# Systematic Review and Meta-Analysis of O-RADS Ultrasound and O-RADS MRI for Risk Assessment of Ovarian and Adnexal Lesions

Qing Zhang, MD[1], Xiaoli Dai, MD[1], Wei Li, MD[2]

**Evidence Synthesis and Decision Analysis · Systematic Review/Meta-Analysis**

**Keywords**

adnexal masses, MRI, O-RADS, systematic review, ultrasound

Submitted: Aug 9, 2022
Revision requested: Aug 30, 2022
Revision received: Dec 19, 2022
Accepted: Jan 23, 2023
First published online: Feb 1, 2023
Version of record: May 3, 2023

An electronic supplement is available online at doi.org/10.2214/AJR.22.28396.

Q. Zhang and X. Dai contributed equally to this work.

The authors declare that there are no disclosures relevant to the subject matter of this article.

**BACKGROUND.** O-RADS ultrasound (US) and O-RADS MRI have been developed to standardize risk stratification of ovarian and adnexal lesions.

**OBJECTIVE.** The purpose of this study was to perform a meta-analysis evaluating the diagnostic performance of O-RADS US and O-RADS MRI for risk stratification of ovarian and adnexal lesions.

**EVIDENCE ACQUISITION.** We searched the Web of Science, PubMed, Cochrane Library, Embase, and Google Scholar databases from January 1, 2020, until October 31, 2022, for studies reporting on the performance of O-RADS US or O-RADS MRI in the diagnosis of malignancy of ovarian or adnexal lesions. Study quality was assessed with QUADAS-2. A hierarchic summary ROC model was used to estimate pooled sensitivity and specificity. Heterogeneity was assessed with the *Q* statistic. Metaregression analysis was performed to explore potential sources of heterogeneity. O-RADS US was compared with the International Ovarian Tumor Analysis (IOTA) simple rules and Assessment of Different Neoplasias in the Adnexa (ADNEX) model in studies providing head-to-head comparisons.

**EVIDENCE SYNTHESIS.** Twenty-six studies comprising 9520 patients were included. O-RADS US was evaluated in 15 and O-RADS MRI in 12 studies; both systems were evaluated in one of the studies. Quality assessment revealed that risk of bias or concern about applicability most commonly related to patient selection. Pooled sensitivity and specificity of O-RADS US were 95% (95% CI, 91–97%) and 82% (95% CI, 76–87%) and of O-RADS MRI were 95% (95% CI, 92–97%) and 90% (95% CI, 84–94%). Analysis with the *Q* statistic revealed significant heterogeneity among studies of O-RADS US in both sensitivity and specificity (both $p < .001$) and among studies of O-RADS MRI in specificity ($p < .001$) but not sensitivity ($p = .07$). In metaregression, no factor was significantly associated with sensitivity or specificity of either system (all $p > .05$). O-RADS US showed no significant difference in sensitivity or specificity versus IOTA simple rules in four studies (sensitivity, 96% vs 93%; specificity, 76% vs 82%) or versus the ADNEX model in three studies (sensitivity, 96% vs 96%; specificity, 79% vs 78%).

**CONCLUSION.** O-RADS US and O-RADS MRI both have high sensitivity for ovarian or adnexal malignancy. O-RADS MRI, but not O-RADS US, also has high specificity.

**CLINICAL IMPACT.** Awareness of the diagnostic performance results regarding O-RADS US and O-RADS MRI will be helpful as these systems are increasingly implemented into clinical practice.

Ovarian cancer is the leading cause of death among gynecologic malignancies [1]. Among patients in the United States, ovarian cancer has a 5-year survival rate of 47% [2]. Imaging of the adnexa can help detect malignant masses to prompt early surgery with the aim of improving survival. This goal must be balanced with efforts to limit surgery and other invasive interventions for benign lesions. Thus, reliable differentiation of benign from malignant ovarian and adnexal lesions is essential to guide selection of optimal treatment strategies [3].

Imaging evaluation of ovarian and adnexal lesions is performed primarily with ultrasound (US). However, radiologists' experience and expertise in pelvic US varies widely [4]. Accordingly, several guidelines or scoring systems have been proposed to standardize radiologists' reporting of ovarian and adnexal lesions detected with US. These include the International Ovarian Tumor Analysis (IOTA) simple rules [5], the Assessment of Different Neoplasias in the Adnexa (ADNEX) model developed by the IOTA Group [6], and

ARRS is accredited by the Accreditation Council for Continuing Medical Education (ACCME) to provide continuing medical education activities for physicians.

The ARRS designates this journal-based CME activity for a maximum of 1.00 AMA PRA Category 1 Credit™. Physicians should claim only the credit commensurate with the extent of their participation in the activity.

To access the article for credit, follow the prompts associated with the online version of this article.

doi.org/10.2214/AJR.22.28396

*AJR* 2023; 221:21–33

ISSN-L 0361–803X/23/2211–21



[1]Department of Clinical Medicine, Jiangsu Vocational College of Medicine, Yancheng, China.

[2]Department of Medical Imaging, Jiangsu Vocational College of Medicine, W Jianjun Rd, Yancheng, 224000, China. Address correspondence to W. Li (hfjs2000@126.com).

the Gynecologic Imaging Reporting and Data System (GI-RADS) [7]. In 2019, the American College of Radiology (ACR) published O-RADS US to guide risk stratification and management of ovarian and adnexal masses [8]. O-RADS US combines lexicon descriptors of US features with a series of risk stratification governing concepts to assign all adnexal masses an O-RADS category that expresses the risk of malignancy of the mass on a 1–5 scale (1, physiologic; 2, almost certainly benign; 3, low risk of malignancy; 4, intermediate risk of malignancy; 5, high risk of malignancy).

A substantial fraction of adnexal lesions cannot be definitively stratified on the basis of US features [9–11]. Thus, some adnexal lesions present a diagnostic challenge in US regardless of the system used for evaluation. MRI has, therefore, been investigated for the assessment of adnexal lesions that are indeterminate on US. MRI affords excellent soft-tissue contrast and a range of tissue contrast mechanisms, helping to improve lesion characterization [12, 13]. In 2013, Thomassin-Naggara et al. [14] introduced the ADNEX MR 5-point scoring system for the MRI classification of adnexal masses that are indeterminate on US; this system has been validated in two multicenter cohort studies [15, 16]. ADNEX MR served as the basis for the initial version of O-RADS MRI, which was released by the ACR in 2019 [15].

Though other imaging classification systems for ovarian lesions have been proposed, O-RADS is unique because of its dual US and MRI arms with a common lexicon and data-supported stratification of malignancy risk. To our knowledge, the diagnostic performance of O-RADS US [17] and O-RADS MRI [18] has been summarized in a single meta-analysis each. Additional studies continue to emerge. The aim of this meta-analysis was to evaluate the diagnostic performance of both O-RADS US and O-RADS MRI in risk stratification of ovarian and adnexal lesions.

