based BC, we performed this prospective multi-institutional research to assess the diagnostic accuracy and agreement of CT and MRI in terms of BCv2019.

Materials and methods

Ethical statement

This research was authorized by the institutional review boards, and each patient who participated in it gave informed consent. We conducted this research following the Declaration of Helsinki's ethical guidelines.

Study population

Between August 2019 and February 2022, we performed this multi-institutional prospective study. Initially, we collected 79 consecutive patients from six institutions (university hospitals

in the same country). Inclusion criteria were patients with complicated CRMs detected during ultrasound examination either incidentally or during the workup of urinary tract symptoms. Exclusion criteria are illustrated in Fig. 1. After exclusions, a cohort of 63 consecutive patients with 67 CRMs (4 (6.3%) patients had bilateral CRMs) was formed. The clinical-pathologic features of the enrolled patients and CRMs are detailed in Table 1 (34 males and 29 females, mean age = 49.5 ± 11.9 years, range = 27–76 years). All patients underwent CT and MRI scans of the kidney and pathology after biopsy or surgery. Sixty (89.6%) out of 67 CRMs were histopathologically examined (41 were RCC and 19 were benign). Clear cell RCC was the most common malignant CRMs (65.8%), and multilocular cystic neoplasm was the most common benign CRMs (31.6%). Seven CRMs with no histopathology (classified as Bosniak I and II by both CT and MRI) were excluded from the final analysis.

CT and MRI techniques

All CT scans were obtained using 64- or 128-MDCT scanners (Philips Healthcare Ingenuity; SOMATOM Definition, Siemens Healthcare; Aquilion Lightning; Toshiba Medical Systems). All MRI examinations were conducted within one week of the CT examinations. All examinations were performed on a 1.5-Tesla MR system (Optima 450 GEM, GE Healthcare; Achieva-class Iia, Philips Medical Systems;

Siemens Magnetom Avanto, Siemens Healthcare) using a phased-array torso coil. Renal CT and MRI techniques are detailed in [Appendix E1](#) in the Electronic Supplementary Material.

Analysis of images

CT and MRI scans were scheduled within 1 month following the ultrasound. All CT and MRI data were sent to the workstations for image processing. Image analysis and interpretation were conducted on the PACS system (PaxeraUltima-paxeramed). The CT and MRI images were separated for interpretation (i.e., the CT image findings were reviewed without knowledge of the MRI findings and vice versa). Three abdominal radiologists with over 10 years of renal CT and MRI experience (N.E., M.B., and M.M.) independently reviewed all CT images. After 1 month (to diminish the memory bias of readers), the same three radiologists (to avoid bias of different readers) independently reviewed all MR images. All radiologists were blinded to the patient data, sonographic reports, and final diagnoses. Before starting the study, the three radiologists received 2 h of lecture-based and hands-on training that completely described BCv2019. The following imaging features were individually evaluated in each CRM at CT and MRI: wall (thickness, irregularity/protrusion, and enhancement);

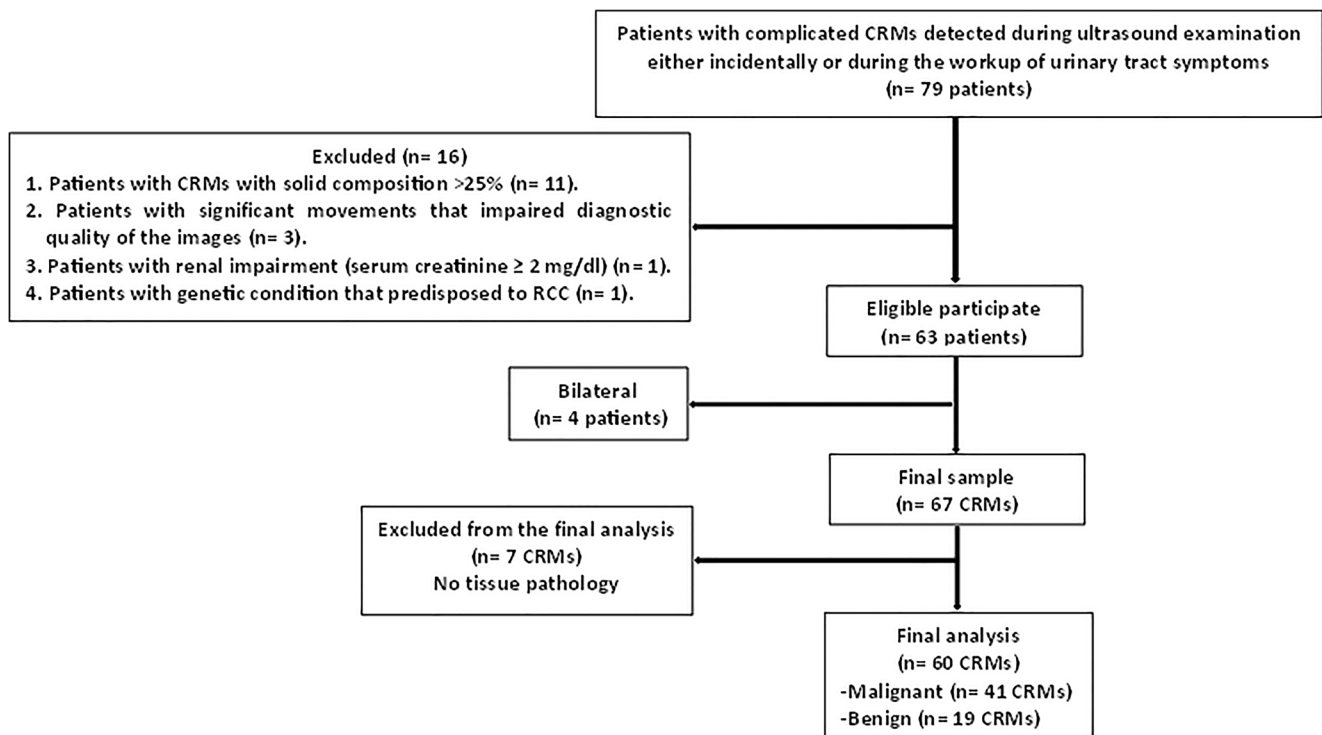


Fig. 1 Flowchart of the study. CRMs, cystic renal masses; RCC, renal cell carcinoma

