EUROPEAN SOCIETY OF RADIOLOGY
European Radiology

UROGENITAL

Ovarian-Adnexal Imaging-Reporting and Data System (O-RADS) ultrasound version 2019: a prospective validation and comparison to updated version (v2022) in pathologically confirmed adnexal masses

Yassir Edrees Almalki[1], Mohammad Abd Alkhalik Basha[2*], Mohamad Gamal Nada[2], Maha Ibrahim Metwally[2], Yasmin Ibrahim Libda[2], Noha Yahia Ebaid[2], Mohamed M. A. Zaitoun[2], Nader E. M. Mahmoud[2], Amgad M. Elsheikh[2], Mohamed Hesham Saleh Saleh Radwan[2], Mohamed I. Amin[2], Elshaimaa Mohamed Mohamed[2], Engy Fathy Tantawy[2], Sameh Saber[2], Walid Mosallam[3], Housseini Mohamed Abdalla[3], Mohamed Abd El-Aziz Mohamed Farag[4], Tamer Mahmoud Dawoud[5], Hamada M. Khater[6], Diaa Bakry Eldib[6], Jehan Ibrahim Altohamy[7], Rasha Taha Abouelkheir[8], Waseem M. El Gendy[9], Sharifa Khalid Alduraibi[10], Majed Saeed Alshahrani[11], Safaa A. Ibrahim[12], Ahmed M. Radwan[12], Ahmed Ali Obaya[13], Ahmed M. Abdelkhalik Basha[14] and Ahmed Mohamed El-Maghraby[2]

Abstract

Objective To evaluate the diagnostic accuracy and reliability of the Ovarian-Adnexal Reporting and Data System (O-RADS) ultrasound v2019 in classifying adnexal masses (AMs) and compare the old and updated systems (v2022).

Patients and methods This prospective study enrolled 977 consecutive women with suspected AMs from three institutions between January 2022 and December 2023. Ultrasound examinations were performed by three experienced radiologists who categorized AMs according to O-RADS ultrasound v2019. The same radiologists retrospectively reviewed the stored ultrasound images and provided the O-RADS ultrasound v2022 classification. Histopathology was used as the reference standard to calculate the diagnostic accuracy of the O-RADS versions in predicting malignant AMs. Inter-observer agreement (IOA) of the O-RADS scoring results was evaluated using the Fleiss kappa (κ) test.

Results The final analysis included 803 women with 855 AMs (219 (25.6%) malignant and 636 (74.4%) benign). Both O-RADS versions demonstrated good diagnostic accuracy, with area under the curve (AUC) values ranging from 0.906 to 0.923 (v2019) and 0.919 to 0.936 (v2022). The updated v2022 showed a slightly higher accuracy (82.5–86.7% vs. 80.7–85.3%), sensitivity (93.6–95.0% vs. 92.2–94.1%), and specificity (78.1–84.1% vs. 76.1–82.9%) compared to v2019. The IOA for the overall O-RADS classification was perfect for both versions ($\kappa = 0.96$–0.97).

Conclusions The O-RADS ultrasound classification system demonstrated good diagnostic accuracy and reliability in predicting malignant AMs, with the updated v2022 showing modest improvements.

*Correspondence:
Mohammad Abd Alkhalik Basha
Mohammad_basha76@yahoo.com
Full list of author information is available at the end of the article

Springer

Key Points

Question *Accurate classification of adnexal masses is essential for management. Can updated O-RADS ultrasound v2022 improve diagnostic accuracy and reliability compared to v2019 in predicting malignancies?*

Findings *O-RADS ultrasound v2022 demonstrated slightly higher diagnostic accuracy for identifying malignant adnexal masses compared to v2019, reflecting modest improvements in risk stratification and clinical decision-making.*

Clinical relevance *The updated O-RADS ultrasound v2022 provides improved risk stratification for adnexal masses, enhancing diagnostic confidence, supporting more precise clinical decision-making, and improving patient outcomes through timely intervention or tailored management strategies in ovarian cancer care.*

Keywords Adnexal diseases, Neoplasms, Gynecology, Ultrasonography

Introduction

Since 2010, several ultrasound-based classification systems have emerged to characterize adnexal masses (AMs). These include the International Ovarian Tumor Analysis (IOTA) simple rules, the Society of Radiologists in Ultrasound (SRU) guidelines, and the Gynecologic Imaging-Reporting and Data System (GI-RADS) [1–5].

In 2016, the American College of Obstetricians and Gynecologists endorsed the use of ultrasound to evaluate AMs, noting that the presence of malignant features based on the IOTA simple rules should increase suspicion of ovarian cancer [6, 7]. In 2018, the American College of Radiology (ACR) established the Ovarian-Adnexal Reporting and Data System (O-RADS) ultrasound, based on IOTA descriptors with North American modifications, to guide risk stratification and management of AMs. O-RADS provides a standardized lexicon and categorizes AMs into six risk categories with corresponding management guidelines [8, 9]. In 2022, an updated version of O-RADS was developed. This version introduced new lexicon descriptors to improve risk stratification accuracy, address deficiencies, and aid in appropriate evaluation, surveillance, or intervention decisions, reflecting ongoing efforts to optimize its clinical utility [10].

Since the initial publication of O-RADS ultrasound in 2019, several validation studies [11–21] have been conducted to assess its diagnostic accuracy and reliability. These studies consistently demonstrated that O-RADS ultrasound has excellent diagnostic accuracy and reliability. However, despite the widespread use of O-RADS ultrasound in clinical settings and the considerable number of studies conducted to validate its diagnostic efficacy, there is still a need for further research in this area. Therefore, we conducted this prospective multicenter study to assess the diagnostic accuracy and reliability of O-RADS ultrasound v2019 in classifying AMs, using histopathology as the reference standard. Additionally, we retrospectively analyzed the stored ultrasound static images and cine-clips and applied the updated O-RADS ultrasound v2022 classification to assess its diagnostic accuracy and compare it with the old version.

Patients and methods

Ethical statement

This prospective study was conducted following international guidelines approved by the Institutional Review Board (approval number: ZU-10352). Informed consent was obtained from all patients before the study. We adhered to the ethical principles of the Declaration of Helsinki while preparing this study.

Study population

Patient enrollment for this study was conducted from January 2022 to December 2023. During this two-year period, we consecutively enrolled 977 women with suspected AMs from three institutions. Patients were enrolled continuously throughout these 24 months, ensuring a representative sample across different times of the year. The exclusion criteria were as follows: (1) patients categorized as O-RADS 1, such as those with simple or corpus luteum cysts ≤ 3 cm ($n = 64$); (2) patients who did not undergo surgery, such as those with cysts that resolved during follow-up ($n = 73$); (3) patients lost to follow-up ($n = 25$); and (4) pregnant patients at the time of examination ($n = 12$). As a result, the final cohort consisted of 803 women. All participants underwent a comprehensive assessment, including a detailed medical history, a complete physical examination, and an ultrasound examination. The flowchart of the study is shown in Fig. 1.

Ultrasound examination

All ultrasound examinations were performed using the S50 ELITE (Sonoscape), GE Logiq 9 (GE Healthcare), and Philips iU22 (Philips Healthcare) ultrasound machines. Transabdominal and transvaginal (TV) ultrasound scans were performed, with cine-clips and static images stored for each patient. TV ultrasound was conducted using a real-time sector scanner with a high-frequency TV probe (5/7.5, 5/9, and 4/8 MHz). Transabdominal ultrasound scanning was performed using a real-time scanner with a low-frequency probe (1/5 MHz). For TV examination, the patient assumed a