## Evidence Acquisition

The findings of this systematic review and meta-analysis are reported in accordance with the PRISMA recommendations [19]. The aim of the meta-analysis was to assess the performance of O-RADS US and O-RADS MRI for the diagnosis of malignancy of ovarian and adnexal lesions.

### Highlights

**Key Finding**

- *Among 26 studies comprising 9520 patients (10,190 masses), O-RADS US had pooled sensitivity of 95% (95% CI, 91–97%) and pooled specificity of 82% (95% CI, 76–87%). O-RADS MRI had pooled sensitivity of 95% (95% CI, 92–97%) and pooled specificity of 90% (95% CI, 84–94%).*

**Importance**

- *O-RADS US and O-RADS MRI both have high sensitivity for malignancy; O-RADS MRI also has high specificity.*

### *Search Strategy and Study Selection*

A systematic search was performed of the PubMed, Embase, Cochrane Library, Web of Science, and Google Scholar databases from January 1, 2020, to October 31, 2022, for articles describing the use of O-RADS US or O-RADS MRI for the evaluation of ovarian or adnexal lesions. The specific search terms were as follows: ([ovarian] OR [adnexal] OR [pelvic]) AND ([cancer] OR [tumor] OR [mass*] OR [lesion*]) AND ([MR*] OR [MRI] OR [US] OR [ultrasound] OR [sonography]) OR ([O-RADS] OR [Ovarian-Adnexal Reporting and Data System]). The search was restricted to articles published in English. The bibliographies of relevant review articles and of included articles were manually screened for potential additional articles.

Studies identified in the literature search were assessed independently by two radiologists (Z.Q. with 6 and X.D. with 8 years of experience in performing systematic reviews and meta-analyses). Disagreements were resolved by discussion with a third radiologist (L.W., who had 12 years of experience in performing

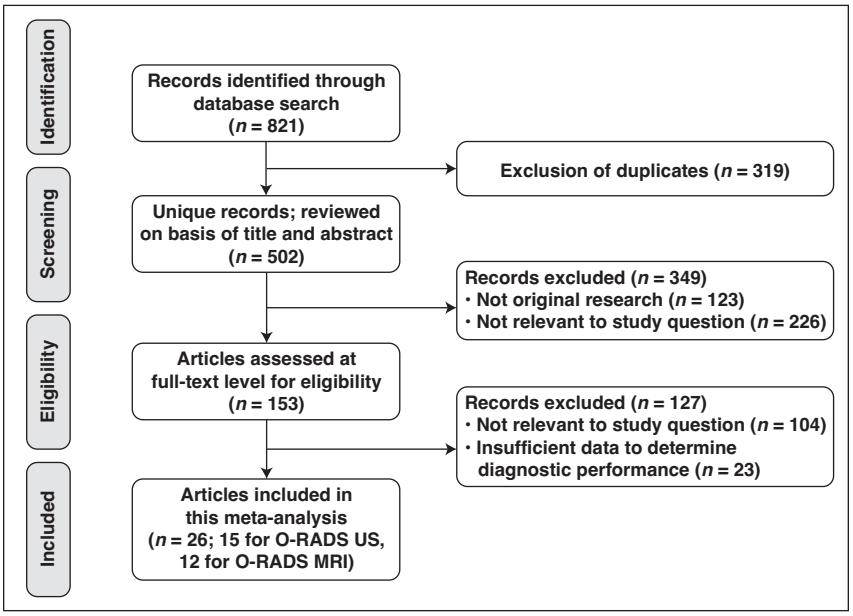



**Fig. 1**—Chart shows study selection process. Both O-RADS ultrasound (US) and O-RADS MRI were evaluated in one study.

**TABLE 1: Clinical Characteristics**

| First Author | O-RADS Arm | Country | Publication Year | No. of Patients | No. of Lesions | | | Patient Age (y)[a] | Lesion Size (mm)[a] | Reference Standard |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Total | Malignant | Borderline | | | |
| Ahmed [40] | US | Egypt | 2021 | 50 | 50 | 34 | NA | 41.8 ± 3.64 | NA | Pathology and 6- to 12-mo imaging follow-up |
| Basha [26] | US | Egypt | 2021 | 609 | 647 | 178 | 32 | 48 ± 13.7 | NA | Pathology and 2-y imaging follow-up |
| Cao [27] | US | China | 2021 | 1054 | 1054 | 304 | 31 | 37.7 ± 10.5<br>53.4 ± 11.4 | NA | Pathology |
| Chen [29] | US | China | 2022 | 85 | 85 | 24 | NA | 46.4 ± 14.8 | 58.0 (42.0–81.3)<br>52.0 (39.0–80.0) | Pathology |
| Chen [28] | US | Taiwan | 2022 | 322 | 322 | 264 | 8 | 44 (20–83) | NA | Pathology |
| Guo [30] | US | China | 2022 | 575 | 592 | 145 | 22 | 36.6 ± 13.9<br>46.5 ± 13.9 | NA | Pathology |
| Guo [38] | US, MRI | United States | 2022 | 54 | 58 | 8 | NA | 37 ± 12 | 7.6 ± 3.2 | Pathology and imaging follow-up (mean, 2.3 y) |
| Hack [31] | US | Canada | 2022 | 227 | 262 | 75 | 18 | 52 ± 16 | 54 (38–84) | Pathology and 2-y imaging follow-up |
| Hiett [32] | US | United States | 2021 | 150 | 150 | 40 | 12 | 48.2 ± 1.7<br>47.5 ± 3.1 | 83.7 ± 5.2<br>101.4 ± 7.7 | Pathology |
| Jha [36] | US | United States | 2022 | 913 | 1014 | 85 | NA | 42.4 ± 13.9 | NA | Pathology and 2-y imaging follow-up |
| Lai [35] | US | China | 2021 | 734 | 734 | 170 | 69 | 35 (29–46)<br>48 (36–55.3) | NA | Pathology |
| Solis Cano [39] | US | Mexico | 2021 | 73 | 73 | 23 | NA | 42 ± 11 | NA | Pathology |
| Wang [34] | US | China | 2022 | 431 | 431 | 173 | 31 | 40.2 ± 16<br>52.1 ± 13.2 | NA | Pathology |
| Wu [37] | US | China | 2022 | 443 | 443 | 131 | 47 | 35.3 | 68–135 | Pathology |
| Xie [33] | US | China | 2022 | 453 | 453 | 269 | 48 | 48.8 ± 13.4 | 105.9 ± 64 | Pathology |
| Aslan [41] | MRI | Turkey | 2021 | 200 | 237 | 28 | 2 | 56.3 ± 17.5 | 46.9 ± 12.4<br>80.1 ± 9.8 | Pathology and 2-y imaging follow-up |
| Assouline [47] | MRI | France | 2022 | 585 | 779 | 173[b] | 38 | 46.5 | NA | Pathology |
| Bang [48] | MRI | South Korea | 2022 | 307 | 331[c] | 178 | 54 | 50.8 ± 15.7 | NA | Pathology |