Table 1 Clinical-pathologic features of patients and CRMs

Characteristic	Values
No. of patients	63
Age (y), mean \pm SD (range)	49.5 \pm 11.9 (27–76)
Sex of patients	
Male	34 (54)
Female	29 (46)
Laterality	
Right	26 (41.3)
Left	33 (52.4)
Bilateral	4 (6.3)
No. of CRMs	67
Size of CRM (mm), mean \pm SD (range)	65.9 \pm 21.0 (31–108)
Procedure for diagnosis	
Radical nephrectomy	38 (56.7)
Partial nephrectomy	10 (14.9)
Core biopsy	4 (7.5)
Fine needle aspiration biopsy	8 (11.9)
Histopathology	60 (89.6)
Benign	19 (31.7)
Multilocular cystic neoplasm	6 (31.6)
Benign cystic nephroma	5 (26.3)
Simple epithelial cyst	5 (26.3)
Mixed epithelial and stromal tumor	3 (15.8)
Malignant (RCC)	41 (68.3)
Clear cell	27 (65.8)
Papillary	8 (19.5)
Clear cell papillary	4 (9.8)
Chromophobe	2 (4.9)

Unless otherwise indicated, data are numbers of patients or CRMs with percentage in parentheses. *CRMs*, cystic renal masses; *SD*, standard deviation; *RCC*, renal cell carcinoma

septa (number, thickness, irregularity/protrusion, and enhancement); nodules; calcifications; and heterogeneity. The radiologists then used the BCv2019 algorithm [13] to assign a Bosniak class to all detected CRMs on CT and MRI. Finally, each CRM had 6 independent Bosniak classes (three by CT and three by MRI).

The inter-reader agreement (IRA) was assessed among the three independent radiologists regarding BCv2019 and imaging features of CRMs. After each radiologist had given its classification, the three radiologists worked together to reach the final Bosniak class of CRM via a collaborative consensus review. In case of disagreement among radiologists, all features were discussed in detail until a final agreement was obtained. The findings of consensus reviewing were used to determine the diagnostic accuracy of BCv2019 with each modality.

Reference standard

The final diagnoses were established based on histopathologic findings after surgery (48 CRMs), core biopsy (4 CRMs), or fine-needle aspiration biopsy (8 CRMs). Two experienced pathologists checked all specimens, and the results were obtained by consensus. The histopathology was obtained in all patients within 1 month from MRI. A biopsy was performed to determine the type of CRMs by the requesting clinician.

Statistical analysis

MedCalc (version 11.1) was used for all statistical analyses. The means and standard deviations were used to describe continuous variables. The counts and proportions were used to describe categorical variables. On a lesion-based analysis, we used the four-fold table test with 95% CIs to assess the diagnostic accuracy of BCv2019 to classify CRMs using histopathology as a reference standard. The IRA of BCv2019 and imaging features of CRMs were evaluated using the intraclass correlation coefficient (ICC) with 95% CIs. The inter-modality agreement of BCv2019 and imaging features between CT and MRI were evaluated using the weighted kappa (κ) statistics with 95% CIs. The κ values were calculated as follows: 0.01–0.20 = poor agreement =; 0.21–0.40 = fair agreement; 0.41–0.60 = moderate agreement; 0.61–0.80 = substantial agreement; and 0.81–1.0 = almost perfect agreement. A *p* value \leq 0.05 was considered to be statistically significant.

Results

Assignment of Bosniak classes

Table 2 summarized the determined Bosniak classes of the CRMs by CT and MRI stratified by readers and histopathology. Based on consensus reviewing, the proportions of malignancy in Bosniak IIF, III, and IV were 61.5% (95% CI = 31.6–86.1%), 88.2% (95% CI = 63.6–98.5%), and 88.9% (95% CI = 65.3–98.6), respectively by CT and 31.3% (95% CI = 11.0–58.7%), 87.5% (95% CI = 61.7–98.5%), and 95.7 (95% CI = 78.1–99.9%), respectively, by MRI with no statically significant differences (*p* = 0.2, 0.6, and 0.8, respectively).

Diagnostic accuracy of BCv2019 with CT and BCv2019 with MRI

Table 3 presented the diagnostic accuracy of BCv2019 with CT and MRI using histopathology as a reference standard. We considered a combination of Bosniak III and IV as definite for RCC diagnosis because the combined Bosniak III and IV produced higher levels of diagnostic accuracy than Bosniak

Table 2 Frequency distributions of Bosniak classes and CT and MRI imaging features of CRMs stratified by readers and histopathology with inter-reader agreement

