

A comparison of the diagnostic performance of the O-RADS, RMI4, IOTA LR2, and IOTA SR systems by senior and junior doctors

Yuyang Guo¹, Baihua Zhao¹, Shan Zhou¹, Lieming Wen¹, Jieyu Liu¹, Yaqian Fu², Fang Xu³, Minghui Liu¹

¹Department of Ultrasound Diagnosis, The Second Xiangya Hospital, Central South University, Changsha; ²Health Management Center, The Second Xiangya Hospital, Central South University, Changsha; ³Department of Ultrasonography, The First Hospital of Changsha, Changsha, China

Purpose: This study compared the diagnostic performance of the Ovarian-Adnexal Reporting and Data System (O-RADS), the Risk of Malignancy Index 4 (RMI4), the International Ovarian of Tumor Analysis Logistic Regression Model 2 (IOTA LR2), and the IOTA Simple Rules (IOTA SR) in predicting the malignancy of adnexal masses (AMs).

Methods: This retrospective study included 575 women with AMs between 2017 and 2020. All clinical messages, ultrasound images, and pathological findings were collected. Two senior doctors (group I) and two junior doctors (group II) used the four systems to classify AMs. The postoperative pathological diagnosis was used as the gold standard to evaluate the diagnostic efficiency. A receiver operating characteristic curve was used to test the diagnostic performance. The interrater agreement between the two groups was tested using kappa values.

Results: Of all 592 AMs, 447 (75.5%) were benign, 123 (20.8%) were malignant, and 22 (3.7%) were borderline. The intergroup consistency test yielded kappa values of 0.71, 0.92, 0.68, and 0.77 for the O-RADS, RMI4, IOTA LR2, and IOTA SR, respectively. To predict malignant lesions, the areas under the curve of the O-RADS, RMI4, IOTA LR2, and IOTA SR systems were 0.90, 0.89, 0.90, and 0.86 for group I and 0.89, 0.87, 0.88, and 0.84 for group II, respectively. The O-RADS had the highest sensitivity (91.0% in group I and 84.8% in group II).

Conclusion: The four diagnostic systems could compensate for junior doctors' inexperience in predicting malignant adnexal lesions. The O-RADS performed best and showed the highest sensitivity.

Keywords: Adnexal masses; Ultrasonography; O-RADS; IOTA; RMI

Key points: This is the first comparison of the diagnostic performance of the Ovarian-Adnexal Reporting and Data System (O-RADS), Risk of Malignancy Index 4 (RMI4), International Ovarian of Tumor Analysis Logistic Regression Model 2 (IOTA LR2), and IOTA Simple Rules (IOTA SR) systems in a large sample from Asian populations. The diagnostic efficiency and reliability of the four systems could compensate for junior doctors' inexperience in predicting the malignancy of adnexal masses. It may make more sense to evaluate and improve those ultrasound predicting models for clinical management and surgical strategy.

ORIGINAL ARTICLE

<https://doi.org/10.14366/usg.21237>
pISSN: 2288-5919 • eISSN: 2288-5943
Ultrasonography 2022;41:511-518

Received: November 7, 2021

Revised: January 21, 2022

Accepted: January 31, 2022

Correspondence to:

Baihua Zhao, MD, Department of Ultrasound Diagnosis, The Second Xiangya Hospital, Central South University, No. 139, Renmin Middle Road, Changsha, Hunan 410011, China

Tel. +86-15116377643

Fax. +86-073185292140

E-mail: zhaobaihua2000@csu.edu.cn,

zhaobaihua2000@126.com

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2022 Korean Society of Ultrasound in Medicine (KSUM)



How to cite this article:

Guo Y, Zhao B, Zhou S, Wen L, Liu J, Fu Y, et al. A comparison of the diagnostic performance of the O-RADS, RMI4, IOTA LR2, and IOTA SR systems by senior and junior doctors. Ultrasonography. 2022 Jul;41(3):511-518.

Introduction

Pelvic ultrasonography is the most widely recognized noninvasive examination for adnexal masses. Ultrasonography is a low-cost and accessible, but highly experience-dependent modality. To improve clinical management and the surgical strategy through accurate predictions of the malignancy of adnexal masses, many guidelines, grading systems, and prediction models have been developed [1,2].

The International Ovarian of Tumor Analysis (IOTA) working group is a multicentric, large-sample, and ongoing study team on adnexal lesions. The IOTA Logistic Regression Model 2 (IOTA LR2) and the IOTA Simple Rule (IOTA SR), proposed from 2002 to 2007, are both prospective study products of the IOTA based on large samples [3]. The IOTA SR can compensate for junior doctors' lack of experience, minimize false negatives, and improve true positives. The high sensitivity of the IOTA LR2 in particular ensures that patients with true positives could be found as much as possible [4–11]. The Risk of Malignancy Index (RMI) system was initially developed from the RMI1 to RMI3 system, and the RMI4 system was proposed by Yamamoto et al. in 2009 [12–16]. It combines menopausal status, ultrasound results, and serum cancer antigen 125 (CA125) levels to provide a simple standard for assessing adnexal masses. The RMI4 score is calculated using the formula: $RMI4 = U \times M \times S \times CA125$, where U is the ultrasound score, M is the menopausal score, S is the tumor size score, and CA125 is the absolute value of serum CA125 levels. The 2020 Ovarian-Adnexal Reporting and Data System (O-RADS) was published by the American College of Radiology. This system includes the O-RADS ultrasound lexicon, risk categories, clinical management, and malignancy risk to reduce ambiguity in the description of lesions in ultrasound reports and to provide corresponding management approaches for patients with different risk grades [17,18].