**(Table 1 continues on next page)**

**TABLE 1: Clinical Characteristics (continued)**

| First Author | O-RADS Arm | Country | Publication Year | No. of Patients | No. of Lesions | | | Patient Age (y)[a] | Lesion Size (mm)[a] | Reference Standard |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Total | Malignant | Borderline | | | |
| Basu [45] | MRI | India | 2022 | 42 | 46 | 13 | 6 | 35.9 (10–75) | NA | Pathology and 4-mo imaging follow-up |
| Crestani [46] | MRI | France | 2020 | 26 | 26 | 9 | 3 | 51 ± 18 | NA | Pathology |
| Elshetry [44] | MRI | Egypt | 2022 | 90 | 116 | 45 | 5 | 39.4 ± 13.8 | 60 ± 45<br>110 ± 82 | Pathology and 6-mo to 1-y imaging follow-up |
| Hottat [42] | MRI | Belgium | 2021 | 163 | 201 | 58 | 16 | 51 ± 17 | NA | Pathology |
| Pereira [49] | MRI | Brazil | 2022 | 243 | 287 | 90 | 17 | 47.1 ± 14.9<br>57.8 ± 13.2 | 91.1 ± 65<br>91.6 ± 65 | Pathology and 1-y imaging follow-up |
| Sahin [43] | MRI | United Kingdom | 2021 | 291 | 350 | 53 | 0 | 46.8 ± 15.4 | 62 (42–94) | Pathology and 1-y imaging follow-up |
| Thomassin-Naggara [15] | MRI | Multiple[d] | 2020 | 1130 | 1130 | 203 | 45 | 49 (18–96) | NA | Pathology and 2-y imaging follow-up |
| Wengert [16] | MRI | Multiple[d] | 2022 | 244 | 320 | 207 | 30 | 55.3 ± 15.8 | 73.8 ± 53 | Pathology |

Note—Numbers in brackets denote references. US = ultrasound, NA = not available.
[a]Patient age and lesion size are mean ± SD or median and range. Where two values are reported, first value is for benign lesions, and second value is for malignant lesions.
[b]Includes 135 invasive lesions and 38 borderline lesions.
[c]MRI performed for 110 lesions; others were evaluated with contrast-enhanced CT and/or FDG PET/CT.
[d]Austria, Croatia, France, Italy, Portugal, Serbia, Switzerland, and United Kingdom.

systematic reviews and meta-analyses). Duplicate search results were excluded. Article titles and abstracts were then reviewed to exclude articles not reporting an original research study (e.g., letter, editorial, abstract, or review article) or not relevant to the study question. The remaining articles were reviewed at the full-text level and excluded if not relevant to the study question or if providing insufficient data to construct a 2 × 2 contingency table to determine diagnostic performance with respect to a reference standard of pathology or the combination of pathology and follow-up imaging when pathology was unavailable.

### Data Extraction

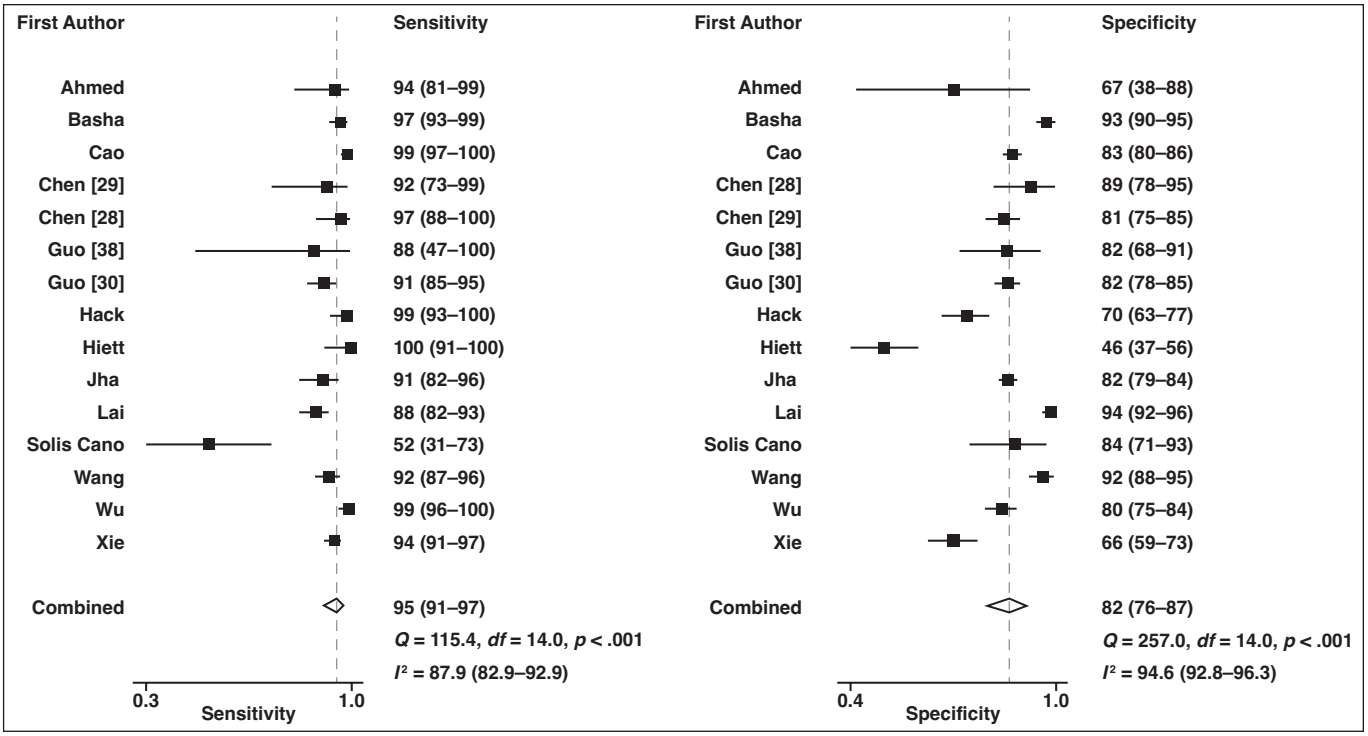
The two previously noted investigators (Z.Q., X.D.) performed data extraction and resolved disagreements by discussion with the third investigator (L.W.). A predefined standardized form was used to extract relevant information from studies in the final analysis, including clinical characteristics (country where study was conducted, number of patients, number of adnexal lesions, number of malignant lesions, number of borderline lesions, patient age, lesion size, and reference standard) and study characteristics (publication year, study design, level of analysis, time period for patient selection, number of readers, reader experience, blinding of readers to the reference standard, technical characteristics [e.g., US mode, MRI field strength, and MRI sequences], interreader agreement [measured with kappa values], and cutoff O-RADS category for malignancy). The number of true-positive, false-positive, false-negative, and true-negative cases for the diagnosis of malignancy at the cutoff O-RADS category in the study were recorded. For studies in which a head-to-head comparison of O-RADS US with either the IOTA simple rules or the ADNEX model was performed, the 2 × 2 contingency data were also extracted for the other system.