CT	Reader 1		Reader 2		Reader 3		Consensus reviewing		Inter-reader agreement
	Benign	RCC	Benign	RCC	Benign	RCC	Benign	RCC	
Bosniak I	2 (100)	0	4 (100)	0	4 (100)	0	4 (100)	0	
Bosniak II	1 (100)	0	3 (60)	2 (40)	2 (66.7)	1 (33.3)	6 (75)	2 (25)	
Bosniak IIF	9 (34.6)	17 (65.4)	6 (50)	6 (50)	8 (74.1)	9 (52.9)	5 (38.5)	8 (61.5) [31.6–86.1]	
Bosniak III	4 (21.1)	15 (78.9)	2 (13.3)	13 (86.7)	4 (16)	21 (84)	2 (11.8)	15 (88.2) [63.6–98.5]	
Irregularity	2 (50)	8 (53.3)	0	8 (61.5)	2 (50)	14 (66.7)	0	10 (66.7) [38.4–88.2]	0.60 [0.38–0.75]
Wall/Septa thick	3 (75)	9 (60)	2 (100)	6 (46.2)	3 (75)	10 (47.6)	2 (100)	8 (53.3) [26.6–78.7]	0.57 [0.43–0.69]
Bosniak IV	3 (25)	9 (66.7)	4 (16.7)	20 (83.3)	1 (9.1)	10 (90.9)	2 (11.1)	16 (88.9) [65.3–98.6]	
Nodule with acute Angle	2 (66.7)	7 (77.8)	2 (50)	15 (75)	1 (100)	8 (60)	2 (100)	12 (75) [47.6–92.7]	0.71 [0.60–0.80]
Nodule > 4 mm with Obtuse angle	1 (33.3)	2 (22.2)	2 (50)	5 (25)	0	2 (40)	0	4 (25) [7.3–52.4]	0.58 [0.43–0.70]
MRI	Reader 1		Reader 2		Reader 3		Consensus reading		Inter-reader agreement
	Benign	RCC	Benign	RCC	Benign	RCC	Benign	RCC	
Bosniak I	2 (100)	0	4 (100)	0	2 (100)	0	1 (100)	0	
Bosniak II	7 (87.5)	1 (12.5)	1 (50)	1 (50)	2 (100)	0	4 (100)	0	
Bosniak IIF	4 (80)	1 (20)	7 (43.8)	9 (56.2)	9 (56.2)	7 (43.8)	11 (68.7)	5 (31.3) [11.0–58.7]	
Bosniak III	3 (13)	20 (87)	1 (8.3)	11 (91.7)	4 (36.4)	7 (63.6)	2 (12.5)	14 (87.5) [61.7–98.5]	
Irregularity	2 (66.7)	13 (65)	0	7 (63.6)	1 (25)	6 (85.7)	1 (50)	11 (78.6) [49.2–95.3]	0.66 [0.53–0.77]
Wall/Septa thick	2 (66.7)	10 (50)	1 (100)	6 (54.5)	4 (100)	4 (57.1)	2 (100)	8 (57.1) [28.9–82.3]	0.62 [0.45–0.74]
Bosniak IV	3 (13.6)	19 (86.4)	6 (23.1)	20 (76.9)	2 (10.5)	27 (89.5)	1 (4.3)	22 (95.7) [78.1–99.9]	
Nodule with acute Angle	2 (66.7)	1 (2 (63.2)	4 (66.7)	13 (65)	1 (50)	18 (66.7)	1 (100)	15 (68.2) [45.1–86.1]	0.76 [0.66–0.84]
Nodule > 4 mm with Obtuse angle	1 (33.3)	7 (36.8)	2 (33.3)	7 (35)	1 (50)	9 (33.3)	0	7 (31.8) [13.9–54.9]	0.70 [0.58–0.79]

Data are the number of CRMs with percentage in parentheses. Data in brackets are 95% confidence intervals. BCv2019 = Bosniak classification version 2019; CRMs, cystic renal masses; CT, computed tomography; MRI, magnetic resonance imaging; RCC, renal cell carcinoma. The inter-reader agreement was calculated using the intraclass correlation coefficient (ICC) with 95% confidence intervals. The k coefficient values were interpreted as follows: 0.00–0.20 = poor agreement; 0.21–0.40 = fair agreement; 0.41–0.60 = moderate agreement; 0.61–0.80 = substantial agreement; and 0.81–1.00 = almost perfect agreement

Table 3 Diagnostic accuracy of BCv2019 with CT and BCv2019 with MRI according to each reader and consensus reading

	Reader 1		Reader 2		Reader 3		Consensus reviewing		<i>p</i> value
	CT	MRI	CT	MRI	CT	MRI	CT	MRI	
Sensitivity (%)	58.5 (24/41) [42.1–73.7]	95.1 (39/41) [83.5–99.4]	80.5 (33/41) [65.1–91.2]	75.6 (31/41) [59.7–87.6]	75.6 (31/41) [59.7–87.6]	82.9 (34/41) [67.9–92.9]	75.6 (31/41) [59.7–87.6]	87.8 (36/41) [73.8–95.9]	0.09
Specificity (%)	63.2 (12/19) [38.4–83.7]	68.4 (13/19) [43.5–87.4]	68.4 (13/19) [43.5–87.4]	63.2 (12/19) [38.4–83.7]	73.7 (14/19) [48.8–90.9]	68.4 (13/19) [43.5–87.4]	78.9 (15/19) [54.4–93.9]	84.2 (16/19) [60.4–96.6]	0.5
Accuracy (%)	60.0 (36/60) [46.5–72.4]	86.7 (52/60) [75.4–94.1]	76.7 (46/60) [64.0–86.6]	71.7 (43/60) [58.6–82.6]	75.0 (45/60) [62.1–85.3]	78.3 (47/60) [65.8–87.9]	76.7 (46/60) [64.0–86.6]	86.7 (52/60) [75.4–94.1]	0.1

Data in parentheses were used to calculate percentages. Data in brackets are 95% confidence intervals. *BCv2019*, Bosniak classification version 2019; *CRMs*, cystic renal masses; *CT*, computed tomography; *MRI*, magnetic resonance imaging

IV alone Based on consensus reviewing, there was no significant difference in the sensitivity, specificity, and accuracy of the BCv2019 with MRI and BCv2019 with CT (87.8%; 95% CI = 73.8–95.9% versus 75.6%; 95% CI = 59.7–87.6%; $p = 0.09$, 84.2%; 95% CI = 60.4–96.6% versus 78.9%; 95% CI = 54.4–93.9%; $p = 0.5$, and 86.7%; 95% CI = 64.0–86.6% versus 76.7%; 95% CI = 75.4–94.1%; $p = 0.1$, respectively).