All these prediction systems were established and tested based on data from European populations. These systems need to be more widely validated in practice with various ethnic populations [1,2]. This study aimed to compare the diagnostic efficiency and reliability among the O-RADS, RMI4, IOTA LR2, and IOTA SR for predicting benign and malignant ovarian tumors by senior and junior doctors.

Materials and Methods

Compliance with Ethical Standards

The study was approved by the Human Research Ethics Committee of the Second Xiangya Hospital (No. 2021-038) and performed in accordance with the principles of the Declaration of Helsinki. Written informed consents were waived.

Study Sample

Data from 575 women who underwent gynecological surgery with preoperative ultrasound examinations and postoperative histological diagnoses of adnexal masses in the Second Xiangya Hospital between January 2017 and October 2020 were retrospectively analyzed. The complete medical records of all patients were obtained, including age, menopausal status, gynecological examination, tumor markers, operation methods, postoperative pathology, and follow-up. A postmenopausal state was defined if women over the age of 50 had undergone hysterectomy or lacked records related to menopause.

The inclusion criteria were (1) an interval of less than 30 days between ultrasonography and gynecological surgery and (2) a definite pathological diagnosis.

The exclusion criteria were (1) pregnant women with adnexal masses, (2) women who had images of poor quality or without diagnostic signs, and (3) women with no clear menopausal status and no test for CA125 levels.

Instruments and Image Analysis

Ultrasound diagnostic systems with 9–15 MHz intracavitary transducers were used for the ultrasound examinations, which were performed by doctors at the attending level and above. Transabdominal ultrasonography was performed if the mass was too large to be observed using transvaginal ultrasonography. All images were stored and collected from the ultrasound working system. The ultrasound characteristics of each mass were assessed. The descriptions included single or bilateral, cystic component, morphology and margins, cyst wall thickness, acoustic shadowing, maximum diameter (maximum diameter of the tumor and maximum diameter of the solid part), solid papillary protrusions, separation, ascites, peritoneal nodules, and color Doppler score.

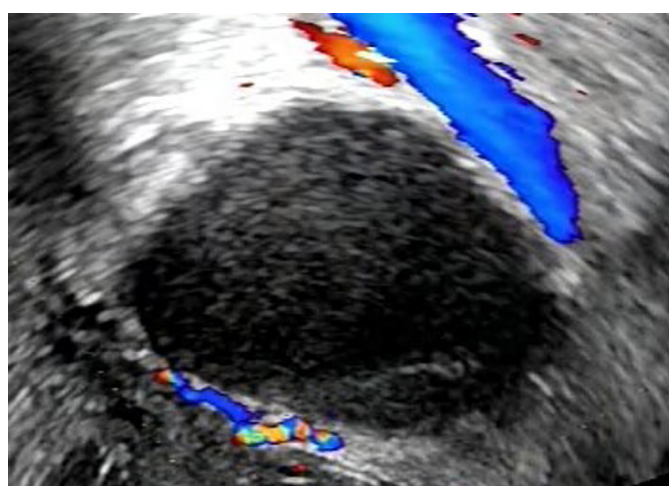
Senior doctors (L.W. and B.Z.; group I), with more than 10 years of ultrasonic diagnosis experience, and junior doctors (Y.G. and S.Z.; group II), with 1 year of ultrasonic diagnosis experience and the diagnosis of 300 adnexal tumors in practice, received theoretical and practical training on the four systems. After training, the authors had good to excellent agreement when applying the O-RADS, RMI4, IOTA LR2, and IOTA SR systems. A series of 40 adnexal masses were randomly selected for a test-retest analysis. In group I, for the O-RADS, RMI4, IOTA LR2, and IOTA SR systems, the intra-reader agreement tests yielded kappa values of 0.92 (95% confidence interval [CI], 0.78 to 1.00), 0.94 (95% CI, 0.83 to 1.00), 0.80 (95% CI, 0.59 to 1.00), and 0.71 (95% CI, 0.48 to 0.95), respectively; for group II, the kappa values were 0.93 (95% CI, 0.81 to 1.00), 0.82 (95% CI, 0.63 to 1.00), 0.87 (95% CI, 0.70 to 1.00), and 0.88 (95% CI, 0.72 to 1.00), respectively.