### Quality Assessment

The two investigators (Z.Q., X.D.) independently conducted the quality assessment, consulting the third investigator (L.W.) about disagreements. QUADAS-2 was used to evaluate study quality in terms of four domains [20]: patient selection, index test, reference standard, and flow and timing. Each study was scored as showing high, low, or unclear risk of bias for each domain and high, low, or unclear concern about applicability for all domains except flow and timing. The quality of studies that included a head-to-head comparison of O-RADS US with either the IOTA simple rules or the ADNEX model was assessed with QUADAS-Comparative (QUADAS-C), an extension of QUADAS-2 designed for comparative diagnostic performance studies [21].

### Data Synthesis and Statistical Analysis

The sensitivity and specificity of each study were computed on the basis of extracted 2 × 2 contingency data. The hierarchic summary ROC (HSROC) model was used to estimate pooled sensitivity and pooled specificity, with 95% CIs, of both O-RADS US and O-RADS MRI for the detection of malignancy [22, 23]. Forest plots and HSROC curves were constructed to present the results for each system graphically. Publication bias for each system was estimated with Deeks funnel plots. Statistical significance of the plot was assessed with the Deeks asymmetry test, $p < .05$ indicating publication bias [24]. Pooled sensitivity and pooled specificity were also determined for the IOTA simple rules and ADNEX MR model.



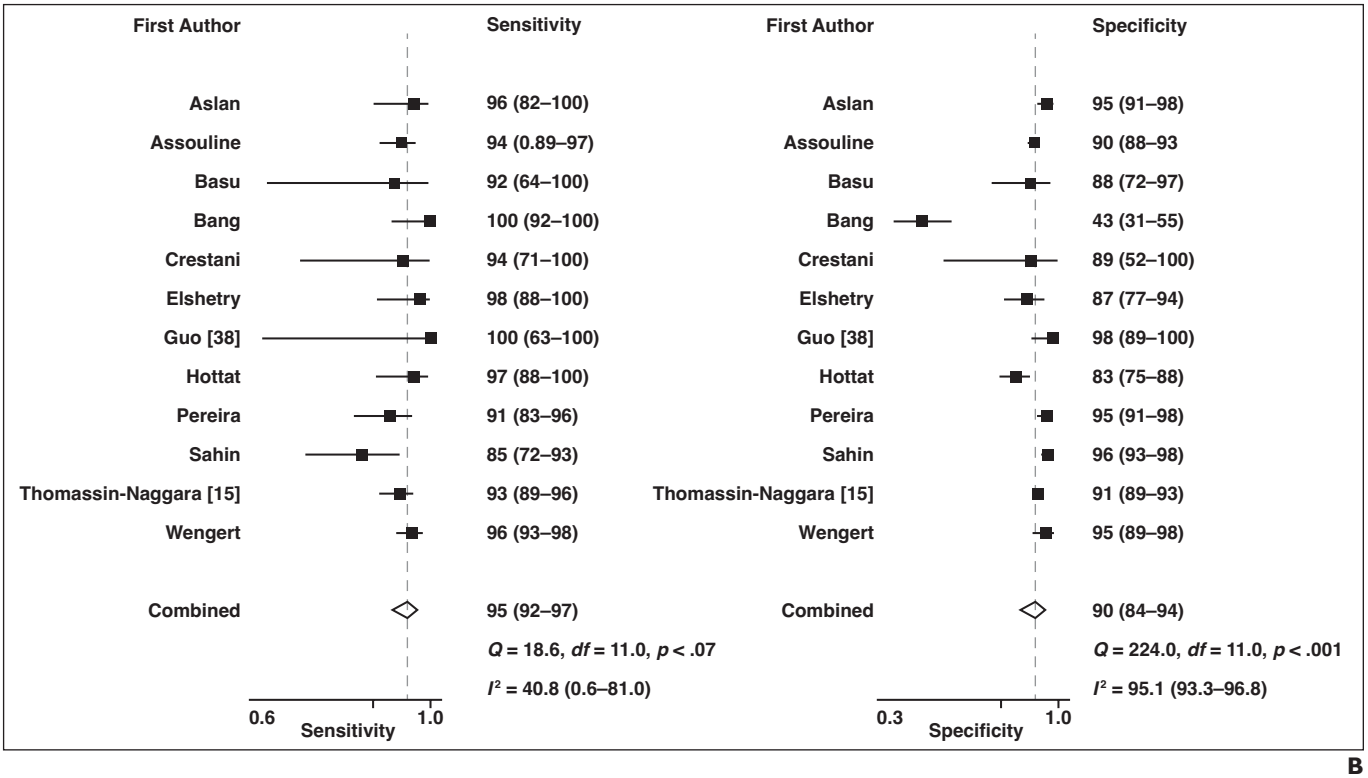




**Fig. 2**—Coupled forest plots of pooled sensitivity and specificity. Values are pooled estimates with 95% CI in parentheses. Boxes indicate point estimates; horizontal bars, 95% CIs of individual studies; diamonds, overall effect estimates (top and bottom peaks, overall point estimate; left and right ends, 95% CI); dashed line, line of no effect.