Discordance between CT and MRI regarding Bosniak classes and imaging features

The data analysis regarding the Bosniak classes indicated discordance between CT and MRI data in 50% (30/60) CRMs. Of these, 53.3% (16/30) were RCC. In comparison to CT, MRI produced 36.7% (22/60) upgrading of the CRMs (6.7% (4/60) in class I, 11.7% (7/60) in class II, 8.3% (5/60) in class IIF, and 10% (6/60) in class III) and 13.3% (8/60) downgrading of the CRMs (1.7% (1/60) in class II, 1.7% (1/60) in class IIF, and 10% (6/60) in class IV). The change

in individual CRM classification on MRI compared to CT is presented in Table 4. Four CRMs were classified as Bosniak I on CT and were upgraded to Bosniak II and IIF on MRI. The increasing number and thickness of septa were the main causes for the upgrade by MRI. Eight CRMs were classified as Bosniak II on CT. Seven out of eight CRMs were upgraded to Bosniak IIF and IV on MRI. The increasing number and thickness of septa in five CRMs and the detection of enhanced nodules in two CRMs were the causes for the upgrade by MRI. Seventeen CRMs were classified as Bosniak III on CT. Six out of 17 CRMs were upgraded to Bosniak IV on MRI. The detection of enhanced nodules was the reason for the upgrade by MRI.

From Table 2, we did not find significant differences between CT and MRI regarding the irregularity (66.7%, 95% CI = 38.4–88.2 versus 78.6%, 95% CI = 49.2–95.3%, $p = 0.7$), the wall/septa thickness (53.3%, 95% CI = 26.6–78.7 versus 57.1%, 95% CI = 28.9–82.3%, $p = 0.8$), the nodules with acute angle (75%, 95% CI = 47.6–92.7% versus 68.2%,

Table 4 BCv2019 for 60 CRMs by CT and MRI with change in individual CRM classification on account of MRI, compared to CT

CT-based Bosniak classification	MRI-based Bosniak classification					Sum of CT
	I	II	IIF	III	IV	
I	0	3 (5)	1 (1.6)	0	0	4 (6.7)
II	1 (1.6)	0	5 (8.3)	0	2 (3.3)	8 (13.3)
IIF	0	1 (1.6)	7 (11.7)	2 (3.3)	3 (5)	13 (21.7)
III	0	0	0	11 (18.3)	6 (10)	17 (28.3)
IV	0	0	3 (5)	3 (5)	12 (20)	18 (30)
Sum of MRI	1 (1.6)	4 (6.7)	16 (26.7)	16 (26.7)	23 (38.3)	60 (100)

Data are the number of CRMs based on consensus reviewing with the percentage in parentheses. *BCv2019*, Bosniak classification version 2019; *CRMs*, cystic renal masses; *CT*, computed tomography; *MRI*, magnetic resonance imaging. Yellow cells indicate agreement in class in the same CRM evaluated with CT and MRI. Green cells indicate instances of MRI upgrade in class compared to CT in the same CRM. Blue cells indicate instances of MRI downgrade in class compared to CT in the same CRM

95% CI = 45.1–86.1%, $p = 0.9$), and the nodules > 4 mm with obtuse angle (25%, 95% CI = 7.3–52.4% versus 31.8%, 95% CI = 13.9–54.9%, $p = 0.9$).

Inter-reader agreement for imaging features of CRMs and BCv2019 within CT and MRI

The inter-reader agreement (IRA) in CT and MRI was moderate to substantial for evaluating imaging features of CRMs and substantial for determining the BCv2019 class ($k = 0.66$ and 0.62 , respectively) (Table 5).

Inter-modality agreement of BCv2019 between CT and MRI

The inter-modality agreement between CT and MR was moderate to substantial for evaluating imaging features of CRMs (Table 5). The inter-modality agreement of BCv2019 classes was moderate ($\kappa = 0.58$, 95% CI = 0.45–0.72) between CT and MRI.

Figures 2, 3, and 4 illustrate examples of our cases.

Discussion

Since its development, many modifications have been performed in order to improve the Bosniak classification (BC) for better classification of CRMs and diagnosis of cystic RCC. Using the latest update of the BC (BCv2019), our study

demonstrated that there is no significant difference in sensitivity, specificity, and accuracy of MRI and CT in the diagnosis of cystic RCC (87.8% versus 75.6%, 84.2% versus 78.9%, and 86.7% versus 76.7%, respectively), which was comparable to the previous research findings [16–21].

Determining the true prevalence of malignancy in many Bosniak classes is problematic [13]. Our study found that the proportions of malignancy in Bosniak IIF, III, and IV were 61.5%, 88.2%, and 88.9%, respectively, according to CT and 31.3%, 87.5%, and 95.7%, respectively, according to MRI. Our results are in line with previous reports [13, 18]. Silverman et al [13] observed a wide range of reported malignancy rates: 0–38% in Bosniak IIF, 25–100% in Bosniak III (approximately 50%), and 56–100% in Bosniak IV (approximately 90%). Sevcenco et al [18] reported malignancy rate pooled estimates of 6.7% in Bosniak IIF, 55.1% in Bosniak III, and 91% in Bosniak IV. An interesting finding in our study was the big difference in the percentage of malignancy in Bosniak IIF between CT and MRI (61.5% versus 31.3%). This may be attributed to the better detection of the irregularity and enhanced nodules by the MRI, which resulted in an upgrade of Bosniak IIF CRMs. This finding is consistent with the finding of Park et al [17], who reported that CRMs upgraded on MRI relative to CT using BCv2019 had a high malignancy rate, supporting the role of MRI for further characterization of Bosniak IIF CRMs on CT.