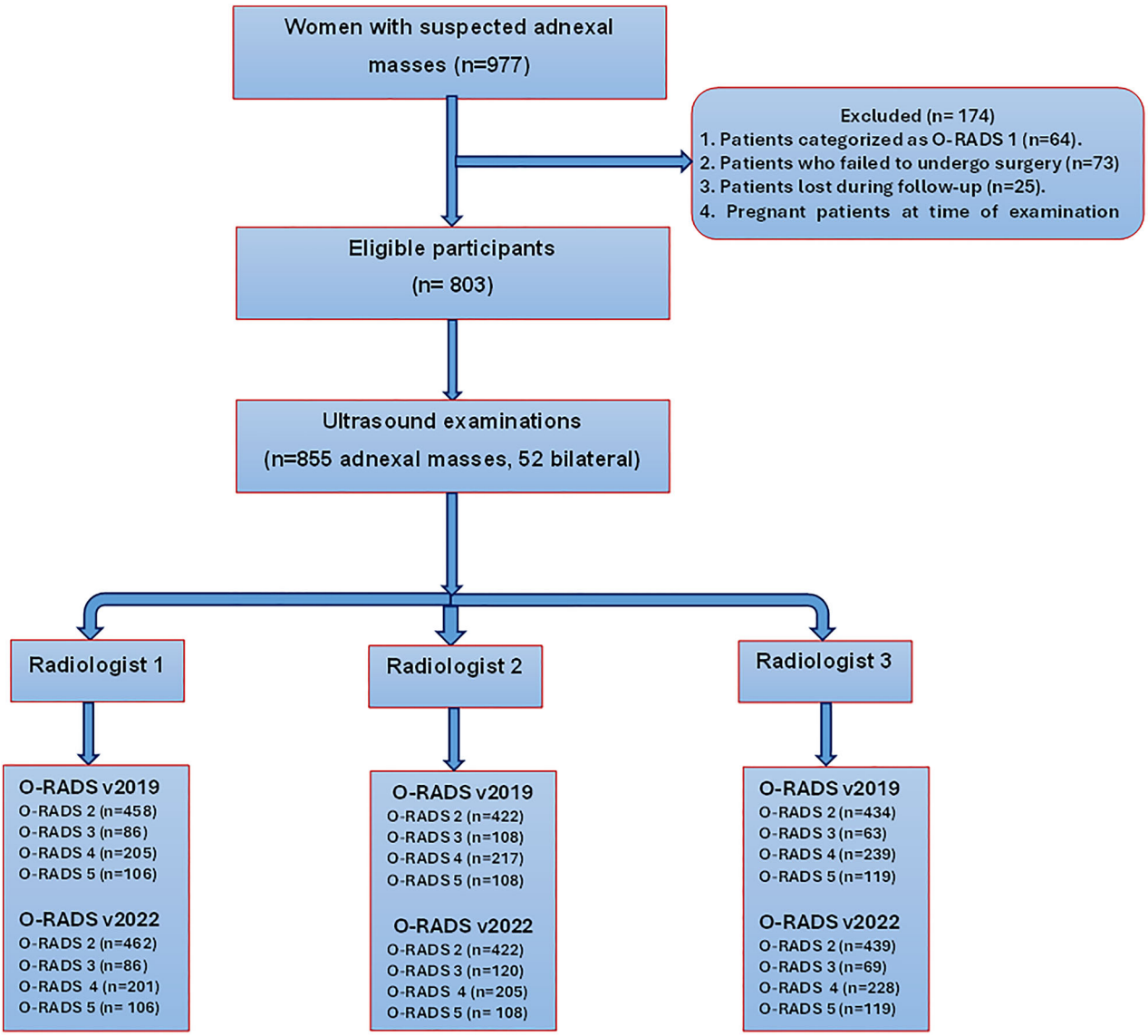

Fig. 1 Flowchart of the study population

lithotomy position after bladder emptying. Radiologists first used B-mode ultrasound to conduct a comprehensive scan and evaluate the lesion. Subsequently, color Doppler flow imaging was performed and the section of the lesion with the highest blood flow was retained. Transabdominal ultrasound was performed for virgo intacta patients ($n = 214$) or for patients with large tumors that could not be fully visualized using the TV approach ($n = 81$). Cine-clips were recorded for each AM to capture the entire lesion by gray-scale and color Doppler modes. This ensured a comprehensive visualization of the masses' morphology and vascularity for later review and analysis.

Examination interpretation

Three highly experienced radiologists (M.A.A.B., H.M.K., and W.M.), each with over 50,000 ultrasound examinations, independently conducted all sonographic examinations while blinded to the patient's clinical data. Before the study commenced, the radiologists received a 5-h lecture and hands-on training session that provided a detailed explanation of the O-RADS. During ultrasound real-time examinations, the radiologists comprehensively evaluated and documented the following imaging features for each AM: Laterality (unilateral or bilateral), maximum diameter, internal echoes and/or incomplete septation, uni/multilocular, typical classic benign ovarian lesions, solid lesions, papillary projections or nodules,

smooth/irregular, acoustic shadowing, ascites and/or peritoneal nodules, and color score. After completing the ultrasound examinations, each radiologist independently assigned an O-RADS category to each AM using the O-RADS ultrasound v2019 criteria [8]. Subsequently, in September 2023, the updated O-RADS ultrasound v2022 was published during the course of this study. At this point, the same three radiologists retrospectively and independently reviewed the stored ultrasound static images and cine-clips for all detected AMs while remaining blinded to the final histopathological diagnosis. Using the updated O-RADS ultrasound v2022 criteria [10], they independently reassigned O-RADS categories to each AM. The time gap between the original ultrasound examination and the retrospective v2022 review ranged from 3 to 18 months, with a mean interval of 8.3 ± 2.9 months.

Reference standard

The final diagnosis of AMs was confirmed through postoperative histopathological examination. A team of specialized gynecological pathologists, who were unaware of the ultrasound findings, reviewed all specimens and reached a consensus on the diagnoses. The AMs were classified based on the histological classification of ovarian tumors by the World Health Organization (WHO) [22]. Borderline AMs were considered to be malignant.

Statistical analysis

Statistical analysis was performed using MedCalc (version 20.022) and SPSS (version 26). Continuous variables were presented as means and standard deviations, while categorical variables were presented as numbers and percentages. To compare categorical variables, we used the chi-square test; for continuous variables, we used the one-way ANOVA test. The receiver operating characteristic (ROC) curve was utilized to identify the best cutoff value and the area under the curve (AUC) for predicting malignant AMs. To evaluate the diagnostic accuracy of both O-RADS versions in categorizing AMs, we employed a four-fold table test with histopathology as the reference standard. We used Fleiss kappa (κ) statistics to assess the inter- and intra-observer agreement of ultrasound imaging features and O-RADS scoring results in predicting AM malignancy. The κ values were interpreted as follows: 0.01–0.20 = poor agreement, 0.21–0.40 = fair agreement, 0.41–0.60 = moderate agreement, 0.61–0.80 = good agreement, and 0.81–1.0 = perfect agreement. Statistical significance was set at $p \leq 0.05$.

Results

Patients and AMs

The study included 803 women with at least one AM detected on ultrasound. Among these patients, 52 (6.5%) had bilateral masses, resulting in 855 AMs examined. Table 1 provides a summary of the clinical and pathological data of both patients and AMs. The mean age at presentation was 39 ± 10.4 years (range, 14–67 years). Of the participants, 259 (32.3%) were nulliparous, and 544 (67.7%) were multiparous. In terms of menopausal status, 564 (70.2%) patients were premenopausal and 239 (29.8%) were postmenopausal. Of the 855 AMs, 219 (25.6%) were malignant, and 636 (74.4%) were benign. The most common benign AM was a hemorrhagic cyst (31.2%), whereas serous cystadenocarcinoma was the most frequently observed malignant AM (31.7%).

Table 1 Clinical-pathologic data of patients and AMs

Characteristic	Value
Total no. of patients	803
Age, years, mean ± SD (range)	39 ± 10.4 (14–67)
Menopausal status	
Premenopausal	564 (70.2)
Postmenopausal	239 (29.8)
Parity	
Multipara	544 (67.7)
Nullipara	259 (32.3)
Laterality of masses	
Unilateral	751 (93.5)
Bilateral	52 (6.5)
Total no. of masses	855
Final diagnosis	
Benign	636 (74.4)
Malignant	219 (25.6)

Unless stated otherwise, data are number of patients or adnexal masses. Data in parentheses are percentages
AMs adnexal masses, *SD* standard deviation

Ultrasound imaging features of AMs

Table 2 provides a detailed overview of the ultrasound imaging features of AMs stratified by radiologists. Significant differences were observed among the radiologists in assessing the following features: classic benign lesions ($p < 0.0001$), solid lesions ($p < 0.0001$), papillary projections/nodules ($p < 0.0001$), smooth/irregular features ($p < 0.0001$), color flow scores ($p < 0.0001$), and internal echoes and/or incomplete septations ($p = 0.002$). However, no significant differences were found among the radiologists when evaluating the maximum diameter ($p = 0.900$), uni/multilocular features ($p = 0.192$), acoustic shadowing ($p = 0.278$), and the presence of ascites and/or peritoneal nodules ($p = 0.332$).