The two groups analyzed the images, evaluated each mass using the four systems, and were blinded to the clinical information and pathological results. In each group, the two doctors worked together to analyze the images. Lesions with O-RADS grades of 1–3 were classified as benign tumors, and lesions with O-RADS grades of 4–5 were classified as malignant tumors. The cutoff value of the RMI4 system was an RMI4 score of 450. The cutoff value for the IOTA LR2 was a malignancy risk of 10%. For the IOTA SR model, a mass with at least one malignant feature and no benign features was considered a malignant tumor, and a mass with only benign features was considered a benign lesion [3,13,17]. For intermediate cases,

with or without benign and malignant features simultaneously, in the IOTA SR, the doctors' subjective judgments were used as the outcomes [3,12,13]. All results were compared with the histological diagnosis, which was classified according to the International Federation of Gynecology and Obstetrics criteria [19], and borderline masses were classified as malignant (Fig. 1).

Statistical Analysis

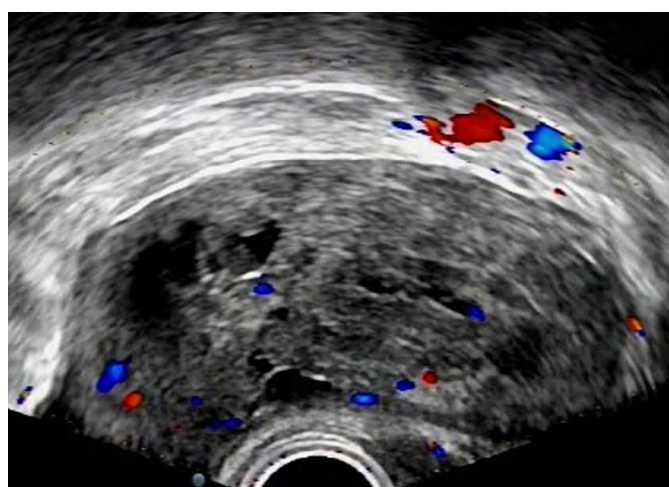
Statistical analysis was performed using SPSS ver. 26.0 (IBM Corp., Armonk, NY, USA) and GraphPad Prism 6.0 (GraphPad Software Inc., San Diego, CA, USA). Categorical variables were compared by



A



B



C

Fig. 1. Images exemplifying benign, borderline, and malignant masses.

A. Image of a pathologically proven ovarian endometrial cyst from a 38-year-old woman (case 129 among the 575 women) is shown. The mass was classified as O-RADS 2, a classic benign lesion with diameter <10 cm; the RMI4 score was 147.29; the malignancy risk of IOTA LR2 was 6%; and the IOTA SR model classified it as a benign lesion. **B.** Image of a 47-year-old woman with a pathologically proven borderline mucinous adenoma (case 7 among the 575 women) is shown. The mass was classified as O-RADS 4, a unilocular cyst with solid component; the RMI4 score was 96.18; the malignancy risk of IOTA LR2 was 20%; and the IOTA SR model categorized it as an intermediate case, which was classified as malignant by senior doctors' subjective judgment. **C.** Image of a postoperative pathological adult granulosa cell tumor from a 51-year-old woman (case 84 among the 575 women) is shown. The mass was classified as O-RADS 4, a smooth solid mass with a color score of 2–3; the RMI4 score was 66.72; the malignancy risk of IOTA LR2 was 48%; and the IOTA SR model classified it as a malignant lesion. O-RADS, Ovarian-Adnexal Reporting and Data System; RMI4, Risk of Malignancy Index model 4; IOTA, International Ovarian of Tumor Analysis; IOTA LR2, IOTA Logistic Regression Model 2; IOTA SR, IOTA Simple Rules.

the chi-square test or the Fisher exact test. Continuous variables were compared by the independent-sample t-test or rank-sum test. Receiver operating characteristic (ROC) curves were drawn to test the diagnostic performance of the four ultrasound classification systems in the two groups. The sensitivity, specificity, negative predictive value, positive predictive value, and Youden index were analyzed. The kappa coefficient was used to assess intergroup agreement. A $\kappa \geq 0.75$ was considered as indicating high repeatability, a $0.40 \leq \kappa < 0.75$ was considered as indicating medium repeatability, and a $\kappa < 0.40$ was considered as indicating low repeatability. A P-value < 0.05 was interpreted as statistically significant.

Results

The mean age of the 575 women was 39.0 ± 14.6 years (range, 6 to 81 years). Of the entire sample, 446 women (77.6%) were premenopausal, and 129 (22.4%) were postmenopausal. One hundred and eighteen (30.5%) women had bilateral lesions. One hundred and eighty-five (32.2%) women had elevated CA125 levels, including 135 (23.5%) premenopausal women and 50 (8.7%) postmenopausal women (Table 1). Eight women (1.4%) had undergone a hysterectomy.