**A,** Plot shows results for O-RADS ultrasound.

**B,** Plot shows results for O-RADS MRI.

**TABLE 2: Study Characteristics**

| First Author | O-RADS Arm | Publication Year | Study Design | Level of Analysis | Study Period | No. of Readers | Reader Experience (y)[a] | Reader Blinding | Technical Parameters[b] | κ | Cutoff O-RADS Category for Malignancy |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Ahmed [40] | US | 2021 | Prospective | Per person | NA | NA | NA | NA | Color Doppler | NA | ≥ 4 |
| Basha [26] | US | 2021 | Retrospective | Per lesion | 5/2016–12/2019 | 5 | 15 | Yes | Color Doppler | 0.77 | ≥ 4 |
| Cao [27] | US | 2021 | Retrospective | Per person | 1/2016–12/2018 | 2 | 5 | Yes | Color Doppler | 0.77 | ≥ 4 |
| Chen [29] | US | 2022 | Retrospective | Per person | 1/2019–12/2019 | 2 | 14 | Yes | Color Doppler | NA | ≥ 5 |
| Chen [28] | US | 2022 | Retrospective | Per person | 1/2020–10/2020 | 5 | ≥ 5 | Yes | Color Doppler | 0.76 | ≥ 4 |
| Guo [30] | US | 2022 | Retrospective | Per lesion | 2017–2020 | 2 | ≥ 10 | Yes[c] | Color Doppler | 0.71 | ≥ 4 |
| Guo [38][d] | US<br>MRI | 2022 | Retrospective | Per lesion | 7/2017–6/2020 | 2<br>2 | 3, 36<br>2, 2 | Yes | Color Doppler<br>1.5 or 3 T, T1/T2/DCE/DWI | 0.43<br>0.58 | ≥ 4<br>≥ 4 |
| Hack [31] | US | 2022 | Retrospective | Per lesion | 8/2015–4/2017 | 2 | 9, 30 | Yes | Color Doppler | NA | ≥ 4 |
| Hiett [32] | US | 2021 | Retrospective | Per person | 3/2018–2/2021 | 2 | 25 | Yes[c] | Color Doppler | NA | ≥ 4 |
| Jha [36] | US | 2022 | Retrospective | Per lesion | 1/2011.1–12/2014 | 8 | 1–20 | Yes | Color Doppler | 0.56 | ≥ 4 |
| Lai [35] | US | 2021 | Retrospective | Per person | 1/2017–11/2020 | 2 | 5 | Yes | Color Doppler | 0.83 | ≥ 4 |
| Solis Cano [39] | US | 2021 | Retrospective | Per person | 2017–2021 | 1 | NA | Yes | Color Doppler | NA | ≥ 3 |
| Wang [34] | US | 2022 | Retrospective | Per person | 2/2020–10/2021 | ≥ 2 | NA | Yes | Color Doppler | NA | ≥ 4 |
| Wu [37] | US | 2022 | Retrospective | Per person | 1/2017–9/2020 | 2 | 5 | Yes | Color Doppler | 0.62 | ≥ 4 |
| Xie [33] | US | 2022 | Retrospective | Per lesion | 1/2018–1/2020 | 2 | ≥ 10 | Yes | Color Doppler | 0.83 | ≥ 4 |
| Aslan [41] | MRI | 2021 | Retrospective | Per lesion | 1/2018–6/2020 | 1 | 10 | Yes | 1.5 T, T1/T2/DCE/DWI | NA | ≥ 4 |
| Assouline [47] | MRI | 2022 | Retrospective | Per lesion | 3/2013–3/2018 | 13 | 6 m–10 y | Yes | 1.5 or 3 T, T1/T2/DCE/DWI | NA | ≥ 4 |
| Bang [48] | MRI | 2022 | Retrospective | Per lesion | 1/2014–7/2020 | NA | NA | Yes | NA for MRI field, T1/T2/DCE/DWI | NA | ≥ 4 |
| Basu [45] | MRI | 2022 | Prospective | Per lesion | 4/2020–6/2021 | 1 | 10 | Yes | 1.5 T, T1/T2/DCE/DWI | NA | ≥ 4 |
| Crestani [46] | MRI | 2020 | Retrospective | Per person | 2014–2018 | 1 | 15 | Yes | NA for MRI field, T1/T2/DCE/DWI | NA | ≥ 4 |
| Elshetry [44] | MRI | 2022 | Prospective | Per person | 4/2020–9/2021 | 2 | 9, 12 | Yes | 1.5 T, T1/T2/DCE/DWI | NA | ≥ 4 |
| Hottat [42] | MRI | 2021 | Prospective | Per lesion | 1/2015–4/2020 | 2 | 10, 13 | Yes | 1.5 or 3 T, T1/T2/DCE/DWI | 0.97 | ≥ 4 |

**(Table 2 continues on next page)**

**TABLE 2: Study Characteristics (continued)**

| First Author | O-RADS Arm | Publication Year | Study Design | Level of Analysis | Study Period | No. of Readers | Reader Experience (y)[a] | Reader Blinding | Technical Parameters[b] | κ | Cutoff O-RADS Category for Malignancy |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Pereira [49] | MRI | 2022 | Prospective | Per lesion | 2/2014–2/2020 | 2 | 9, 10 | Yes | 1.5 T, T1/T2/DCE/DWI | NA | ≥ 4 |
| Sahin [43] | MRI | 2021 | Retrospective | Per lesion | 1/2008–12/2018 | 2 | 6, 8 | Yes[c] | 1.5 T, T1/T2/DWI | 0.73 | ≥ 4 |
| Thomassin-Naggara [15] | MRI | 2020 | Prospective | Per person | 3/2013–3/2016 | 2 | 6–12 m, ≥ 10 y | Yes | 1.5 or 3 T, T1/T2/DCE/DWI | 0.78 | ≥ 4 |
| Wengert [16] | MRI | 2022 | Prospective | Per lesion | 3/2013–3/2018 | ≥ 2 | ≥ 10 | Yes | 1.5 or 3 T, T1/T2/DCE/DWI | NA | ≥ 4 |

Note—Numbers in brackets denote references. US = ultrasound, NA = not available, T1 = T1-weighted imaging, T2 = T2-weighted imaging, DCE = dynamic contrast-enhanced imaging.

[a]Years unless otherwise indicated.

[b]For MRI, examinations included all listed sequences.

[c]Readers were aware of partial clinical information.

[d]Cells listing two values report information for US arm followed by information for MRI arm.

Heterogeneity of included studies was assessed with the *Q* statistic and $I^2$. Heterogeneity based on the *Q* statistic was assessed by use of the associated *p* value. Heterogeneity based on $I^2$ was classified as follows on the basis of overlapping intervals found in the *Cochrane Handbook for Systematic Reviews of Interventions* [25]: 0–40%, unimportant; 30–60%, moderate; 50–90%, substantial; 75–100%, considerable. Heterogeneity was also assessed visually on the basis of a comparison of the 95% confidence region and the 95% prediction region of the HSROC curve.

Metaregression was performed to explore sources of heterogeneity. The following study characteristics were used as covariates for O-RADS US: number of masses (≤ 400 vs > 400), publication year (2021 vs 2022), level of analysis (per lesion vs per patient), reference standard (pathology vs pathology and follow-up imaging), and malignancy rate (< 28% vs ≥ 28%). The following study characteristics were used as covariates for O-RADS MRI: number of masses (≤ 200 vs > 200), publication year (before 2020 vs 2020 or later), study design (prospective vs retrospective), reference standard (pathology vs pathology and follow-up imaging), level of analysis (per lesion vs per patient), and malignancy rate (< 25% vs ≥ 25%). For ordinal or continuous covariates, the thresholds were selected empirically to obtain two groups with similar sizes.

Diagnostic performance measures of at least 90% were considered to indicate high performance. Values of $p < .05$ were considered statistically significant. Lack of overlap of 95% CIs also was considered evidence of a significant difference. All analyses were performed with Stata Statistical Software (release 16.0, StataCorp).

## Evidence Synthesis

### *Literature Search and Study Selection*

Figure 1 shows the flow of study selection. The initial literature search identified 821 articles, from which 319 duplicate records were excluded. The title and abstract were reviewed for the remaining 502 unique articles. Of these, 349 were excluded because they did not report on original research ($n = 123$) or were not relevant to the study question ($n = 226$). The full text of the remaining 153 potentially eligible articles was reviewed. Of these, 127 were excluded because they were not relevant to study question ($n = 104$) or contained insufficient data to construct 2 × 2 tables to determine diagnostic performance ($n = 23$). These exclusions resulted in a final sample of 26 studies comprising 9520 participants with a total of 10,190 masses. Of these 26 studies, 15 were evaluations of O-RADS US [26–40], and 12 were evaluations of O-RADS MRI [15, 16, 38, 41–49]; both O-RADS US and O-RADS MRI were evaluated in 1 of the 26 studies [38].