Our analysis supports the findings of Tse et al [15] and Chan et al [16], who found that the same patient with CRMs might be classified differently by Bosniak with CT and

Table 5 Inter-reader and inter-modality agreement for imaging features of CRMs and BCv2019 within CT and MRI

Imaging features	Inter-reader agreement		Inter-modality agreement	<i>p</i> value
	CT	MRI		
Wall thickness	0.51 (0.26–0.68)	0.57 (0.37–0.73)	0.63 (0.48–0.77)	0.38
Wall irregularity/protrusion	0.61 (0.42–0.75)	0.69 (0.54–0.80)	0.62 (0.35–0.89)	0.25
Wall enhancement	0.51 (0.26–0.68)	0.49 (0.24–0.67)	0.62 (0.34–0.89)	0.013*
Septa number	0.53 (0.30–0.70)	0.63 (0.44–0.76)	0.43 (0.26–0.60)	< 0.0001*
Septa thickness	0.55 (0.32–0.71)	0.54 (0.31–0.70)	0.44 (0.28–0.59)	0.24
Septa irregularity/protrusion	0.64 (0.45–0.76)	0.71 (0.56–0.81)	0.52 (0.31–0.73)	0.8
Septa enhancement	0.45 (0.18–0.64)	0.59 (0.38–0.73)	0.54 (0.28–0.81)	0.004*
Nodules	0.75 (0.63–0.84)	0.77 (0.66–0.85)	0.52 (0.29–0.74)	0.78
Calcifications	0.79 (0.68–0.86)	0.73 (0.59–0.82)	0.46 (0.30–0.62)	< 0.0001*
Heterogeneity	0.72 (0.58–0.82)	0.76 (0.63–0.84)	0.57 (0.38–0.77)	0.11
BCv2019	0.66 (0.54–0.76)	0.62 (0.50–0.73)	0.58 (0.45–0.72)	0.02*

Data are k coefficients and numbers in parentheses are 95% confidence intervals. *BCv2019*, Bosniak classification version 2019; *CRMs*, cystic renal masses; *CT*, computed tomography; *MR*, magnetic resonance imaging; * = significant. The p value is related to the comparison of CT to MRI. The inter-reader agreement was calculated using the intraclass correlation coefficient (ICC) with 95% confidence intervals. The inter-modality agreement between CT and MRI was calculated using the weighted kappa (κ) statistics with 95% confidence intervals. The k values were interpreted as follows: 0.00–0.20 = poor agreement; 0.21–0.40 = fair agreement; 0.41–0.60 = moderate agreement; 0.61–0.80 = substantial agreement; and 0.81–1.00 = almost perfect agreement

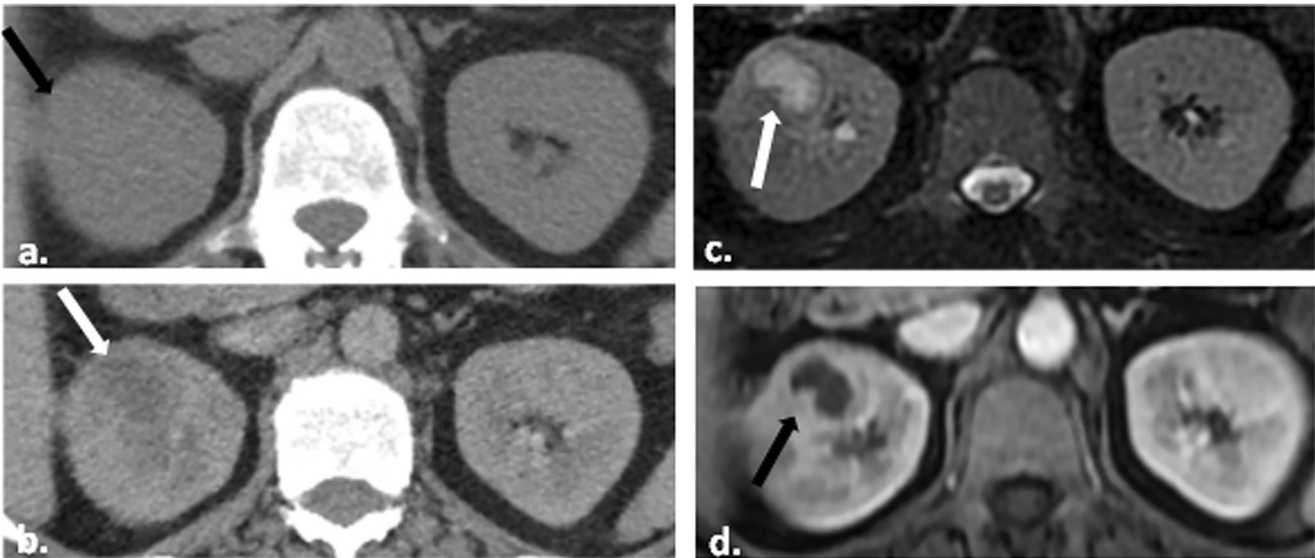


Fig. 2 A 51-year-old male patient with right CRM. **a** Axial unenhanced CT image shows an ill-defined hypodense mass (arrow). **b** Axial contrast-enhanced CT image (nephrographic phase) shows a homogeneous non-enhancing mass (21 HU) (arrow). **c** Axial T2-weighted fat-saturated MR image shows a well-defined hyperintense mass with a small isointense

nodule at its posterolateral wall (arrow). **d** Axial contrast-enhanced fat-saturated T1-weighted MR image shows enhanced nodule (arrow). The CRM was classified as Bosniak II on CT and Bosniak IV on MRI. The patient underwent total nephrectomy, and histopathology revealed papillary RCC

Bosniak with MRI. In our study, when compared to CT, the MRI produced a 36.7% upgrade and 13.3% downgrade of the CRMs. The number of septa and their thickness and the presence of enhanced nodules were accounted for the majority of

variations in the Bosniak classes between CT and MRI. This disparity between CT and MRI in the assessment of imaging features is most likely owing to the higher contrast resolution of MRI compared to CT. Our findings are consistent with