Assignment of O-RADS categories

The frequency distributions of the O-RADS categories stratified by the radiologists, O-RADS versions, and histopathological diagnoses are presented in Table 3. O-RADS 2 was the most common category, constituting

Table 2 Ultrasound imaging features of 855 AMs stratified by radiologists

Parameter	O1	O2	O3	*p*-value
Maximum diameter, cm, mean ± SD (range)	6.9 ± 4.1 (3.0–22.2)	7.1 ± 4.1 (2.1–22.0)	7.2 ± 4.2 (2.1–22.0)	0.900
Internal echoes and/or incomplete septation				0.002
Yes	787 (92.0)	821 (96.0)	804 (94.0)	
No (simple cyst)	68 (8.0)	34 (4.0)	51 (6.0)	
Uni/multilocular				0.192
Unilocular	667 (78.0)	667 (78.0)	641 (75.0)	
Bilocular	63 (7.4)	80 (9.4)	82 (9.6)	
Multilocular	125 (14.6)	108 (12.6)	132 (15.4)	
Typical benign ovarian lesions				< 0.0001
Yes	547 (64.0)	456 (53.3)	433 (50.6)	
No	308 (36.0)	399 (46.7)	422 (49.4)	
Solid lesions				< 0.0001
Yes	143 (16.7)	80 (9.4)	97 (11.3)	
No	712 (83.3)	775 (90.6)	758 (88.7)	
Papillary projections or nodules				< 0.0001
No	701 (82.0)	627 (73.3)	656 (76.7)	
< 4	120 (14.0)	160 (18.7)	171 (20.0)	
≥ 4	34 (4.0)	68 (8.0)	28 (3.3)	
Smooth/irregular				< 0.0001
Smooth	667 (78.0)	587 (68.7)	570 (66.7)	
Irregular	188 (22.0)	268 (31.3)	385 (33.3)	
Acoustic shadowing				0.278
Yes	80 (9.4)	97 (11.3)	80 (9.4)	
No	775 (90.6)	758 (88.7)	775 (90.6)	
Ascites and/or peritoneal nodules				0.332
Yes	28 (3.3)	34 (4.0)	40 (4.7)	
No	827 (96.7)	821 (96.0)	815 (95.3)	
Color score				< 0.0001
1 (none)	616 (72.0)	604 (70.6)	502 (58.7)	
2 (minimal flow)	188 (22.0)	182 (21.3)	228 (26.7)	
3 (moderate flow)	40 (4.7)	58 (6.8)	108 (12.6)	
4 (very-strong flow)	11 (1.3)	11 (1.3)	17 (2.0)	

Unless stated otherwise, data are number of patients or adnexal masses. Data in parentheses are percentages
AMs adnexal masses, *O-RADS* Ovarian-Adnexal Reporting and Data System, *SD* standard deviation, *O* observer

50.8–53.6% (v2019) and 49.4–52.6% (v2022) of AMs, with the vast majority of AMs being benign (98.3–98.9%) across all radiologists and both O-RADS versions. O-RADS 3 was the least common category, constituting 7.4–12.6% (v2019) and 8.1–14.0% (v2022) of AMs with more variability, but still a predominance of benign AMs (88.4–93.3%). In the critical O-RADS 4 category, the updated O-RADS v2022 resulted in higher malignancy rates (43.4–52.2%) compared to the original O-RADS v2019 (42.7–50.2%) across the three radiologists. For O-RADS 5, the malignancy rates ranged from 87.4 to 95.3% across the three radiologists using both versions.

Change in individual lesion categorization on account of O-RADS ultrasound v2022, compared to v2019 stratified by observer

Observer 1 reported upgrading 0.5% (4/855) of AMs (all originally O-RADS 2) and downgrading 0.5% (4/855) of AMs (originally O-RADS 4). Observer 2 documented upgrading 1.4% (12/855) of AMs (all originally O-RADS 3) and downgrading 1.4% (12/855) of AMs (originally O-RADS 4). Observer 3 recorded upgrading in 1.3% (11/855) of AMs (five O-RADS 2 and six O-RADS 3) and downgrading was seen in 1.3% (11/855) of AMs (all originally O-RADS 4) (Table 3). The reclassification of these lesions was attributed to the additional descriptors

Table 3 Frequency distributions of O-RADS categories for 855 AMs stratified by radiologists, O-RADS ultrasound versions, and histopathological diagnosis

O-RADS		O-RADS 2			O-RADS 3			O-RADS 4			O-RADS 5		
		Total	Benign	Malignant	Total	Benign	Malignant	Total	Benign	Malignant	Total	Benign	Malignant
v2019	O1	458 (53.6)	451 (98.5)	7 (1.5)	86 (10.1)	76 (88.4)	10 (11.6)	205 (24.0)	102 (49.8)	103 (50.2)	106 (12.4)	7 (6.6)	99 (93.4)
	O2	422 (49.4)	415 (98.3)	7 (1.7)	108 (12.6)	99 (91.7)	9 (8.3)	217 (25.4)	110 (50.7)	107 (49.3)	108 (12.6)	12 (11.1)	96 (88.9)
	O3	434 (50.8)	428 (98.6)	6 (1.4)	63 (7.4)	56 (88.9)	7 (11.1)	239 (28.0)	137 (57.3)	102 (42.7)	119 (13.9)	15 (12.6)	104 (87.4)
v2022	O1	462 (54.0)	457 (98.9)	5 (1.1)	86 (10.0)	78 (90.7)	8 (9.3)	201 (23.5)	96 (47.8)	105 (52.2)	106 (12.4)	5 (4.7)	101 (95.3)
	O2	422 (49.4)	416 (98.6)	6 (1.4)	120 (14.0)	112 (93.3)	8 (6.7)	205 (24.0)	98 (47.8)	107 (52.2)	108 (12.6)	10 (9.3)	98 (90.7)
	O3	439 (51.3)	433 (98.6)	6 (1.4)	69 (8.1)	64 (92.8)	5 (7.2)	228 (26.7)	129 (56.6)	99 (43.4)	119 (13.9)	10 (8.4)	109 (91.6)

Data are number of adnexal masses with the percentage in parenthesis
O-RADS Ovarian-Adnexal Reporting and Data System, *AMs* adnexal masses, *O* observer

introduced in O-RADS ultrasound v2022, specifically for bilocular cysts and solid/smooth lesions with shadowing.

Diagnostic accuracy of O-RADS ultrasound v2019 and v2022 in predicting malignant AMs

ROC curve analysis (Fig. 2) demonstrated that the best cutoff value for predicting malignant AMs was > O-RADS 3 for all reviewers. AUCs ranged from 0.906 to 0.923 for v2019 and 0.919 to 0.936 for v2022. Table 4 presents a comparative analysis of the diagnostic accuracy of O-RADS v2019 and v2022 in predicting malignant AMs across three radiologists. Both versions demonstrated good accuracy, with values ranging from 80.7 to 85.3% (v2019) and 82.5 to 86.7% (v2022) ($p = 0.051–0.185$). Both versions also exhibited high sensitivity, ranging from 92.2 to 94.1% (v2019) and 93.6 to 95.0% (v2022) ($p = 0.688–0.936$). The specificity values ranged from 76.1 to 82.9% (v2019) and 78.1 to 84.1% (v2022) ($p = 0.035–0.317$).

Inter- and intra-observer agreement for imaging features and O-RADS ultrasound categorization

The inter- and intra-observer agreement for individual imaging features and O-RADS categorization are summarized in Table 5. The inter-observer agreement (IOA) for the overall O-RADS classification was perfect ($\kappa = 0.96–0.97$) for both versions across all radiologists. For most individual imaging features, the agreement was good ($\kappa = 0.61–0.77$). Moderate to good IOA was observed for solid lesions and color score ($\kappa = 0.45–0.66$). The highest agreement was observed for the uni/multilocular features ($\kappa = 0.77–0.79$), and ascites and/or peritoneal nodules ($\kappa = 0.71–076$). The lowest agreement was observed for the solid lesion feature ($\kappa = 0.45–0.66$) and color score ($\kappa = 0.50–0.57$).

The intra-observer agreement shows a high level of consistency for each observer. Observer 1 demonstrated perfect agreement in all parameters ($\kappa = 0.85–0.98$). Observer 2 also exhibited perfect agreement ($\kappa = 0.88–0.97$). Similarly, Observer 3 showed perfect agreement in all parameters ($\kappa = 0.87–0.96$). Overall, the O-RADS classification displayed consistently perfect intra-observer agreement between the two versions: Observer 1 ($\kappa = 0.99$, 95% CI = 0.98–1.0), observer 2 ($\kappa = 0.98$, 95% CI = 0.96–1.0), and observer 3 ($\kappa = 0.98$, 95% CI = 0.96–1.0).

Representative cases in our study are shown in Figs. 3–6.

Discussion

Numerous studies have highlighted the importance of the O-RADS ultrasound classification in accurately characterizing and stratifying the risk of AMs [11–20]. These