### *Characteristics of Included Studies*

The clinical characteristics of the included studies are summarized in Table 1. The study characteristics are summarized in Table 2. For O-RADS US, 14 studies were retrospective [26–39], and one study [40] was prospective. For O-RADS MRI, six studies were prospective [15, 16, 42, 44, 46, 49], and six studies [38, 41, 43, 46–48] were retrospective. Study sample size ranged from 50 to 1054 patients in studies of O-RADS US and from 26 to 1130 patients in studies of O-RADS MRI. Mean or median patient age ranged from 35.0 to 53.4 years in studies of O-RADS US and from 35.9 to 57.8 years in studies of O-RADS MRI. The malignancy rate ranged from 11.9% to 65.4% in studies of O-RADS US, and from 8.0% to 64.7% in studies of O-RADS MRI. Examinations were interpreted by a single reader in four studies, by two readers in 17 studies, and by at least three readers in three studies; in the other two studies, the number of readers was not reported. Kappa values ranged from 0.43 to 0.83 for O-RADS US risk categories and from 0.58 to 0.78 for O-RADS MRI risk categories.

For studies of O-RADS US, the reference standard was pathology in 10 studies and pathology or imaging follow-up (6 months–2 years) in five studies. For studies of O-RADS MRI, the reference standard was pathology in five studies, and pathology or imaging follow-up (4 months–2 years) in seven studies. For diagnosis of malignancy, one study had a threshold O-RADS category 3 or higher [39], one study had a threshold O-RADS category of 5 or higher [29], and the other 24 studies had

**TABLE 3: Distribution of O-RADS Categories and Number of Malignancies for Surgically Resected Lesions**

| First Author | O-RADS Arm | No. of Masses | O-RADS Category | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 1 | | 2 | | 3 | | 4 | | 5 | |
| | | | Assigned | Malignant | Assigned | Malignant | Assigned | Malignant | Assigned | Malignant | Assigned | Malignant |
| Ahmed [40] | US | 50 | 0 | 0 | 0 | 0 | 13 | 2 | 18 | 15 | 21 | 19 |
| Basha [26] | US | 647 | 0 | 0 | 262 | 1 | 179 | 5 | 37 | 11 | 169 | 161 |
| Cao [27] | US | 1054 | 0 | 0 | 446 | 2 | 182 | 3 | 148 | 51 | 278 | 249 |
| Chen [29] | US | 85 | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| Chen [28] | US | 322 | 0 | 0 | 149 | 0 | 68 | 2 | 95 | 26 | 68 | 30 |
| Guo [30] | US | 592 | 1 | 0 | 299 | 6 | 79 | 7 | 153 | 80 | 60 | 52 |
| Guo [38] | US | 58 | 0 | 0 | 31 | 0 | 11 | 1 | 11 | 3 | 5 | 4 |
| Hack [31] | US | 262 | 0 | 0 | 100 | 0 | 32 | 1 | 63 | 22 | 67 | 52 |
| Hiett [32] | US | 150 | 0 | 0 | 17 | 0 | 34 | 0 | 66 | 14 | 33 | 26 |
| Jha [36] | US | 1014 | 0 | 0 | 657 | 3 | 112 | 5 | 155 | 18 | 90 | 59 |
| Lai [35] | US | 734 | 0 | 0 | 369 | 5 | 71 | 65 | 94 | 90 | 73 | 69 |
| Solis Cano [39] | US | 73 | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| Wang [34] | US | 443 | 0 | 0 | 179 | 0 | 72 | 1 | 121 | 65 | 71 | 65 |
| Wu [37] | US | 431 | 0 | 0 | 79 | 1 | 168 | 8 | 86 | 70 | 98 | 93 |
| Xie [33] | US | 453 | 0 | 0 | 78 | 4 | 59 | 11 | 134 | 82 | 182 | 172 |
| Aslan [41] | MRI | 237 | 12 | 0 | 111 | 0 | 77 | 1 | 20 | 10 | 17 | 17 |
| Assouline [47] | MRI | 585 | 25 | 0 | 391 | 2 | 142 | 9 | 99 | 56 | 122 | 106 |
| Bang [48] | MRI | 307 | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| Basu [45] | MRI | 42 | 1 | 0 | 19 | 0 | 9 | 1 | 4 | 1 | 12 | 11 |
| Crestani [46] | MRI | 26 | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| Elshetry [44] | MRI | 90 | 0 | 0 | 39 | 0 | 24 | 1 | 26 | 18 | 27 | 26 |
| Guo [38] | MRI | 58 | | | 36 | 0 | 13 | 0 | 5 | 4 | 4 | 4 |
| Hottat [42] | MRI | 201 | 6 | 0 | 32 | 0 | 88 | 2 | 42 | 21 | 39 | 35 |
| Pereira [49] | MRI | 243 | 4 | 0 | 123 | 0 | 31 | 4 | 25 | 21 | 54 | 54 |
| Sahin [43] | MRI | 201 | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |
| Thomassin-Naggara [15] | MRI | 1130 | 91 | 10 | 571 | 2 | 213 | 12 | 122 | 60 | 133 | 119 |
| Wengert [16] | MRI | 320 | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |

Note—Numbers in brackets denote references. US = ultrasound, NA = not applicable.

a threshold O-RADS category of 4 or higher. Four studies compared the diagnostic performance of O-RADS US with IOTA simple rules [26, 30, 32, 33], and three studies compared the diagnostic performance of O-RADS US with the ADNEX model [28, 32, 35].

### Quality Assessment

Tables S1 and S2 (available in the online supplement) summarize the results of the quality assessment by means of QUADAS-2 and QUADAS-C. In the patient selection domain, eight studies were assessed as having high risk of bias [27, 30, 33, 34, 38, 41, 46, 48], most commonly because the study excluded patients without a CA 125 measurement, and eight studies were assessed as unclear concern of applicability, in all cases because of a high malignancy rate. In the index test domain, one study was assessed as having unclear risk of bias because the study did not include information regarding reader blinding [40]. Four studies were assessed as having unclear concern about applicability because the readers knew partial clinical information about study patients [30, 32, 41, 43]. In the reference standard domain, one study was assessed as having high risk of bias because the reference standard for some patients was only 4-month follow-up imaging [45]. No study was assessed as having high or unclear concern about applicability. In the flow and timing domain, one study was assessed as having high risk of bias because results were not reported for O-RADS category 5 lesions [48].