With the inclusion of 17 (3.0%) bilateral adnexal lesions, data on a total of 592 adnexal masses were collected. These numbers included 447 (75.5%) benign, 123 (20.8%) malignant, and 22 (3.7%) borderline tumors, confirmed by postoperative pathologic

diagnoses. An analysis of the histological findings showed that the most frequent benign tumor was mature teratoma, while the most common malignant tumor was serous adenocarcinoma. The details are shown in Table 2.

The detailed outcomes of the two groups regarding the classification of the 592 adnexal masses using the four systems are demonstrated in Fig. 2. In the O-RADS system, groups I and II classified 379 (64.0%) and 388 (65.5%) cases as benign tumors, respectively, while 213 (36.0%) and 204 (34.5%) cases were classified as malignant. The diagnostic malignancy rates of O-RADS grades 1 to 5 for 592 adnexal masses were 0% (0/1), 2.0% (6/299), 8.9% (7/79), 52.3% (80/153), and 86.7% (52/60) in group I and 0% (0/4), 2.8% (8/283), 13.9% (14/101), 42.9% (48/112), and 81.5% (75/92) in group II, respectively. Using the RMI4 system to classify the adnexal lesions, groups I and II classified 483 (81.6%) and 491 (82.9%) cases as benign masses and 109 (18.4%) and 101 (17.1%) cases as malignant masses, respectively. There were 426 (72.0%) and 397 (67.1%) benign cases and 166 (28.0%) and 195 (32.9%) malignant cases, respectively, using the IOTA LR2 system. Groups I and II applied the IOTA SR and classified 442 (74.7%) and 410 (69.3%) cases as benign tumors, and 150 (25.3%) and 182 (30.7%) cases as malignant tumors. In addition, using the IOTA SR system, groups I and II classified 48 (8.1%) and 79 (13.1%) cases as

Table 1. Demographic and clinical characteristics of the 575 women

Characteristic	Women with benign lesions (n=433)	Women with malignant tumors (n=142)	P-value ^{a)}
Age (year)	36.6±13.9	46.5±13.9	<0.001 ^{b)}
Postmenopausal status			
Yes	66 (15.2)	63 (44.4)	<0.001
No	367 (84.8)	79 (55.6)	
Bilateral involvement			
Yes	72 (16.6)	46 (32.4)	<0.001
No	361 (83.4)	96 (67.6)	
CA125			
Increased	81 (18.7)	104 (73.2)	<0.001
Normal	352 (81.3)	38 (26.8)	<0.001

Values are presented as a mean±SD or number (%). CA125, cancer antigen 125; SD, standard deviation. ^{a)}Chi-square test. ^{b)}Wilcoxon rank-sum test.

Table 2. Pathological diagnoses of the 592 adnexal masses

Pathology	No. (%)
Benign masses	447 (75.5)
Mature teratoma	203 (34.3)
Ovarian endometrial cyst	73 (12.3)
Serous cystadenoma and mucinous cystadenoma	63 (10.6)
Ovarian cyst and embryonic residual cyst	43 (7.3)
Adnexal inflammatory mass	42 (7.1)
Other benign tumors	17 (2.9)
Benign mixture ovarian tumor	6 (1.0)
Malignant masses	145 (24.5)
Serous cystadenocarcinoma and mucinous cystadenocarcinoma	80 (13.5)
Other malignant tumors	23 (4.0)
Borderline cystadenoma	21 (3.5)
Endometrioid adenocarcinoma	6 (1.0)
Yolk sac tumor	6 (1.0)
Immature teratoma and Malignant transformation of mature cystic teratoma	5 (0.8)
Dysgerminoma	2 (0.3)
Malignant mixed germ cell tumor	1 (0.2)
Borderline Brenner tumor	1 (0.2)

indeterminate lesions, respectively. The malignancy rates of the IOTA SR for benign, malignant, and uncertain groups were 5.8% (24/417), 78.0% (99/127), and 45.8% (22/48) in group I and 5.6% (20/355), 66.5% (105/158), and 25.3% (20/79) in group II, respectively.

The ROC curves for each system in the two groups are shown in Fig. 3. The O-RADS had the highest area under the curve (AUC), with 0.90 in group I and 0.89 in group II. The IOTA SR had the

lowest AUC, with 0.86 in group I and 0.84 for group II (Table 3). The sensitivity, specificity, Youden index, positive predictive value (PPV), and negative predictive value for the four systems are presented in Table 3. Of the four systems, the O-RADS had the highest Youden index (0.73 in group I) and the highest sensitivity (0.91 and 0.85 in groups I and II), respectively.

