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

Yassir Edrees Almalki 1 • Mohammad Abd Alkhaliq Basha 2 • Rania Refaat 3 • Sharifa Khalid Alduraibi 8 • Ahmed A. El-Hamid M. Abdalla 2 • Hala Y. Youssef 2 • Mohamed M. A. Zaitoun 2 • Saeed Bakry Elsayed 2 • Nader E. M. Mahmoud 2 • Nader Ali Alayouty 2 • Susan Adil Ali 3 • Ahmad Abdullah Alnaggar 2 • Sameh Saber 2 • Ahmed Mohamed El-Maghraby 2 • Amgad M. Elsheikh 2 • Mohamed Hesham Saleh Saleh Radwan 2 • Ahmed Gamil Ibrahim Abdelmegid 2 • Sameh Abdelaziz Aly 5 • Waleed S. Abo Shanab 6 • Ahmed Ali Obaya 7 • Shaimaa Farouk Abdelhai 7 • Shereen Elshorbagy 8 • Yasser M. Haggag 9 • Hwaide M. Mokhtar 10 • Nesreen M. Sabry 11 • Jehan Ibrahim Althohamy 12 • Rasha Taha Abouelkheir 13 • Tawfik Omran 14 • Ahmed Shalan 5 • Youssef H. Algazzar 15 • Maha Ibrahim Metwally 2

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 Alkhaliq 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

Published online: 13 August 2022
(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 BCv219 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 (46.3%) 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);
The radiologists then used the BCv2019 algorithm 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 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 ≤ 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>
<thead>
<tr>
<th>CT Reader 1</th>
<th>Reader 2</th>
<th>Reader 3</th>
<th>Consensus reviewing</th>
<th>Inter-reader agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCC</td>
<td>RCC</td>
<td>RCC</td>
<td>RCC</td>
<td>RCC</td>
</tr>
<tr>
<td>Benign</td>
<td>Benign</td>
<td>Benign</td>
<td>Benign</td>
<td>Benign</td>
</tr>
<tr>
<td>Bosniak I</td>
<td>2 (100)</td>
<td>0</td>
<td>4 (100)</td>
<td>0</td>
</tr>
<tr>
<td>Bosniak II</td>
<td>1 (100)</td>
<td>0</td>
<td>3 (60)</td>
<td>2 (40)</td>
</tr>
<tr>
<td>Bosniak IIF</td>
<td>9 (34.6)</td>
<td>17 (65.4)</td>
<td>6 (50)</td>
<td>6 (50)</td>
</tr>
<tr>
<td>Bosniak III</td>
<td>4 (21.1)</td>
<td>15 (78.9)</td>
<td>2 (13.3)</td>
<td>13 (86.7)</td>
</tr>
<tr>
<td>Irregularity</td>
<td>2 (50)</td>
<td>8 (53.3)</td>
<td>0</td>
<td>8 (61.5)</td>
</tr>
<tr>
<td>Wall/Septa thick</td>
<td>3 (75)</td>
<td>9 (60)</td>
<td>2 (100)</td>
<td>6 (46.2)</td>
</tr>
<tr>
<td>Bosniak IV</td>
<td>3 (25)</td>
<td>9 (66.7)</td>
<td>4 (16.7)</td>
<td>20 (83.3)</td>
</tr>
<tr>
<td>Nodule with acute Angle</td>
<td>2 (66.7)</td>
<td>7 (77.8)</td>
<td>2 (50)</td>
<td>15 (75)</td>
</tr>
<tr>
<td>Nodule &gt; 4 mm with Obtuse angle</td>
<td>1 (33.3)</td>
<td>2 (22.2)</td>
<td>2 (50)</td>
<td>5 (25)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MRI Reader 1</th>
<th>Reader 2</th>
<th>Reader 3</th>
<th>Consensus reading</th>
<th>Inter-reader agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCC</td>
<td>RCC</td>
<td>RCC</td>
<td>RCC</td>
<td>RCC</td>
</tr>
<tr>
<td>Benign</td>
<td>Benign</td>
<td>Benign</td>
<td>Benign</td>
<td>Benign</td>
</tr>
<tr>
<td>Bosniak I</td>
<td>2 (100)</td>
<td>0</td>
<td>4 (100)</td>
<td>0</td>
</tr>
<tr>
<td>Bosniak II</td>
<td>7 (87.5)</td>
<td>1 (12.5)</td>
<td>1 (50)</td>
<td>1 (50)</td>
</tr>
<tr>
<td>Bosniak IIF</td>
<td>4 (80)</td>
<td>1 (20)</td>
<td>7 (43.8)</td>
<td>9 (56.2)</td>
</tr>
<tr>
<td>Bosniak III</td>
<td>3 (13)</td>
<td>20 (87)</td>
<td>1 (8.3)</td>
<td>11 (91.7)</td>
</tr>
<tr>
<td>Irregularity</td>
<td>2 (66.7)</td>
<td>13 (65)</td>
<td>0</td>
<td>7 (63.6)</td>
</tr>
<tr>
<td>Wall/Septa thick</td>
<td>2 (66.7)</td>
<td>10 (50)</td>
<td>1 (100)</td>
<td>6 (54.5)</td>
</tr>
<tr>
<td>Bosniak IV</td>
<td>3 (13.6)</td>
<td>19 (86.4)</td>
<td>6 (23.1)</td>
<td>20 (76.9)</td>
</tr>
<tr>
<td>Nodule with acute Angle</td>
<td>2 (66.7)</td>
<td>1 2 (63.2)</td>
<td>4 (66.7)</td>
<td>13 (65)</td>
</tr>
<tr>
<td>Nodule &gt; 4 mm with Obtuse angle</td>
<td>1 (33.3)</td>
<td>7 (36.8)</td>
<td>2 (33.3)</td>
<td>7 (35)</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Reader 1</th>
<th>Reader 2</th>
<th>Reader 3</th>
<th>Consensus reviewing</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT</td>
<td>MRI</td>
<td>CT</td>
<td>MRI</td>
<td></td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td></td>
<td>58.5 (24/41)</td>
<td>95.1 (39/41)</td>
<td>80.5 (33/41)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[42.1–73.7]</td>
<td>[83.5–99.4]</td>
<td>[65.1–91.2]</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td></td>
<td>63.2 (12/19)</td>
<td>68.4 (13/19)</td>
<td>63.2 (12/19)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[38.4–83.7]</td>
<td>[43.5–87.4]</td>
<td>[38.4–83.7]</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td></td>
<td>60.0 (36/60)</td>
<td>86.7 (52/60)</td>
<td>76.7 (46/60)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[46.5–72.4]</td>
<td>[75.4–94.1]</td>
<td>[64.0–86.6]</td>
</tr>
</tbody>
</table>

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.

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

<table>
<thead>
<tr>
<th>CT-based Bosniak classification</th>
<th>MRI-based Bosniak classification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
</tr>
<tr>
<td>I</td>
<td>0</td>
</tr>
<tr>
<td>II</td>
<td>1 (1.6)</td>
</tr>
<tr>
<td>IIF</td>
<td>0</td>
</tr>
<tr>
<td>III</td>
<td>0</td>
</tr>
<tr>
<td>IV</td>
<td>0</td>
</tr>
<tr>
<td>Sum of MRI</td>
<td>1 (1.6)</td>
</tr>
</tbody>
</table>

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

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

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 (\( \kappa \)) 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.
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. 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.

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.
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

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.
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 Alkhalkal 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

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

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