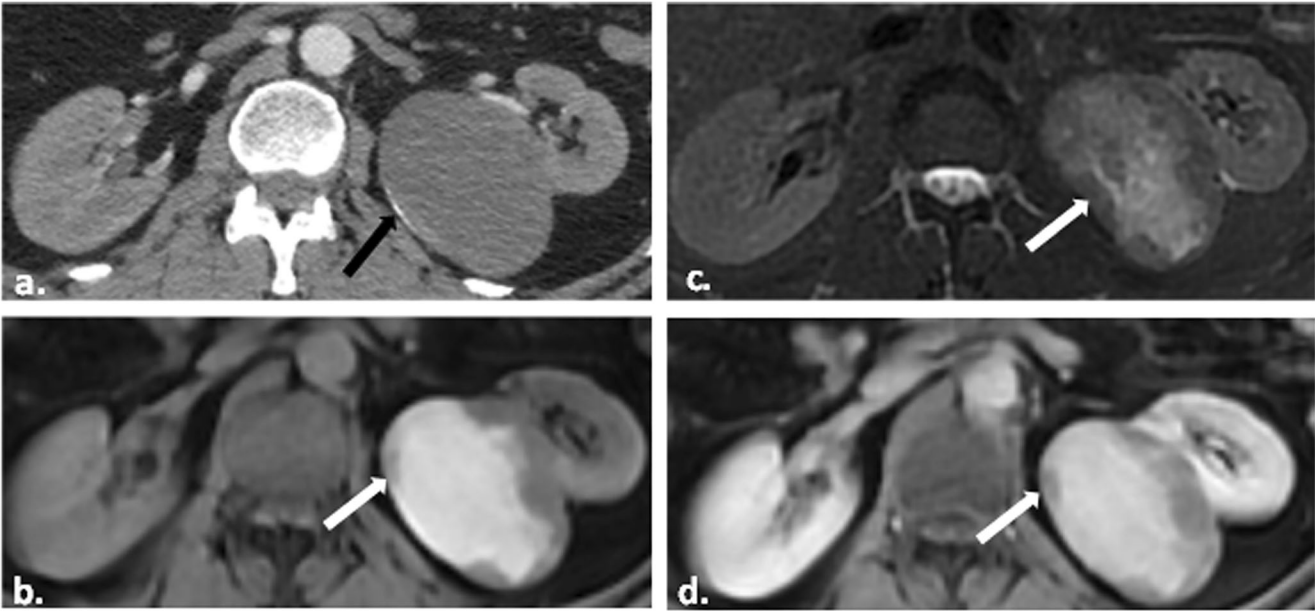


Fig. 3 A 67-year-old male patient with left CRM. **a** Axial contrast-enhanced CT image (nephrographic phase) shows a well-defined homogeneous mass with a peripheral rim of calcification (arrow). **b** Axial fat-saturated T1-weighted MR image shows a well-defined heterogeneous mass with high signal intensity central and isointense peripheral nodules (arrow). **c** Axial fat-saturated T2-weighted MR image shows the mass

with high signal intensity central and iso to low signal intensity peripheral nodules (arrow). **d** Axial contrast-enhanced fat-saturated T1-weighted MR image shows mildly enhanced nodules (arrow). The CRM was classified as Bosniak II on CT and Bosniak IV on MRI. The patient underwent total nephrectomy, and histopathology revealed hemorrhagic papillary RCC