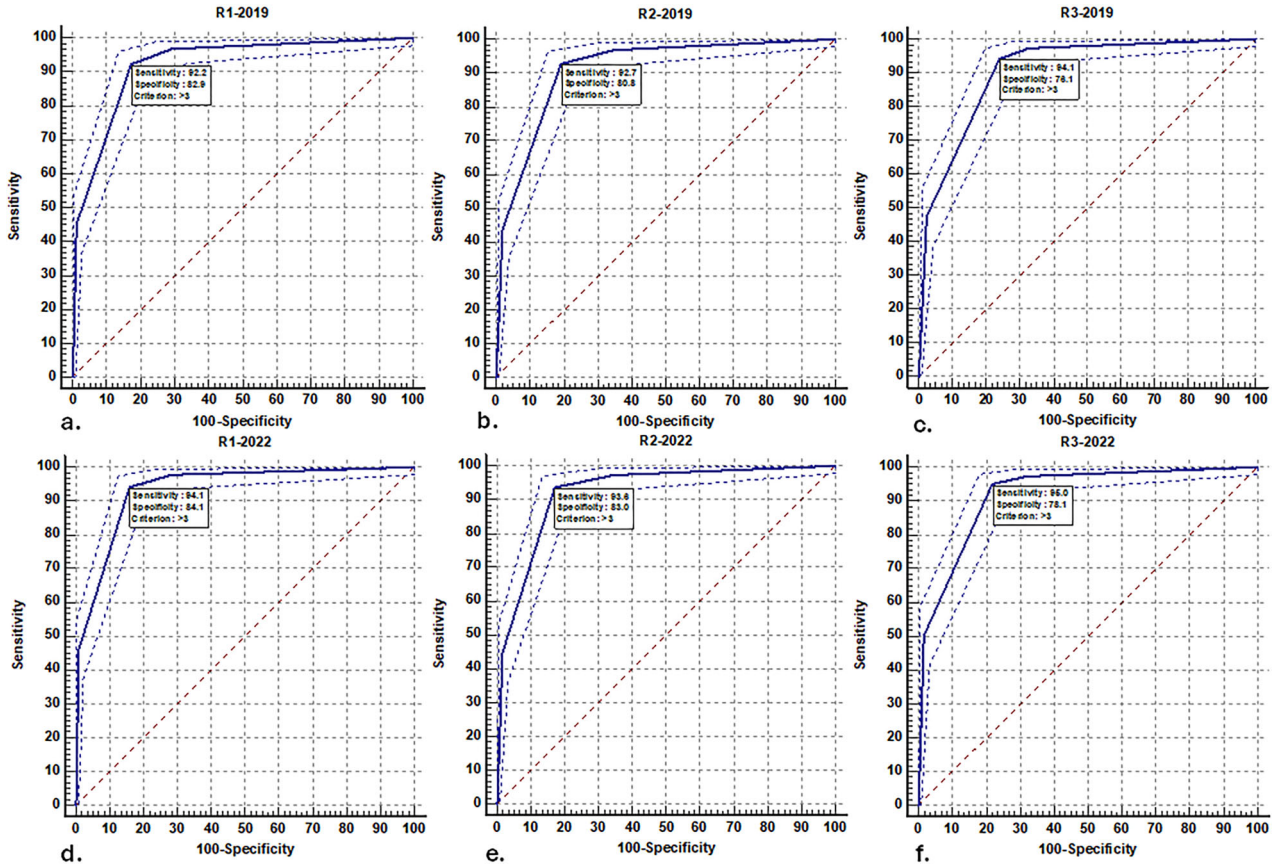

Fig. 2 The ROC of the diagnostic accuracy of the O-RADS v2019 (**a**, **b**, **c**) and 2022 (**d**, **e**, **f**) in predicting malignancy of adnexal masses as evidenced by histopathology as a reference standard and according to each radiologist

studies played a crucial role in guiding clinical management and optimizing patient outcomes. By comparing the diagnostic accuracy of both versions of the O-RADS ultrasound, researchers can observe the evolution and refinement of the system over time, potentially leading to improved risk assessment and decision-making. This prospective multicenter study included 803 women with 855 AMs, and histopathology was used as the reference standard for evaluating the diagnostic accuracy and reliability of O-RADS ultrasound v2019 and v2022. The results indicated that both versions of the O-RADS ultrasound demonstrated good diagnostic accuracy in predicting malignant AMs. O-RADS ultrasound v2022 showed slightly higher accuracy (82.5–86.7% vs. 80.7–85.3%), sensitivity (93.6–95.0% vs. 92.2–94.1%), and specificity (78.1–84.1% vs. 76.1–82.9%) compared to v2019. The overall IOA was perfect for both versions, with v2022 having a slightly higher agreement than v2019 ($\kappa = 0.97$ vs. 0.96). These findings highlight the reliability and consistency of both versions in assessing the risk of AMs. The improvements introduced in O-RADS ultrasound v2022 contribute to its enhanced diagnostic accuracy, reinforcing its potential to improve patient care and outcomes.

The diagnostic value of O-RADS ultrasound v2019 observed in our study aligns with previous validation studies that consistently demonstrated the good diagnostic accuracy and reliability of the O-RADS ultrasound system. These studies reported sensitivity values ranging from 90.6 to 98.7% for detecting malignant AMs [11, 14, 20, 21]. The minor variations in sensitivity across different studies may be attributed to factors such as differences in study populations, prevalence of malignancy, and experience level of the interpreting radiologists. Despite these variations, the overall findings consistently support the high sensitivity of O-RADS ultrasound in identifying malignant AMs, underscoring its utility as a reliable diagnostic tool in clinical practice.

Since its introduction in 2019, several attempts have been made to enhance the diagnostic accuracy of O-RADS ultrasound. Coa et al [11] sub-classified the O-RADS 4 category, improving specificity and facilitating

Table 4 Diagnostic accuracy of O-RADS ultrasound v2019 and v2022 in predicting malignancy of AMs according to each radiologist

Parameters	O-RADS v2019			O-RADS v2022		
	O1	O2	O3	O1	O2	O3
Cutoff	> O-RADS 3	> O-RADS 3	> O-RADS 3	> O-RADS 3	> O-RADS 3	> O-RADS 3
Number of true-positive findings	202	203	206	206	205	208
Number of false-negative findings	17	16	13	13	14	11
Number of false-positive findings	109	122	152	101	108	139
Number of true-negative findings	527	514	484	535	528	497
Accuracy (%)	85.3 (729/855) [82.7–87.6]	83.9 (717/855) [81.2–86.3]	80.7 (690/855) [77.9–83.3]	86.7 (723/855) [84.2–88.97]	85.7 (733/855) [83.2–88.0]	82.5 (705/855) [79.7–85.0]
Sensitivity (%)	92.2 (202/219) [87.9–95.4]	92.7 (203/219) [88.4–95.8]	94.1 (206/219) [90.1–96.8]	94.1 (206/219) [90.1–96.8]	93.6 (205/219) [89.5–96.5]	95.0 (208/219) [91.2–97.5]
Specificity (%)	82.9 (527/636) [89.7–85.7]	80.8 (514/636) [77.5–83.8]	76.1 (484/636) [72.6–79.4]	84.1 (535/636) [81.0–86.9]	83.0 (528/636) [79.9–85.9]	78.1 (497/636) [74.7–81.3]
Positive predictive value (%)	65.0 (202/311) [60.9–68.8]	62.5 (203/325) [58.6–66.2]	57.5 (206/358) [54.0–61.0]	67.1 (206/307) [62.0–70.0]	65.5 (205/313) [61.4–69.3]	59.9 (208/347) [56.3–63.5]
Negative predictive value (%)	96.9 (527/544) [95.2–98.0]	97.0 (514/530) [95.2–98.1]	97.4 (484/497) [95.6–98.4]	97.6 (535/548) [96.0–98.6]	97.4 (528/542) [95.8–98.4]	97.8 (497/508) [96.2–98.8]
AUC	0.923 [0.904–0.940]	0.912 [0.891–0.930]	0.906 [0.884–0.925]	0.936 [0.917–0.951]	0.923 [0.903–0.940]	0.919 [0.899–0.937]

Data in parentheses were used to calculate percentages. Data in brackets are 95% confidence intervals
O-RADS Ovarian-Adnexal Reporting and Data System, *AMs* adnexal masses, *O* observer, *AUC* area under curve

better risk stratification for surgical planning. Hack et al [16] incorporated acoustic shadowing as a benign finding, increasing the AUC to 0.94 with 99% sensitivity and 70% specificity. Additionally, recent studies have explored the combination of the O-RADS with other imaging modalities and biomarkers to further improve its efficacy in diagnosing AMs. Some studies have combined O-RADS with contrast-enhanced ultrasound (CEUS) and found that the addition of CEUS significantly enhances the diagnostic accuracy of O-RADS [23–26]. Other studies have combined O-RADS with serum cancer antigen 125 (CA125) and found that the inclusion of CA125 in the O-RADS improves its accuracy [27, 28]. These findings demonstrate the potential benefits of incorporating complementary techniques and biomarkers into the O-RADS framework to enhance diagnostic capabilities. A unique aspect of our study was the introduction of the updated O-RADS ultrasound v2022 [10]. While the difference was not statistically significant, O-RADS ultrasound v2022 demonstrated slightly better diagnostic accuracy compared to v2019.

Notably, we observed lower malignancy rates with the O-RADS ultrasound v2022 criteria compared to v2019 for the O-RADS 2 category (1.1–1.4% vs. 1.4–1.7%) and the O-RADS 3 category (6.7–9.3% vs. 8.3–11.6%). However, we observed higher malignancy rates with v2022 than v2019 for the O-RADS 4 category (43.4–52.2% vs. 42.7–50.2%) and the O-RADS 5 category (90.7–95.3% vs. 87.4–93.4%). These findings indicate that the updated O-RADS ultrasound v2022 criteria may have the potential for improved estimation of risk of malignancy compared to the previous version. Our findings align with Su et al [29], who reported higher accuracy (89.4%) and specificity (86.1%) for O-RADS ultrasound v2022 than v2019 (84.4% and 79.5%, respectively), with similar sensitivity (100%). These results suggest that the refinements introduced in v2022 could contribute to enhanced risk assessment and decision-making for AMs.