The two groups had moderate agreement ($\kappa=0.71$ and $\kappa=0.68$,

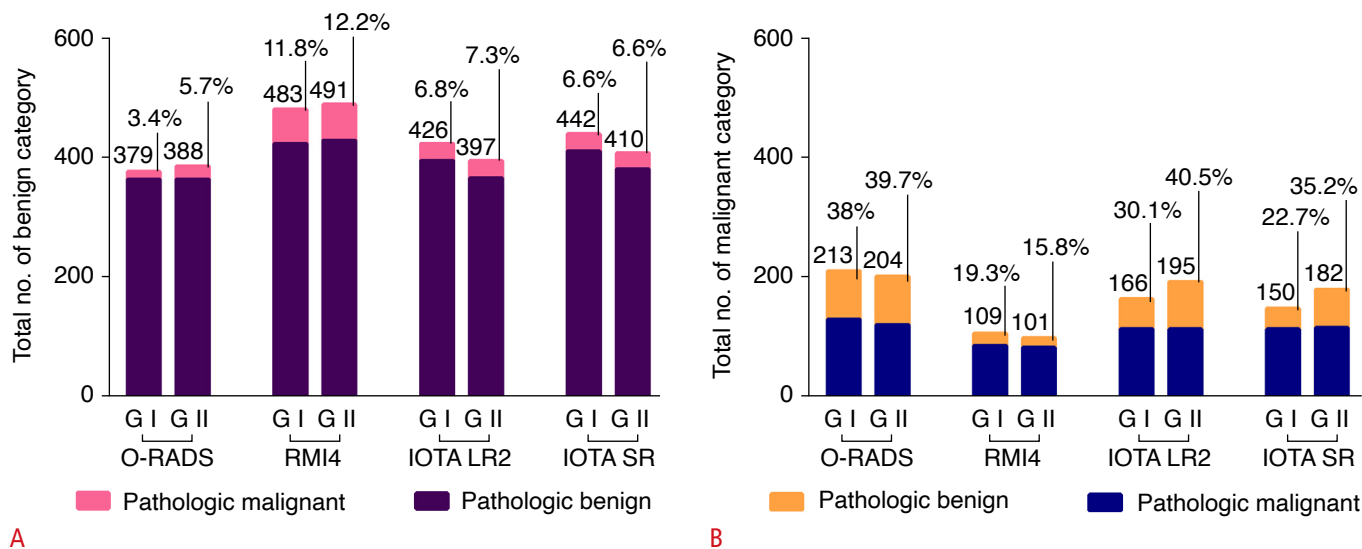


Fig. 2. Frequency distributions and malignancy rates of the benign and malignancy grades of 592 adnexal masses of groups I and II for the four systems.

A. Frequency distributions of the masses with benign classifications of groups I and II for the four systems is shown. **B.** Frequency distributions of the masses with malignant classifications of groups I and II for the four systems is shown. O-RADS, Ovarian-Adnexal Reporting and Data System; RMI4, Risk of Malignancy Index 4; IOTA, International Ovarian of Tumor Analysis; IOTA LR2, IOTA Logistic Regression Model 2; IOTA SR, IOTA Simple Rules; G I, group I; G II, group II.

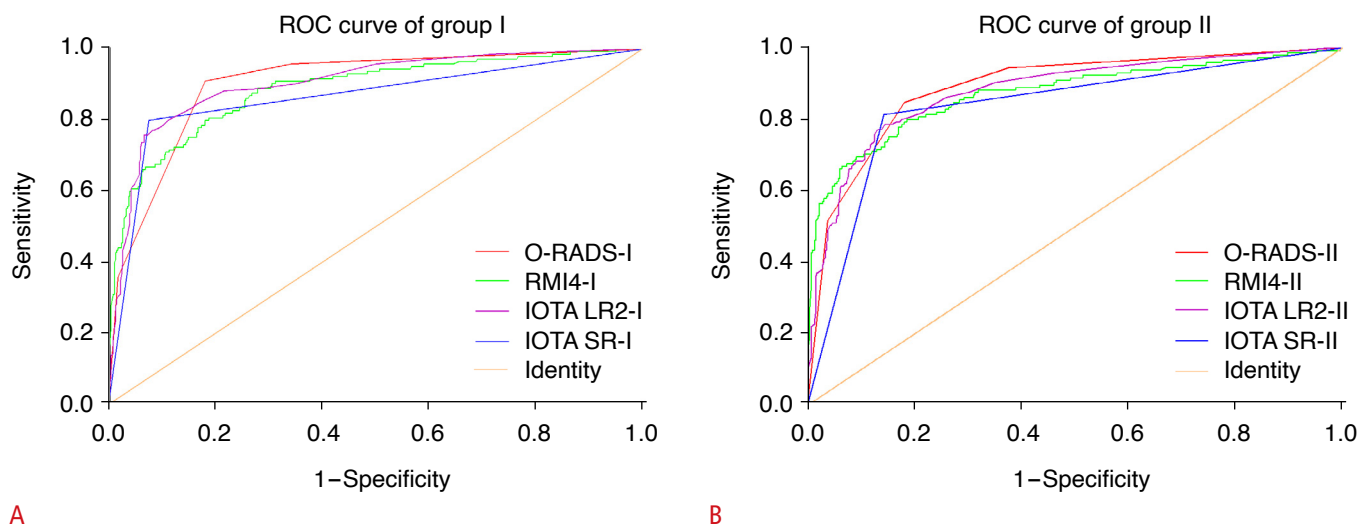


Fig. 3. The ROC curves of the four ultrasound classification systems used in groups I (A) and II (B). ROC, receiver operating characteristic; O-RADS, Ovarian-Adnexal Reporting and Data System; RMI4, Risk of Malignancy Index 4; IOTA, International Ovarian of Tumor Analysis; IOTA LR2, IOTA Logistic Regression Model 2; IOTA SR, IOTA Simple Rules.