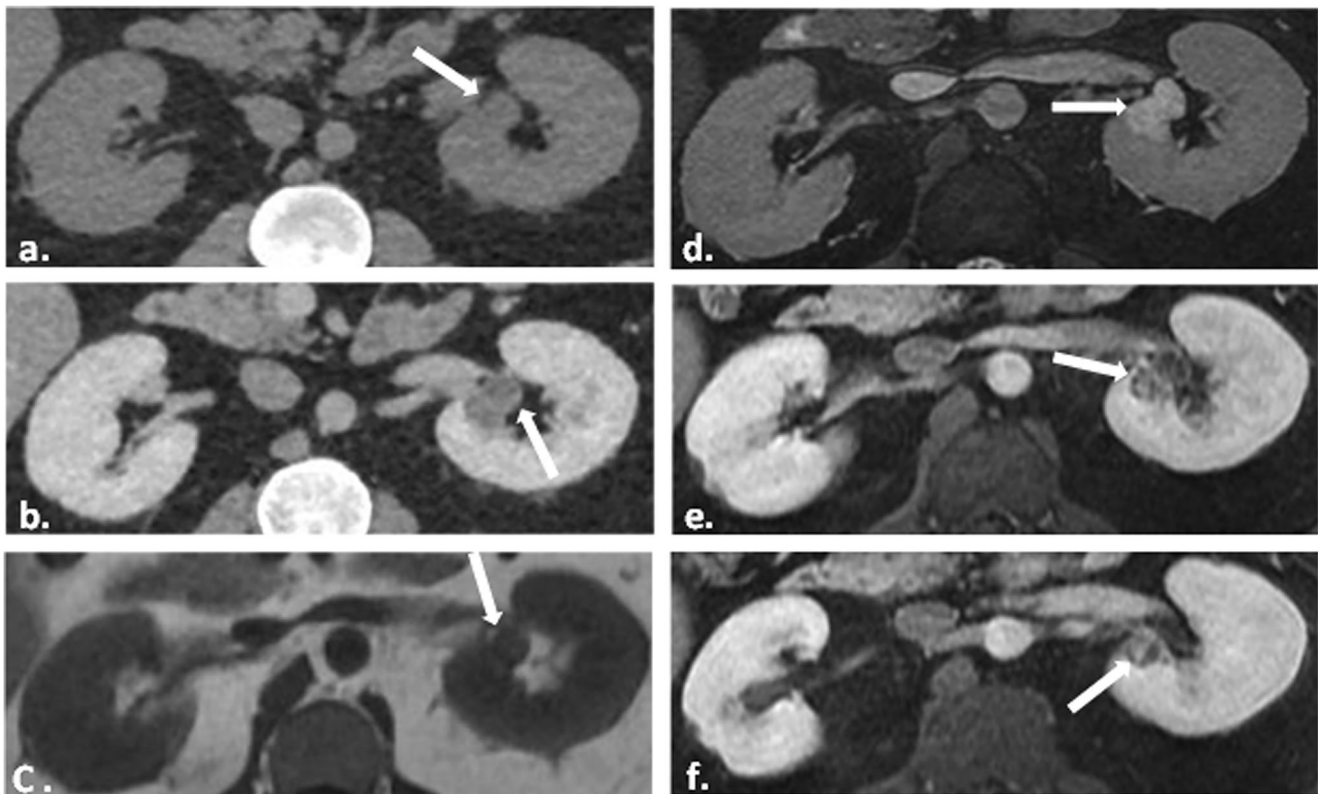


Fig. 4 A 57-year-old female patient with left CRM. **a** Axial unenhanced CT image shows a hypodense mass (arrow). **b** Axial contrast-enhanced CT image (nephrographic phase) shows an irregular thickened enhanced wall and septa (arrow). **c** Axial fat-saturated T1-weighted MR image shows iso to low signal intensity mass (arrow). **d** Axial fat-saturated T2-weighted MR image shows high signal intensity mass (arrow). **e**

and **f** Axial contrast-enhanced fat-saturated T1-weighted MR images show irregular thickened enhanced wall and septa with enhanced nodule (arrow). The CRM was classified as Bosniak III on CT and Bosniak IV on MRI. The patient underwent total nephrectomy and histopathology revealed clear cell RCC

those of previous studies [11, 12, 16, 22–29] showing a better depiction of the septa, protrusions, and enhancement with MRI than CT. This may cause CRMs to be classified in a higher class with MRI than CT.

Recent research published by Tse et al [15] found that MRI and CT both resulted in higher classes in relatively equal frequency, most commonly due to protrusions identified by one imaging modality but not the other. The authors of this research concluded that MRI and CT were concordant in the majority of cases with no statistically significant systematic category upgrade in either modality. Similar findings were reported by Chan et al [16] in this regard; however, the results were somewhat different. Our results are in line with these findings and demonstrated no significant differences between CT and MRI regarding irregularity, wall/septa thickness, nodules with acute angles, and nodules > 4 mm with obtuse angles.

Although limited research has compared CT to MRI in Bosniak classification [15–17, 22], our study is a prospective one in which we assessed the diagnostic accuracy and agreement of CT and MRI in BCv2019. We found moderate to substantial agreement between the readers for the proposed

imaging features and determination of Bosniak classes in CT and MRI. This result was comparable to the results of preliminary studies [16, 19, 21, 30–32], which evaluated inter-observer agreement using BCv2019 and showed moderate (MRI) to substantial (CT and MRI) agreement between observers. Our results also showed a moderate inter-modality agreement of the Bosniak classes between CT and MRI. Park et al [17] reported substantial inter-modality agreement between CT and MRI.

Finally, in keeping with our results, several authors [18] have tested BC and considered this system a valuable classification system of CRMs. Moreover, some authors [10, 15–17, 19, 33] are trying to enhance the diagnostic power of BC by adding MRI and contrast-enhanced ultrasound. Therefore, the BCv2019 can stand alone and works like RADS as a universal system that helps the clinician to go from one imaging technique such as ultrasound to CT and MRI. However, the BCv2019 still needs some improvement to become more specific and comprehensive in all pertinent descriptors and definitions.

Our research has limitations. First, the sample size is still small. Second, the calculated diagnostic accuracy of

BCv2019 was restricted to CRMs with pathological confirmation, resulting in unavoidable selection bias toward a higher percentage of Bosniak IIF–IV CRMs and malignant results, and subsequently impacted the values of diagnostic accuracy. Third, we used FNAB and core biopsy to confirm the diagnosis in 12 CRMs (CRMs that were classified as Bosniak I and II and are unlikely to be surgically resected). The need for tissue confirmation in all CRMs was likely biased the patient population toward fewer Bosniak classes I and II than would be anticipated in the general patient population. Fourth, the order of image analysis may create a selection bias and could impact results. However, we separated each set by 1 month to diminish the memory bias of readers. Finally, all analyses were performed using BCv2019 and did not compare with BCv2005

In conclusion, this research confirms the presence of disparity between CT and MRI in the classification of CRMs in terms of BCv2019; however, both imaging modalities are comparable with no significant difference in diagnostic accuracy and reliability.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00330-022-09044-3>.

Acknowledgements The authors of this research would like to express their gratitude to the Ministry of Education and the Deanship of Scientific Research, Najran University. They also thank the Kingdom of Saudi Arabia for its financial and technical support under code number NU/NRP/MRC/11/21.

Funding This study has received funding from the Deanship of Scientific Research, Najran University, Kingdom of Saudi Arabia, grant number NU/NRP/MRC/11/21.

Declarations

Guarantor The scientific guarantor of this publication is Mohammad Abd Alkhalik Basha.

Conflict of interest The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry One of the authors has significant statistical expertise.

Informed consent Written informed consent was obtained from all subjects (patients) in this study.

Ethical approval Institutional Review Board approval was obtained.