The updated descriptors in O-RADS ultrasound v2022 played a crucial role in reclassifying AMs. These descriptors include locularity, solid/smooth with shadowing, internal versus wall or septal color Doppler, and punctate wall foci of endometriosis. These descriptors were crucial in refining the categorization of AMS and were carefully applied in our analysis. Our analysis showed that these new descriptors resulted in upgrades and downgrades in a small percentage of AMs across observers. Particularly, Observer 1 had a 0.5% change rate, Observer 2 had a 1.4% change rate, and Observer 3 had a 1.3% change rate. These changes were driven by the improved specificity of the updated descriptors, which provided clearer guidelines for AM

Table 5 Inter- and intra-observer agreement for ultrasound imaging features and O-RADS ultrasound classification

Observer	Parameter	O1	O2	O3	All observers
O1	Internal echoes and/or incomplete septation	0.95 (0.86–1.00)	0.65 (0.39–0.91)	0.64 (0.40–0.89)	
	Uni/multilocular	0.98 (0.96–1.00)	0.77 (0.65–0.68)	0.71 (0.59–0.74)	
	Classic benign ovarian lesions	0.97 (0.93–1.00)	0.66 (0.55–0.77)	0.63 (0.51–0.75)	
	Solid lesions	0.91 (0.81–0.99)	0.45 (0.24–0.65)	0.51 (0.31–0.70)	
	Papillary projections or nodules	0.90 (0.83–0.99)	0.64 (0.51–0.78)	0.62 (0.48–0.75)	
	Smooth/irregular	0.88 (0.81–0.98)	0.66 (0.53–0.80)	0.62 (0.49–0.76)	
	Acoustic shadowing	0.85 (0.71–0.99)	0.67 (0.48–0.87)	0.61 (0.38–0.83)	
	Ascites and/or peritoneal nodules	0.91 (0.72–1.00)	0.72 (0.41–1.00)	0.65 (0.34–0.97)	
	Color score	0.89 (0.83–0.96)	0.52 (0.41–0.63)	0.57 (0.47–0.67)	
	O-RADS v2019		0.94 (0.93–0.96)	0.92 (0.90–0.94)	
	Updated O-RADS v2022		0.94 (0.93–0.96)	0.94 (0.93–0.96)	
O2	Internal echoes and/or incomplete septation		0.92 (0.76–1.00)	0.65 (0.37–0.94)	
	Uni/multilocular		0.97 (0.74–1.00)	0.73 (0.62–0.84)	
	Classic benign ovarian lesions		0.96 (0.92–1.00)	0.68 (0.56–0.78)	
	Solid lesions		0.87 (0.72–1.00)	0.66 (0.48–0.87)	
	Papillary projections or nodules		0.92 (0.86–0.98)	0.67 (0.54–0.80)	
	Smooth/irregular		0.88 (0.81–0.97)	0.61 (0.39–0.82)	
	Acoustic shadowing		0.89 (0.77–1.00)	0.69 (0.48–0.89)	
	Ascites and/or peritoneal nodules		0.92 (0.76–1.00)	0.76 (0.51–1.00)	
	Color score		0.88 (0.81–0.94)	0.50 (0.37–0.62)	
	O-RADS v2019			0.94 (0.93–0.96)	
	Updated O-RADS v2022			0.94 (0.92–0.95)	
O3	Internal echoes and/or incomplete septation			0.94 (0.83–1.00)	
	Uni/multilocular			0.96 (0.93–1.00)	
	Classic benign ovarian lesions			0.96 (0.92–1.00)	
	Solid lesions			0.89 (0.77–1.00)	
	Papillary projections or nodules			0.88 (0.80–0.97)	
	Smooth/irregular			0.94 (0.88–0.99)	
	Acoustic shadowing			0.87 (0.72–1.00)	
	Ascites and/or peritoneal nodules			0.93 (0.79–1.00)	
	Color score			0.91 (0.87–0.97)	
	O-RADS v2019				
	Updated O-RADS v2022				
All observers	Internal echoes and/or incomplete septation				0.65 (0.57–0.72)
	Uni/multilocular				0.79 (0.74–0.84)
	Typical benign ovarian lesions				0.65 (0.58–0.72)
	Solid lesions				0.53 (0.44.62)
	Papillary projections or nodules				0.69 (0.62–0.76)
	Smooth/irregular				0.66 (0.59–0.73)
	Acoustic shadowing				0.63 (0.55–0.71)
	Ascites and/or peritoneal nodules				0.71 (0.64–0.77)
	Color score				0.65 (0.58–0.73)
	O-RADS v2019				0.96 (0.96–0.97)
	Updated O-RADS v2022				0.97 (0.96–0.97)

Data are Kappa values. Data in parentheses are 95% confidence intervals. The κ 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 = good agreement; and 0.81–1.00 = perfect agreement
O-RADS Ovarian-Adnexal Reporting and Data System, *O* observer

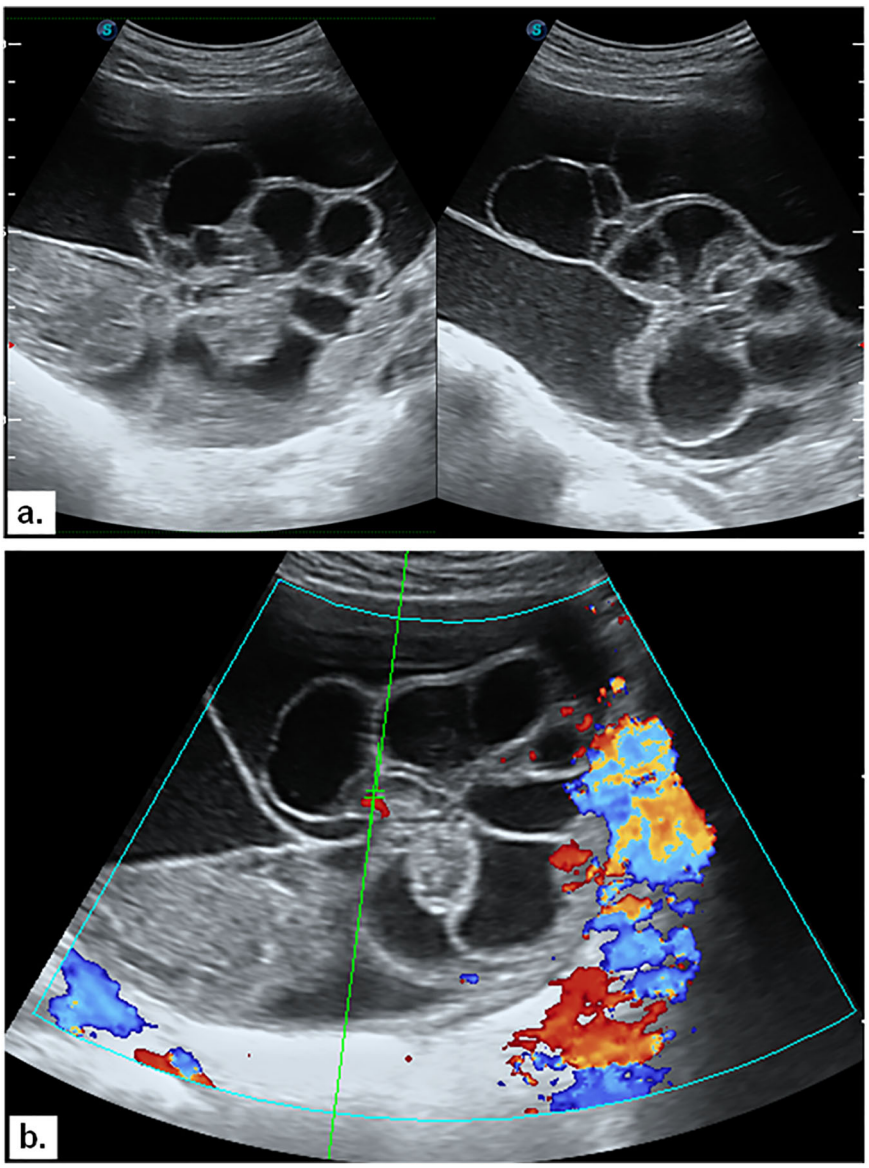

Fig. 3 A 16-year-old female presented with pelvic pain. **a** Transverse and longitudinal transabdominal gray-scale ultrasound images reveal a 15.5-cm right adnexal large, well-defined, multilocular cystic lesion with smooth walls, fine turbid fluid, multiple internal thick septa, and intramural echogenic content. **b** A transverse color Doppler ultrasound image shows minimal flow in the septa (color score = 2). The adnexal mass was categorized as O-RADS 4 by all reviewers using both the 2019 and 2022 versions of O-RADS. Surgical intervention was performed, and the histopathological examination revealed an ovarian borderline mucinous tumor

categorization. The updated criteria also led to more accurate risk stratification, particularly for bilocular smooth cysts and solid smooth lesions with shadowing. As a result, the percentage of pathologically proven malignant lesions decreased in the O-RADS 2 category and increased in the O-RADS 4 and 5 categories. The enhanced specificity of v2022 descriptors contributed to a more nuanced classification system, allowing finer distinctions between lesion types. This improvement is evident in the slight increase in overall specificity from 76.1–82.9% in v2019 to 78.1–84.1% in v2022. These findings emphasize the importance of the new descriptors in enhancing lesion assessment precision and improving the reliability of the O-RADS classification system. The refined criteria ensure a more comprehensive and consistent categorization of lesions, ultimately leading to better clinical decision-making and patient care.