Table 3. The diagnostic validity and consistency analysis of the four ultrasound classification systems in the two groups using consensus data

Category	AUC (95% CI)	Youden index	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Kappa (95% CI)
O-RADS							
Group I	0.90 (0.87–0.93)	0.73	91.0 (84.9–94.9)	81.9 (77.9–85.3)	62.0 (55.1–68.4)	96.6 (94.1–98.1)	0.71 (0.64–0.77)
Group II	0.89 (0.86–0.92)	0.67	84.8 (77.7–90.0)	81.9 (77.9–85.3)	60.3 (53.2–67.0)	94.3 (91.4–96.3)	
RMI4							
Group I	0.89 (0.85–0.92)	0.56	60.7 (52.2–68.6)	95.3 (92.8–97.0)	80.7 (71.8–87.4)	88.2 (84.9–90.9)	0.92 (0.87–0.96)
Group II	0.87 (0.83–0.91)	0.55	58.6 (50.1–66.6)	96.4 (94.1–97.9)	84.2 (75.2–90.4)	87.8 (84.5–90.5)	
IOTA LR2							
Group I	0.90 (0.87–0.93)	0.69	80.0 (72.4–86.0)	88.8 (85.4–91.5)	69.9 (62.2–76.6)	93.2 (90.3–95.3)	0.68 (0.61–0.74)
Group II	0.88 (0.85–0.91)	0.62	80.0 (72.4–86.0)	82.3 (78.4–85.7)	59.5 (52.2–66.4)	92.7 (89.6–95.0)	
IOTA SR							
Group I	0.86 (0.82–0.90)	0.72	80.0 (72.4–86.0)	92.4 (89.4–94.6)	77.3 (69.6–83.6)	93.4 (90.6–95.5)	0.77 (0.71–0.82)
Group II	0.84 (0.79–0.88)	0.67	81.4 (73.9–87.2)	85.7 (82.0–88.7)	64.8 (57.4–71.7)	93.4 (90.4–95.5)	

AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; O-RADS, Ovarian-Adnexal Reporting and Data System; RMI4, Risk of Malignancy Index 4; IOTA, International Ovarian Tumor Analysis; IOTA LR2, IOTA Logistic Regression Model 2; IOTA SR, IOTA Simple Rules.

respectively) in using the O-RADS and IOTA LR2 systems and high agreement ($\kappa=0.92$ and $\kappa=0.77$, respectively) for the RMI4 and IOTA SR systems (Table 3).

Discussion

Many ultrasound systems for diagnosing adnexal masses have been launched internationally, some of which have undergone prospective or retrospective external validation [20–25]. This study focused on the comparison between the latest proposed O-RADS and other validated classification systems. In the present study, the good to excellent inter-reader agreement in each group ensured consistency of understanding and using the systems and controlled the confounding factors caused by different interpretations of the terms. The present analysis of a large sample of 592 adnexal masses had reliable outcomes.

The results showed that the four systems all had excellent AUCs for the diagnosis of adnexal masses. The RMI4 system had the simplest ultrasound rules, with excellent intergroup agreement ($\kappa=0.92$). It had a higher AUC, but the lowest Youden index. The CA125 level is one of the strongest indicators of malignancy in the RMI4 system, but might also increase in women with ovarian endometrial cyst and pelvic inflammatory disease. It is plausible that an ovarian endometrial cyst or inflammatory mass had a high RMI score and would be misdiagnosed as a malignant mass. Hence, for the proposed cutoff score of 450, the diagnostic efficiency should be affected by sample bias. Since ovarian endometrial cysts and inflammatory masses accounted for 19.4% of the sample, it was not

surprising that the RMI4 system had the lowest sensitivity but the highest specificity in this study.

The malignancy rates of the IOTA SR benign, malignant, and uncertain groups were consistent with the recommended values in the previous literature [2]. However, the malignancy rate of groups I and II was as high as 45.8% and 25.3% in the indeterminate group, respectively, suggesting that inexperienced doctors might not correctly diagnose these tumors if they cannot obtain assistance from other diagnostic models or obtain a consultation. The uncertain group is the most obvious shortcoming of the IOTA SR. Although the IOTA LR2 could be applied to classify all adnexal masses, its sensitivity in detecting malignant masses was not greatly improved.