### Diagnostic Performance of O-RADS Ultrasound and O-RADS MRI

For O-RADS US, the sensitivity of individual studies ranged from 52% to 100% with a pooled summary estimate of 95% (95% CI, 91–97%), and the specificity of individual studies ranged from 46% to 94% with a pooled summary estimate of 82% (95% CI, 76–87%). The coupled forest plots for O-RADS US are shown in Figure 2A. Table 3 shows the distribution of O-RADS US categories among lesions that were surgically resected, along with the number of malignancies in each category. The *Q* test revealed significant heterogeneity among studies in terms of sensitivity and specificity (both $p < .001$). The $I^2$ statistic indicated considerable heterogeneity in terms of sensitivity ($I^2 = 87.9\%$) and specificity ($I^2 = 94.6\%$). On the HSROC curve (Fig. 3A), a large visual difference between the 95% confidence region and prediction region suggested substantial heterogeneity among studies. Figure 4A shows the Deeks funnel plot; $p = .09$ for the slope coefficient in the Deeks test indicates a low likelihood of publication bias.

For O-RADS MRI, the sensitivity of individual studies ranged from 85% to 100% with a pooled summary estimate of 95% (95% CI, 92–97%, and the specificity of individual studies ranged from 43% to 88% with a pooled summary estimate of 90% (95% CI, 84–94%). The coupled forest plots for O-RADS MRI are shown in Figure 2B. The *Q* test revealed significant heterogeneity among studies in terms of specificity ($p < .001$) but not of sensitivity ($p = .07$). The $I^2$ statistic indicated moderate heterogeneity in terms of sensitivity ($I^2 = 40.8\%$) and considerable heterogeneity in terms of specificity ($I^2 = 95.1\%$). On the HSROC curve (Fig. 3B), the difference between the 95% confidence region and the 95% prediction region was visually smaller than for O-RADS US, suggesting lesser heterogeneity. Figure 4B shows the Deeks funnel plot; $p = .32$ for the slope coefficient in the Deeks test indicates low likelihood of publication bias.

### Metaregression

The results of metaregression analyses performed to investigate the sources of heterogeneity among studies are summarized in Table 4. For O-RADS US, none of the assessed factors (number of masses, publication year, level of analysis, reference standard, malignancy rate) were found to be associated with het-

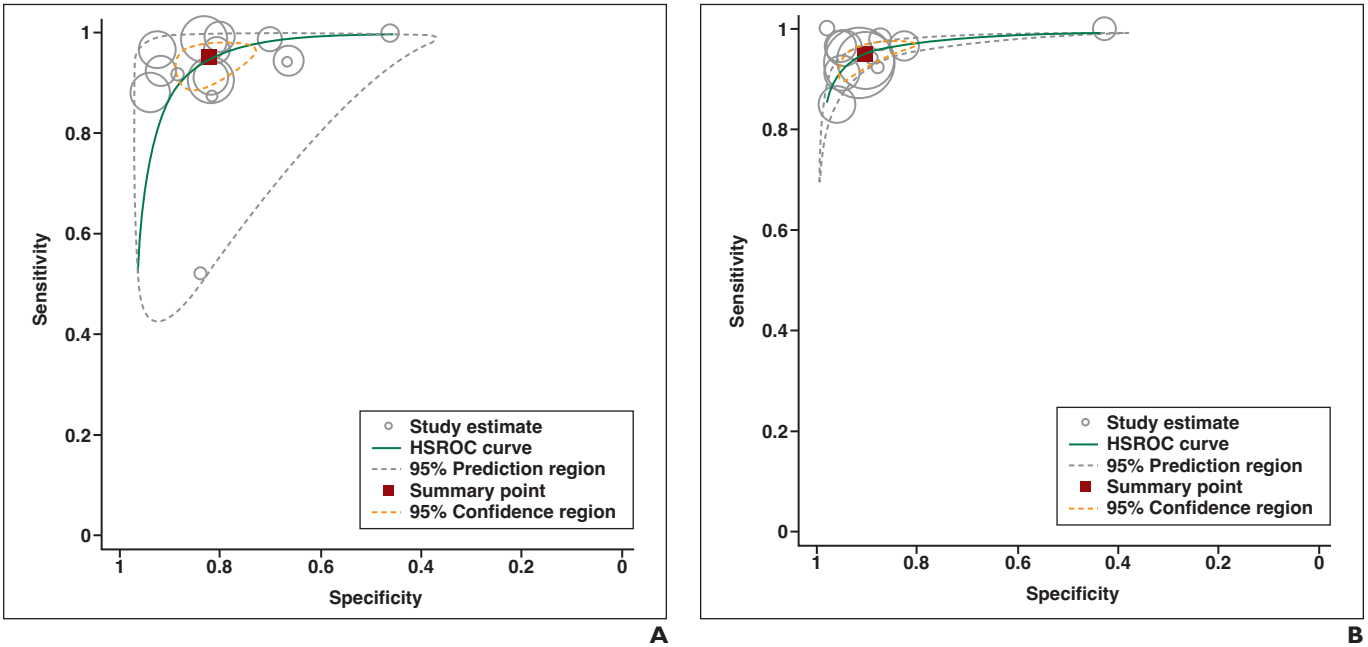



**Fig. 3**—Hierarchic summary ROC (HSROC) plots with summary point and 95% confidence regions.
**A,** Plot shows results for O-RADS ultrasound.
**B,** Plot shows results for O-RADS MRI.