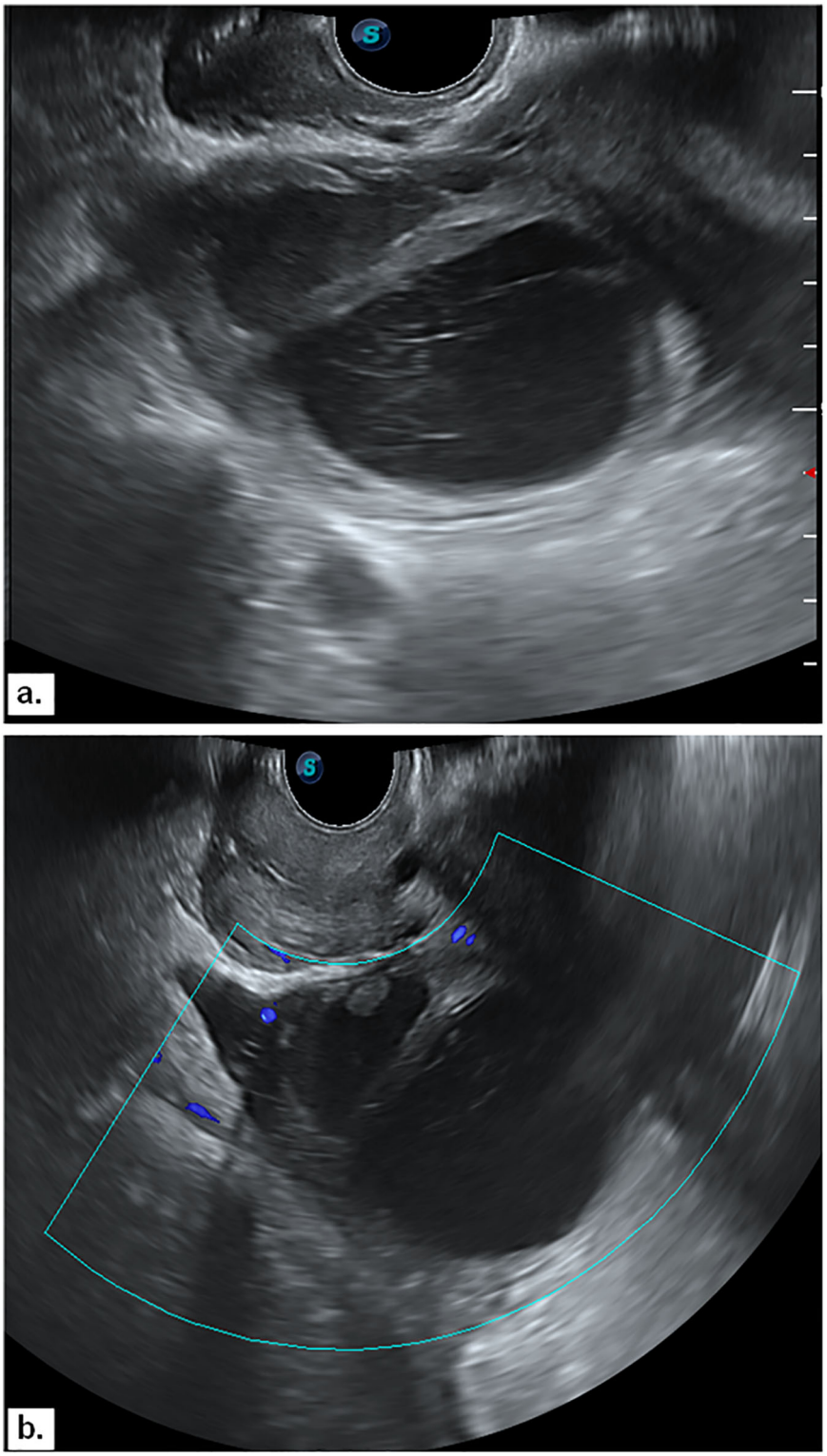

Fig. 4 A 28-year-old woman presented with pelvic pain. **a** A longitudinal transvaginal gray-scale ultrasound image reveals a 7-cm left adnexal well-defined, bilocular cystic lesion with a thick irregular inner wall, a single thick septum, fine turbid fluid, and fine reticulations. **b** A transvaginal color Doppler ultrasound image shows no flow (color score = 1). The adnexal mass was categorized as O-RADS 3 by O1 and O3, and as O-RADS 4 by O2 according to the 2019 version of O-RADS. According to the 2022 version, the adnexal mass was categorized as O-RADS 4 by O1 and O2, and as O-RADS 2 by O3. Surgical intervention was performed, and the histopathological examination confirmed the diagnosis of an endometrioma

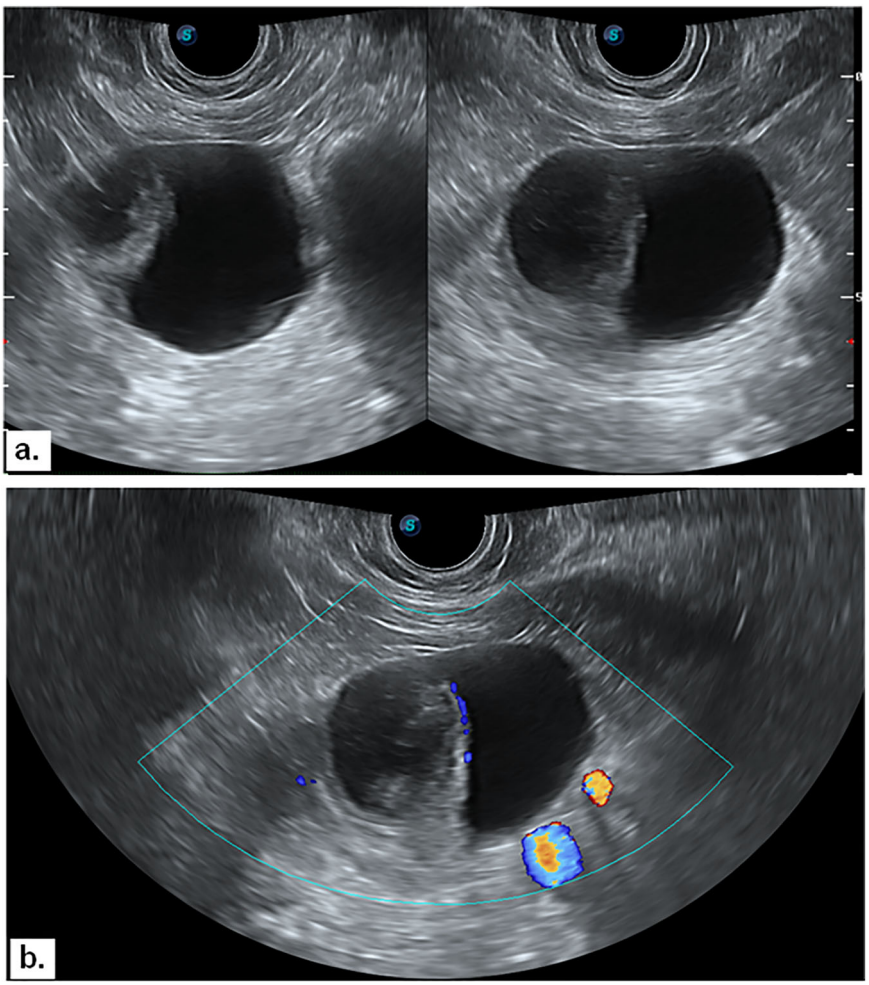

Fig. 5 A 41-year-old female presented with pelvic pain. **a** Transverse and longitudinal transvaginal gray-scale ultrasound images reveal a 6.4-cm right adnexal bilocular cystic lesion with a thick irregular septum and fine turbid fluid. **b** A longitudinal transvaginal color Doppler ultrasound image shows minimal flow within the septum (color score = 2). The adnexal mass was categorized as O-RADS 3 by O1 and as O-RADS 4 by O2 and O3 according to the 2019 version of the O-RADS. According to the 2022 version, the adnexal mass was categorized as O-RADS 4 by all observers. Surgical intervention was performed, and the histopathological examination revealed a serous cystadenocarcinoma

Although the updated O-RADS ultrasound v2022 shows improved specificity (78.1–84.1%) compared to v2019 (76.1–82.9%), it is still relatively low compared to some other ACR reporting and data systems. This result is in line with the study by Coa et al [11], who also found a low specificity (83.2%) for the O-RADS ultrasound v2019 compared to its counterparts. Similarly, Jha et al [14] reported a lower specificity (81.9%) due to a lower cancer prevalence in their study population. In contrast, Pi et al [15] reported excellent specificity (92–100%). It is important to recognize that differences in study populations, cancer prevalence, and interpretation criteria may contribute to differences in specificity between studies. The lower specificity of the O-RADS system may be attributed to the fact that it prioritizes sensitivity over specificity as it is designed for the detection of low-prevalence adnexal malignancies [14]. However, this trade-off may lead to a higher false-positive rate, necessitating further diagnostic investigations or unnecessary interventions. This balance between sensitivity and specificity remains a challenge for optimizing the clinical utility of O-RADS ultrasound. This approach is particularly important for a general population without high risk, as opposed to a high-risk population as assessed by LI-RADS. Although higher specificity would be an advantage, it is difficult to achieve both high sensitivity and high specificity in this context. To increase specificity, additional methods can be integrated into the assessment process. For example, the widely validated IOTA/ADNEX mathematical model [30] can provide specific risk scores and probabilities of malignancy types,