The O-RADS system had the highest AUC and Youden index. At the cost of decreased specificity, its detailed explanations of characteristics and descriptions of benign and malignant lesions ensured the highest sensitivity in detecting malignant masses. However, the simple diagnostic indices involved in the RMI4, IOTA SR, and IOTA LR2 systems easily misdiagnosed some tumors without typical malignant features. The O-RADS could be used to identify actual malignant lesions as much as possible to reduce the severe consequences of missing diagnoses. This advantage corresponds to an important capability of a malignant tumor predictive model, because discovering a possibly malignant lesion is the primary step for patients with adnexal masses. For patients with a high suspicion of malignancy, the O-RADS proposes the following management recommendations. Subsequent examinations and clinical measures are advised for these patients with suspected malignant lesions [17]. A magnetic resonance imaging examination or ultrasound

expert consultation should be arranged for patients with suspected malignant lesions. Patients with a high suspicion of malignancy should be referred to a gynecologic oncologist and treated in a timely manner. In contrast, the other three diagnostic models do not provide corresponding management measures to identify false-negative patients.

The finding of good intergroup agreement showed that the four diagnostic systems could compensate for junior doctors' inexperience to some extent. However, the intergroup agreement values for the O-RADS and IOTA systems were much lower than those for the RMI4 system, which included the simplest image parameters. Experience is needed to ensure a better understanding and application of the detailed definitions of diagnostic signs in practice. Artificial intelligence may help to resolve the issue of a long learning curve.

There are many limitations of this study. First, this retrospective study could not obtain dynamic images to evaluate each adnexal mass sufficiently, leading to misjudgments of certain ultrasound features. Second, the low malignancy rate (24.5%) in the present study sample may account for the lower specificity and PPV of the O-RADS. In previous studies, the malignancy rate was 27.5% to 28.8% [2,26]. Third, prospective studies are needed to further test the performance of the management recommendations.

In conclusion, to a certain extent, all four diagnostic systems could compensate for junior doctors' inexperience in the diagnosis of adnexal masses. The O-RADS performed best and had the highest sensitivity for detecting malignant lesions. It may make sense to use the O-RADS for clinical diagnosis and therapy.

ORCID: Yuyang Guo: <https://orcid.org/0000-0001-5630-6912>; Baihua Zhao: <https://orcid.org/0000-0001-7185-292X>; Shan Zhou: <https://orcid.org/0000-0001-5795-8277>; Lieming Wen: <https://orcid.org/0000-0003-3283-888X>; Jieyu Liu: <https://orcid.org/0000-0003-4203-9985>; Yaqian Fu: <https://orcid.org/0000-0003-3683-304X>; Fang Xu: <https://orcid.org/0000-0002-2326-6288>; Minghui Liu: <https://orcid.org/0000-0002-6548-3264>

Author Contributions

Conceptualization: Zhao B, Wen L, Liu M. Data acquisition: Guo Y, Zhao B, Zhou S, Liu J, Fu Y, Xu F. Data analysis or interpretation: Guo Y, Zhao B, Zhou S, Wen L. Drafting of the manuscript: Guo Y, Zhao B, Zhou S, Liu J, Fu Y, Xu F. Critical revision of the manuscript: Guo Y, Zhao B, Wen L, Liu M. Approval of the final version of the manuscript: all authors.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

The authors thank Professor Jiang Ouyang from the Department of Public Health, Changsha Medical College, for assisting in statistical guidance.

References

- Konecny J, Czekierdowski A, Florczak M, Poziemski P, Stachowicz N, Borowski D. The use of sonographic subjective tumor assessment, IOTA logistic regression model 1, IOTA Simple Rules and GI-RADS system in the preoperative prediction of malignancy in women with adnexal masses. *Ginekol Pol* 2017;88:647-653.
- Basha MA, Metwally MI, Gamil SA, Khater HM, Aly SA, El Sammak AA, et al. Comparison of O-RADS, GI-RADS, and IOTA simple rules regarding malignancy rate, validity, and reliability for diagnosis of adnexal masses. *Eur Radiol* 2021;31:674-684.
- Kaijser J, Bourne T, Valentin L, Sayasneh A, Van Holsbeke C, Vergote I, et al. Improving strategies for diagnosing ovarian cancer: a summary of the International Ovarian Tumor Analysis (IOTA) studies. *Ultrasound Obstet Gynecol* 2013;41:9-20.
- Ning CP, Ji X, Wang HQ, Du XY, Niu HT, Fang SB. Association between the sonographer's experience and diagnostic performance of IOTA simple rules. *World J Surg Oncol* 2018;16:179.
- Sayasneh A, Kaijser J, Preisler J, Johnson S, Stalder C, Husicka R, et al. A multicenter prospective external validation of the diagnostic performance of IOTA simple descriptors and rules to characterize ovarian masses. *Gynecol Oncol* 2013;130:140-146.
- Nunes N, Yazbek J, Ambler G, Hoo W, Naftalin J, Jurkovic D. Prospective evaluation of the IOTA logistic regression model LR2 for the diagnosis of ovarian cancer. *Ultrasound Obstet Gynecol* 2012;40:355-359.
- Meys E, Rutten I, Kruitwagen R, Slangen B, Lambrechts S, Mertens H, et al. Simple rules, not so simple: the use of International Ovarian Tumor Analysis (IOTA) terminology and simple rules in inexperienced hands in a prospective multicenter cohort study. *Ultraschall Med* 2017;38:633-641.
- Dakhly DM, Gaafar HM, Sediek MM, Ibrahim MF, Momtaz M. Diagnostic value of the International Ovarian Tumor Analysis (IOTA) simple rules versus pattern recognition to differentiate between malignant and benign ovarian masses. *Int J Gynaecol Obstet* 2019;147:344-349.
- Alcazar JL. Ultrasound-based IOTA simple rules allow accurate malignancy risk estimation for adnexal masses. *Evid Based Med* 2016;21:197.
- Nunes N, Ambler G, Foo X, Widschwendter M, Jurkovic D. Prospective evaluation of IOTA logistic regression models LR1 and LR2 in comparison with subjective pattern recognition for diagnosis of ovarian cancer in an outpatient setting. *Ultrasound Obstet Gynecol* 2018;51:829-835.