Methodology

- prospective
- diagnostic or prognostic study
- multi-center study

References

1. Herts BR, Silverman SG, Hindman NM et al (2018) Management of the incidental renal mass on CT: a white paper of the ACR Incidental Findings Committee. *J Am Coll Radiol* 15(2):264–273
2. Wood CG III, Stromberg LJ III, Harmath CB et al (2015) CT and MR imaging for evaluation of cystic renal lesions and diseases. *Radiographics* 35(1):125–141
3. Silverman SG, Israel GM, Herts BR, Richie JP (2008) Management of the incidental renal mass. *Radiology* 249(1):16–31
4. Chandrasekar T, Ahmad AE, Fadaak K et al (2018) Natural history of complex renal cysts: clinical evidence supporting active surveillance. *J Urol* 99(3):633–640
5. Nicolau C, Antunes N, Paño B, Sebastia C (2021) Imaging characterization of renal masses. *Medicina (Kaunas)* 57(1):51
6. Hartman DS, Choyke PL, Hartman MS (2004) From the RSNA refresher courses: a practical approach to the cystic renal mass. *Radiographics* 24(suppl_1):S101–S115
7. Hélénon O, Crosnier A, Verkarre V, Merran S, Méjean A, Correas JM (2018) Simple and complex renal cysts in adults: classification system for renal cystic masses. *Diagn Interv Imaging* 99(4):189–218
8. Tsili AC, Andriotis E, Gkeli MG et al (2021) The role of imaging in the management of renal masses. *Eur J Radiol* 141:109777
9. Bosniak MA (1968) The current radiological approach to renal cysts. *Radiology* 158(1):1–10
10. Israel GM, Hindman N, Bosniak MA (2004) Evaluation of cystic renal masses: comparison of CT and MR imaging by using the Bosniak classification system. *Radiology* 231(2):365–371
11. Israel GM, Bosniak MA (2005) An update of the Bosniak renal cyst classification system. *Urology* 66(3):484–488
12. Bosniak MA (2012) The Bosniak renal cyst classification: 25 years later. *Radiology* 262(3):781–785
13. Silverman SG, Pedrosa I, Ellis JH et al (2019) Bosniak classification of cystic renal masses, version 2019: an update proposal and needs assessment. *Radiology* 292(2):475–488
14. Schieda N, Davenport MS, Krishna S et al (2021) Bosniak classification of cystic renal masses, version 2019: a pictorial guide to clinical use. *Radiographics* 41(3):814–828
15. Tse JR, Shen J, Shen L, Yoon L, Kamaya A (2021) Bosniak classification of cystic renal masses version 2019: comparison of categorization using CT and MRI. *AJR Am J Roentgenol* 216(2):412–420
16. Chan J, Yan JH, Munir J et al (2021) Comparison of Bosniak classification of cystic renal masses version 2019 assessed by CT and MRI. *Abdom Radiol (NY)* 46(11):5268–5276
17. Park MY, Park KJ, Kim MH, Kim JK (2021) Bosniak classification of cystic renal masses version 2019: comparison to version 2005 for class distribution, diagnostic performance, and interreader agreement using CT and MRI. *AJR Am J Roentgenol* 217(6):1367–1376
18. Sevenco S, Spick C, Helbich TH et al (2017) Malignancy rates and diagnostic performance of the Bosniak classification for the diagnosis of cystic renal lesions in computed tomography—a systematic review and meta-analysis. *Eur Radiol* 27(6):2239–2247
19. Tse JR, Shen J, Yoon L, Kamaya A (2020) Bosniak classification version 2019 of cystic renal masses assessed with MRI. *AJR Am J Roentgenol* 215(2):413–419
20. Ziani I, Ibrahim A, Dergamoun H et al (2020) Atypical renal cysts: is MRI a compulsory examination in 2020 before therapeutic management? *Open J Urol* 10(10):253
21. Bai X, Sun SM, Xu W et al (2020) MRI-based Bosniak Classification of cystic renal masses, version 2019: interobserver agreement, impact of readers' experience, and diagnostic performance. *Radiology* 297(3):597–605

22. Ferreira AM, Reis RB, Kajiwarra PP et al (2016) MRI evaluation of complex renal cysts using the Bosniak classification: a comparison to CT. *Abdom Radiol (NY)* 41(10):2011–2019
23. Kim WB, Lee SW, Doo SW et al (2012) Category migration of renal cystic masses with use of gadolinium-enhanced magnetic resonance imaging. *Korean J Urol* 53(8):573–576
24. Graumann O, Osther SS, Karstoft J, Hørlyck A, Osther PJ (2016) Bosniak classification system: a prospective comparison of CT, contrast-enhanced US, and MR for categorizing complex renal cystic masses. *Acta Radiol* 57(11):1409–1417
25. Krishna S, Schieda N, Pedrosa I et al (2021) Update on MRI of cystic renal masses including Bosniak version 2019. *J Magn Reson Imaging* 54(2):341–356
26. Israel GM, Bosniak MA (2008) Pitfalls in renal mass evaluation and how to avoid them. *Radiographics* 28(5):1325–1338
27. Israel GM, Bosniak MA (2005) How I do it: evaluating renal masses. *Radiology* 236(2):441–450
28. Zhong J, Cao F, Guan X, Chen J, Ding Z, Zhang M (2017) Renal cyst masses (Bosniak category II-III) may be over evaluated by the Bosniak criteria based on MR findings. *Medicine (Baltimore)* 96(51):e9361
29. Vogel DW, Kiss B, Heverhagen JT et al (2021) Prospective comparison of contrast-enhanced ultrasound and magnetic resonance imaging to computer tomography for the evaluation of complex cystic renal lesions. *Urology* 154:320–325
30. Sun M, Wang C, Jiang F, Fang X, Guo B (2019) Diagnostic value and clinical significance of ultrasound combined with CT in cystic renal cell carcinoma. *Oncol Lett* 18(2):1395–1401
31. Shampain KL, Shankar PR, Troost JP et al (2021) Interrater agreement of Bosniak classification version 2019 and version 2005 for cystic renal masses at CT and MRI. *Radiology* 302(2):357–366
32. Lucocq J, Pillai S, Oparka R, Nabi G (2021) Complex renal cysts (Bosniak \geq IIF): interobserver agreement, progression and malignancy rates. *Eur Radiol* 31(2):901–908
33. Barr RG, Peterson C, Hindi A (2014) Evaluation of indeterminate renal masses with contrast-enhanced US: a diagnostic performance study. *Radiology* 271(1):133–142

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.