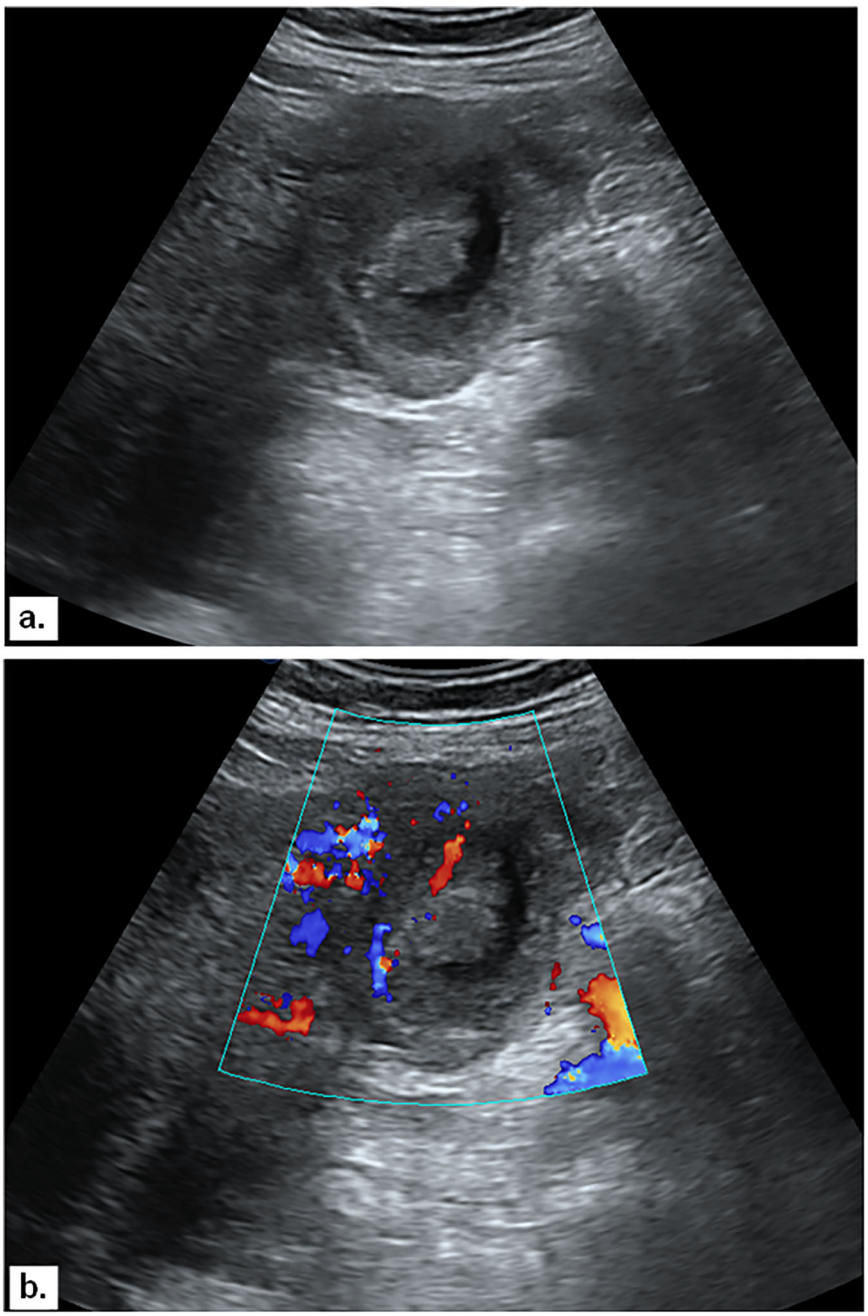

Fig. 6 A 62-year-old female presented with pelvic pain. **a** A longitudinal transabdominal gray-scale ultrasound image reveals a 6.2-cm left adnexal unilocular cystic lesion with an irregular thick wall and an intramural solid component. **b** A longitudinal transabdominal color Doppler ultrasound image shows moderate flow (color score = 3). The adnexal mass was categorized as O-RADS 4 by all observers using both the 2019 and 2022 versions of the O-RADS. Surgical intervention was performed, and the histopathological examination revealed a mucinous cystadenocarcinoma

offering a quantitative complement to the qualitative O-RADS assessments. In addition, treatment recommendations often include subjective assessments from experienced ultrasound specialists, which have been shown to be highly accurate. Incorporating these expert assessments into the O-RADS framework can improve diagnostic utility and accuracy. These results highlight the need to further refine and optimize the O-RADS system to improve its specificity, particularly in populations with a lower prevalence of cancer. Continued research and investigation of

additional imaging findings, integration of other modalities, and inclusion of specific patient characteristics may improve the specificity and overall accuracy of O-RADS in evaluating AMs.

The present study reported perfect inter-observer agreement (IOA) for both O-RADS ultrasound versions, despite some variability in interpreting individual features. This high agreement indicates robust risk stratification, unaffected by minor differences in imaging feature interpretation. Previous studies have consistently demonstrated good IOA for O-RADS, underscoring its reliability and reproducibility in clinical practice. For example, Cao et al [11] reported good agreement ($\kappa = 0.714$) between less-experienced and expert radiologists, Pi et al [15] observed very good overall agreement ($\kappa = 0.82$), and Wu et al [21] found good agreement among experienced sonologists ($\kappa = 0.749–0.773$). Our findings align with these results, highlighting the reproducibility of O-RADS, which ensures consistent risk stratification and management recommendations for adnexal masses. Additionally, the comparable agreement for v2019 and v2022 in our study suggests that the updates in v2022 did not impact the system's reproducibility. Similarly, our study confirmed the high intra-observer agreement, consistent with previous findings. Wu et al [21] reported good to excellent intra-observer agreements of O-RADS ultrasound analysis for the four sonologists ($\kappa = 0.661–0.841$). These consistent results reinforce the robustness of the O-RADS framework in clinical practice. Our study adds evidence that the updated O-RADS v2022 maintains or slightly improves the high intra-observer agreement seen with v2019, enhancing diagnostic confidence and clinical decision-making.

Our study had several notable strengths. First, we included a large number of AMs from multiple centers, which enhances the generalizability of our findings. Second, we used histopathology as the reference standard, providing a robust basis for evaluating the diagnostic accuracy of O-RADS ultrasound. Third, the radiologists involved in the assessment were blinded to the clinical data, thereby minimizing the potential bias in their interpretations. Fourth, we directly compared the accuracy of O-RADS ultrasound v2019 and v2022 within the same cohort, allowing for a more accurate assessment of the updates. However, certain limitations should be acknowledged. First, the retrospective application of O-RADS ultrasound v2022 to previously acquired images may introduce recall or interpretation biases, potentially impacting the comparative analysis. Second, TV ultrasound was not feasible in all patients, which may have affected diagnostic accuracy. Third, our study involved only experienced radiologists; therefore, the results may not reflect the performance of less-experienced radiologists. Fourth, the use of different ultrasound machines across participating centers could have introduced variability in the image quality and characteristics, potentially affecting the consistency of the results. Fifth, the lack of long-term outcomes or follow-up data limited our ability to assess the clinical utility of the O-RADS ultrasound in terms of overall survival, recurrence rates, and its impact on treatment decisions. Finally, the lack of inclusion of AMs that resolved on follow-up without surgical intervention, skewed our data toward a higher prevalence of malignancy. However, the use of histopathologic examination as the reference standard provided a high diagnostic confidence level for the cases included in our analysis.

Conclusion

This study provides evidence supporting the diagnostic accuracy and reliability of the O-RADS ultrasound classification system. Both versions demonstrated good diagnostic accuracy and reliability in predicting malignant AMs, with slight improvements observed in v2022. Further research and validation studies are warranted to strengthen the evidence base and refine the O-RADS for wider clinical application.

Abbreviations

AMs	Adnexal masses
CA125	Cancer antigen 125
CEUS	Contrast-enhanced ultrasound
IOA	Inter-observer agreement
O-RADS	Ovarian-Adnexal Reporting and Data System
TV	Transvaginal

Acknowledgements

The authors express their gratitude to all the staff members and colleagues in the Radiology and Obstetrics & Gynecology Departments at Zagazig University, Egypt, for their valuable cooperation and support.

Funding

The authors state that this work has not received any funding.

Compliance with ethical standards

Guarantor

The scientific guarantor of this publication is the corresponding author.