11. Hidalgo JJ, Llueca A, Zolfaroli I, Veiga N, Ortiz E, Alcazar JL. Comparison of IOTA three-step strategy and logistic regression model LR2 for discriminating between benign and malignant adnexal masses. *Med Ultrason* 2021;23:168-175.
12. Jacobs I, Oram D, Fairbanks J, Turner J, Frost C, Grudzinskas JG. A risk of malignancy index incorporating CA 125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. *Br J Obstet Gynaecol* 1990;97:922-929.
13. Yamamoto Y, Yamada R, Oguri H, Maeda N, Fukaya T. Comparison of four malignancy risk indices in the preoperative evaluation of patients with pelvic masses. *Eur J Obstet Gynecol Reprod Biol* 2009;144:163-167.
14. Dora SK, Dandapat AB, Pande B, Hota JP. A prospective study to evaluate the risk malignancy index and its diagnostic implication in patients with suspected ovarian mass. *J Ovarian Res* 2017;10:55.
15. Zhang S, Yu S, Hou W, Li X, Ning C, Wu Y, et al. Diagnostic extended usefulness of RMI: comparison of four risk of malignancy index in preoperative differentiation of borderline ovarian tumors and benign ovarian tumors. *J Ovarian Res* 2019;12:87.
16. Hada A, Han LP, Chen Y, Hu QH, Yuan Y, Liu L. Comparison of the predictive performance of risk of malignancy indexes 1-4, HE4 and risk of malignancy algorithm in the triage of adnexal masses. *J Ovarian Res* 2020;13:46.
17. Andreotti RF, Timmerman D, Strachowski LM, Froyman W, Benacerraf BR, Bennett GL, et al. O-RADS US risk stratification and management system: a consensus guideline from the ACR Ovarian-Adnexal Reporting and Data System Committee. *Radiology* 2020;294:168-185.
18. Hiett AK, Sonek J, Guy M, Reid TJ. 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* 2022;59:668-676.
19. Heintz AP, Odicino F, Maisonneuve P, Quinn MA, Benedet JL, Creasman WT, et al. Carcinoma of the ovary. FIGO 26th Annual Report on the Results of Treatment in Gynecological Cancer. *Int J Gynaecol Obstet* 2006;95 Suppl 1:S161-S192.
20. Qian L, Du Q, Jiang M, Yuan F, Chen H, Feng W. Comparison of the diagnostic performances of ultrasound-based models for predicting malignancy in patients with adnexal masses. *Front Oncol* 2021;11:673722.
21. Lai HW, Lyu GR, Kang Z, Li LY, Zhang Y, Huang YJ. Comparison of O-RADS, GI-RADS, and ADNEX for diagnosis of adnexal masses: an external validation study conducted by junior sonologists. *J Ultrasound Med* 2022;41:1497-1507.
22. Froyman W, Timmerman D. Methods of assessing ovarian masses: international ovarian tumor analysis approach. *Obstet Gynecol Clin North Am* 2019;46:625-641.
23. Pi Y, Wilson MP, Katlariwala P, Sam M, Ackerman T, Paskar L, et al. 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)* 2021;46:4967-4973.
24. Wong VK, Kundra V. Performance of O-RADS MRI score for classifying indeterminate adnexal masses at US. *Radiol Imaging Cancer* 2021;3:e219008.
25. Meys EM, Jeelof LS, Achten NM, Slangen BF, Lambrechts S, Kruitwagen R, et al. Estimating risk of malignancy in adnexal masses: external validation of the ADNEX model and comparison with other frequently used ultrasound methods. *Ultrasound Obstet Gynecol* 2017;49:784-792.
26. Cao L, Wei M, Liu Y, Fu J, Zhang H, Huang J, et al. Validation of American College of Radiology Ovarian-Adnexal Reporting and Data System Ultrasound (O-RADS US): analysis on 1054 adnexal masses. *Gynecol Oncol* 2021;162:107-112.