Conflict of interest

M.A.A.B. is a member of the Scientific Editorial Board of *European Radiology* (section: urogenital). As such they have not participated in the selection nor review processes for this article. The remaining authors of this manuscript declare no relevant conflicts of interest and no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry

The corresponding author has significant statistical expertise.

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

Ethical approval
Institutional review board approval was obtained (approval number: ZU-10352).

Study subjects or cohorts overlap
Our study cohorts are unique and have not been previously reported.

Methodology

- Prospective
- Diagnostic or prognostic study
- Multicenter study

Author details
[1]Division of Radiology, Department of 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 Radio-diagnosis, Faculty of Human Medicine, Suez Canal University, Esmaelia, Egypt. [4]Department of Radio-diagnosis, Faculty of Human Medicine, Al Azhar University, Cairo, Egypt. [5]Department of Diagnostic Radiology, Faculty of Human Medicine, Tanta University, Tanta, Egypt. [6]Department of Diagnostic Radiology, Faculty of Human Medicine, Benha University, Benha, Egypt. [7]Department of Diagnostic Radiology, National Institute of Urology and Nephrology, Cairo, Egypt. [8]Department of Diagnostic Radiology, Urology and Nephrology Center, Mansoura University, Mansoura, Egypt. [9]Department of Radiology, Kobry Al Kobba Military Hospital, Cairo, Egypt. [10]Department of Radiology, College of Medicine, Qassim University, Buraidah, Kingdom of Saudi Arabia. [11]Department of Obstetrics and Gynecology, Faculty of Medicine, Najran University, Najran, Saudi Arabia. [12]Department of Obstetrics and Gynecology, Faculty of Human Medicine, Zagazig University, Zagazig, Egypt. [13]Department of Clinical Oncology, Faculty of Human Medicine, Zagazig University, Zagazig, Egypt. [14]Faculty of General Medicine, St. Petersburg State University, Egypt Branch, Cairo, Egypt.

Received: 20 March 2024 Revised: 11 October 2024 Accepted: 27 October 2024
Published online: 28 November 2024

References

1. Levine D, Brown DL, Andreotti RF et al (2010) Management of asymptomatic ovarian and other adnexal cysts imaged at US Society of Radiologists in Ultrasound consensus conference statement. Ultrasound Q 26:121–131
2. Timmerman D, Ameye L, Fischerova D et al (2010) Simple ultrasound rules to distinguish between benign and malignant adnexal masses before surgery: prospective validation by IOTA group. BMJ 341:c6839
3. Amor F, Alcázar JL, Vaccaro H, León M, Iturra A (2011) GI-RADS reporting system for ultrasound evaluation of adnexal masses in clinical practice: a prospective multicenter study. Ultrasound Obstet Gynecol 38:450–455
4. Suh-Burgmann E, Flanagan T, Osinski T, Alavi M, Herrinton L (2018) Prospective validation of a standardized ultrasonography-based ovarian cancer risk assessment system. Obstet Gynecol 32:1101–1111
5. Timmerman S, Valentin L, Ceusters J et al (2023) External validation of the Ovarian-Adnexal Reporting and Data System (O-RADS) lexicon and the International Ovarian Tumor Analysis 2-step strategy to stratify ovarian tumors into O-RADS risk groups. JAMA Oncol 9:225–233
6. American College of Obstetricians and Gynecologists' Committee on Practice Bulletins—Gynecology (2016) Practice bulletin no. 174: evaluation and management of adnexal masses. Obstet Gynecol 128:e210–e226
7. Timmerman D, Van Calster B, Testa A et al (2016) Predicting the risk of malignancy in adnexal masses based on the Simple Rules from the International Ovarian Tumor Analysis group. Am J Obstet Gynecol 214:424–437
8. Andreotti RF, Timmerman D, Benacerraf BR et al (2019) Ovarian-adnexal reporting lexicon for ultrasound: a white paper of the ACR Ovarian-Adnexal Reporting and Data System Committee. J Am Coll Radiol 2018;15:1415–1429. [Published correction appears in J Am Coll Radiol 16:403–406]
9. Andreotti RF, Timmerman D, Strachowski LM et al (2020) O-RADS US risk stratification and management system: a consensus guideline from the ACR Ovarian-Adnexal Reporting and Data System Committee. Radiology 294:168–185
10. Strachowski LM, Jha P, Phillips CH et al (2023) O-RADS US v2022: an update from the American College of Radiology's Ovarian-Adnexal Reporting and Data System US Committee. Radiology 308:e230685
11. Cao L, Wei M, Liu Y et al (2021) Validation of American College of Radiology Ovarian-Adnexal Reporting and Data System Ultrasound (O-RADS US): analysis on 1054 adnexal masses. Gynecol Oncol 162:107–112
12. Lai HW, Lyu GR, Kang Z, Li LY, Zhang Y, Huang YJ (2021) Comparison of O-RADS, GI-RADS, and ADNEX for diagnosis of adnexal masses: an external validation study conducted by junior sonologists. J Ultrasound Med 21:1497–1507
13. Guo Y, Zhou S, Zhao B, Wen L, Liu M (2022) Ultrasound findings and O-RADS malignancy risk stratification of ovarian collision tumors. J Ultrasound Med 41:2325–2331
14. Jha P, Gupta A, Baran TM et al (2022) Diagnostic performance of the Ovarian-Adnexal Reporting and Data System (O-RADS) ultrasound risk score in women in the United States. JAMA Network Open 5:e2216370–e2216370
15. Pi Y, Wilson MP, Katlariwala P et al (2021) Diagnostic accuracy and inter-observer reliability of the O-RADS scoring system among staff radiologists in a North American academic clinical setting. Abdom Radiol (NY) 46:4967–4973
16. Hack K, Gandhi N, Bouchard-Fortier G et al (2022) External validation of O-RADS US risk stratification and management system. Radiology 304:114–120
17. Solis Cano DG, Cervantes Flores HA, De Los Santos Farrera O, Guzman Martinez NB, Soria C´ espedes D (2021) Sensitivity and specificity of ultrasonography using Ovarian-Adnexal Reporting and Data System classification versus pathology findings for ovarian cancer. Cureus 13:e17646
18. Basha MAA, Metwally MI, Gamil SA et al (2021) Comparison of O-RADS, GI-RADS, and IOTA simple rules regarding malignancy rate, validity, and reliability for diagnosis of adnexal masses. Eur Radiol 31:674–684
19. Hiett AK, Sonek J, Guy M, Reid TJ (2022) Performance of IOTA Simple Rules, Simple Rules Risk Assessment, ADNEX model and O-RADS in discriminating between benign and malignant adnexal lesions in North American population. Ultrasound Obstet Gynecol 59:668–676
20. Vara J, Manzour N, Chacón E et al (2022) Ovarian Adnexal Reporting Data System (O-RADS) for classifying adnexal masses: a systematic review and meta-analysis. Cancers (Basel) 14:3151
21. Wu M, Zhang M, Cao J et al (2023) Predictive accuracy and reproducibility of the O-RADS US scoring system among sonologists with different training levels. Arch Gynecol Obstet 308:631–637
22. Meinhold-Heerlein I, Fotopoulou C, Harter P et al (2016) The new WHO classification of ovarian, fallopian tube, and primary peritoneal cancer and its clinical implications. Arch Gynecol Obstet 293:695–700
23. Shi Y, Li H, Wu X, Li X, Yang M (2023) O-RADS combined with contrast-enhanced ultrasound in risk stratification of adnexal masses. J Ovarian Res 16:153
24. Yuan K, Huang YJ, Mao MY et al (2023) Contrast-enhanced US to improve diagnostic performance of O-RADS US risk stratification system for malignancy. Radiology 308:e223003
25. Wang T, Cui W, Nie F et al (2023) Comparative study of the efficacy of the Ovarian-Adnexa Reporting and Data System ultrasound combined with contrast-enhanced ultrasound and the ADNEX MR scoring system in the diagnosis of adnexal masses. Ultrasound Med Biol 49:2072–2080
26. Xu J, Huang Z, Zeng J et al (2023) Value of contrast-enhanced ultrasound parameters in the evaluation of adnexal masses with Ovarian-Adnexal Reporting and Data System ultrasound. Ultrasound Med Biol 49:1527–1534
27. Wang R, Li X, Li S et al (2023) Clinical value of O-RADS combined with serum CA125 and HE4 for the diagnosis of ovarian tumours. Acta Radiol 64:821–828
28. Yang Y, Ju H, Huang Y (2023) Diagnostic performance of IOTA SR and O-RADS combined with CA125, HE4, and risk of malignancy algorithm to

distinguish benign and malignant adnexal masses. Eur J Radiol 165:110926
29. Su N, Yang Y, Liu Z et al (2023) Validation of the diagnostic efficacy of O-RADS in adnexal masses. Sci Rep 13:15667
30. Van Calster B, Van Hoorde K, Valentin L et al (2014) Evaluating the risk of ovarian cancer before surgery using the ADNEX model to differentiate between benign, borderline, early and advanced stage invasive, and secondary metastatic tumours: prospective multicentre diagnostic study. BMJ 15:349

Publisher's Note