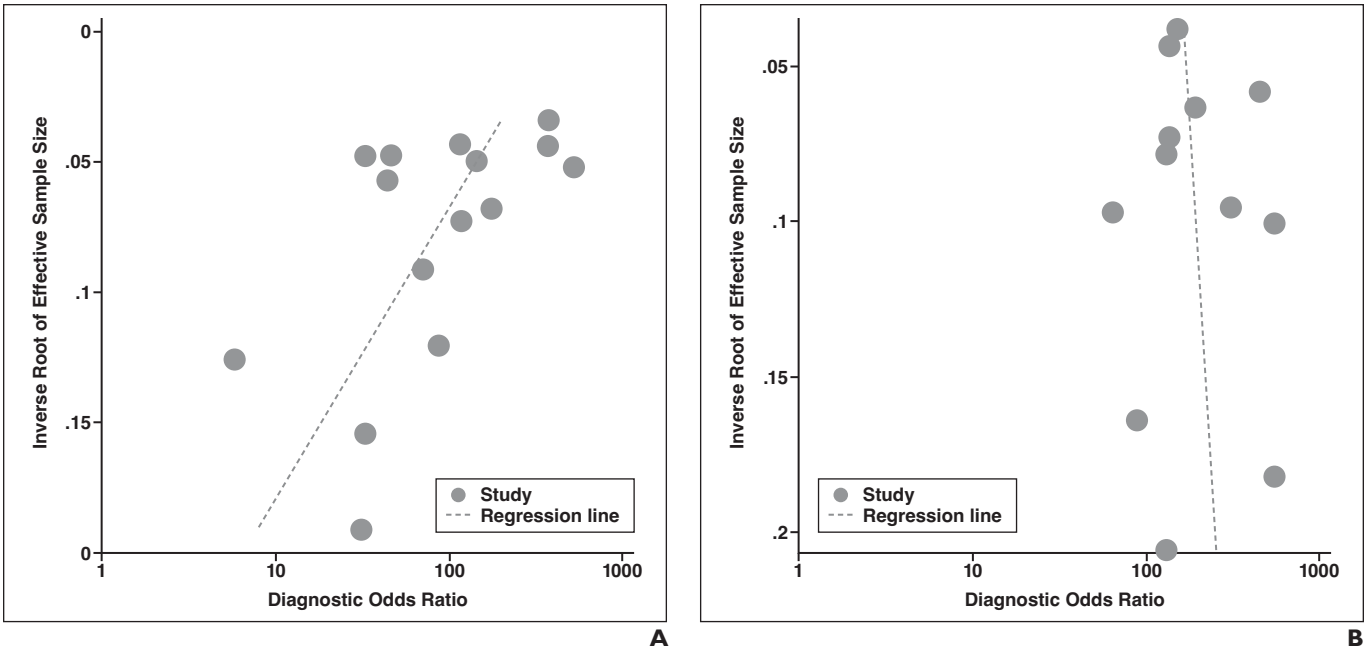



**Fig. 4**—Deeks funnel plot asymmetry tests.
**A,** Plot shows results for O-RADS ultrasound ($p = .09$).
**B,** Plot shows results for O-RADS MRI ($p = .63$).

erogeneity (all $p > .05$). For O-RADS MRI, studies with a reference standard of pathology combined with imaging follow-up, compared with studies that had a pathology-alone reference standard, had lower sensitivity (92% and 97%) and higher specificity (94% and 83%), although these differences were not statistically significant (both $p > .05$). In addition, studies with a malignancy rate of 25% or higher, compared with those with a malignancy rate less than 25%, had higher sensitivity (97% and 91%) but lower specificity (86% and 94%), although these differences were not statistically significant (both $p > .05$). The other covariates (study design, number of masses, publication year, and level of analysis) did not have a significant association with heterogeneity for O-RADS MRI (all $p > .05$).

### *Comparison of O-RADS Ultrasound With Other Systems*

Four studies included a head-to-head comparison of O-RADS US and the IOTA simple rules [26, 30, 32, 33]. Among these studies, the pooled sensitivity was not significantly different between O-RADS US (96% [95% CI, 93–98%]) and the IOTA simple rules (93% [95% CI, 83–97%]), nor was pooled specificity significantly different between O-RADS US (76% [95% CI, 54–89%]) and the IOTA simple rules (82% [95% CI 62–93%]). Three studies included a head-to-head comparison of O-RADS US and the ADNEX model [28, 32, 35]. The pooled sensitivity was not significantly different between O-RADS US (96% [95% CI, 79–99%]) and the ADNEX model (96% [95% CI, 93–99%]), nor was pooled specificity significantly different between O-RADS (79% [95% CI, 50–92%]) and the ADNEX model (78% [95% CI, 66–91%]).

## Discussion

In this systematic review and meta-analysis, we assessed the diagnostic performance of O-RADS US (15 studies) and O-RADS MRI (12 studies) for risk stratification of ovarian and adnexal lesions. Pooled estimates indicated that both systems had high sensitivity (95%) for detection of malignancy. O-RADS MRI had high specificity (90%), but O-RADS US did not (82%). In head-to-head comparisons, O-RADS US had no significant difference in sensitivity or specificity from earlier US risk stratification systems (IOTA simple rules and ADNEX model).

In a prior meta-analysis of the diagnostic performance of O-RADS US, 11 studies were evaluated; pooled sensitivity was 97%, and pooled specificity was 77% [17]. Our meta-analysis included five studies not included in the earlier meta-analysis and showed lower sensitivity but higher specificity than that study did. Unlike the earlier investigators, we compared O-RADS US with other stratification systems in head-to-head comparisons. In a prior meta-analysis of the diagnostic performance of O-RADS MRI, 12 studies were evaluated; pooled sensitivity was 92%, and pooled specificity was 91% [18]. Our meta-analysis included three studies not included in the earlier meta-analysis and showed slightly higher sensitivity than and similar specificity to those in the prior study. In contrast to our meta-analysis, the earlier meta-analysis included studies in which ADNEX MR (a predecessor to O-RADS MRI) was used to derive pooled estimates.

Significant heterogeneity was observed across studies for both O-RADS US and O-RADS MRI in terms of specificity and for O-RADS US in terms of sensitivity. In metaregression, for O-RADS MRI, nonsignificant associations were observed between the pooled estimates of sensitivity and specificity and the reference standards and malignancy rates in the studies. For O-RADS US, no covariate was associated with heterogeneity. The range of evaluated factors was impacted by the number of available studies and the reported features in the studies. Additional factors, beyond those explored, may have contributed to the observed heterogeneity.

O-RADS represents an effort to improve the evaluation, reporting, and management of ovarian and adnexal lesions. Compared with other scoring systems, such as IOTA simple rules and GI-RADS, O-RADS includes a comprehensive set of descriptors and interpretation algorithms for determining which adnexal lesions require no follow-up, imaging follow-up, or surgical resection. Moreover, O-RADS US classifies ovarian lesions on the basis of US features only, facilitating implementation in practice. In comparison, the other systems combine US findings with clinical variables, such as patient age and CA 125 level, to classify lesions. For example, the ADNEX model derives a risk score based on the combination of three clinical variables and six US variables. For O-RADS US, the lower specificity compared with sensitivity may reflect an intent for the system to maximize sensitivity to avoid missed cancers. As with other imaging-based scoring systems, the current limitations of O-RADS may be addressed through ongoing system updates guided by results of further studies.

**TABLE 4: Results of Metaregression**

| Covariate | Sensitivity (%) | *p* | Specificity (%) | *p* |
|---|---|---|---|---|
| Ultrasound | | | | |
| Publication year | | .96 | | .71 |
| 2021 | 95 (81–98) | | 84 (69–90) | |
| 2022 | 95 (93–98) | | 81 (76–88) | |
| No. of masses | | .78 | | .10 |
| ≤ 400 | 94 (75–98) | | 76 (65–85) | |
| > 400 | 95 (92–98) | | 86 (79–91) | |
| Reference standard | | .31 | | .84 |
| Pathology | 94 (86–97) | | 83 (73–89) | |
| Pathology or imaging follow-up | 97 (93–99) | | 82 (73–89) | |
| Level of analysis | | .58 | | .74 |
| Per person | 96 (93–98) | | 81 (71–88) | |
| Per lesion | 95 (83–98) | | 83 (77–90) | |
| Malignancy rate | | .97 | | .75 |
| < 28% | 95 (91–98) | | 83 (72–90) | |
| ≥ 28% | 95 (88–98) | | 81 (74–88) | |
| MRI | | | | |
| Publication year | | .22 | | .51 |
| < 2021 | 93 (88–96) | | 92 (87–97) | |
| ≥ 2021 | 96 (92–98) | | 89 (77–94) | |
| No. of masses | | .58 | | .70 |
| ≤ 200 | 96 (92–99) | | 92 (86–96) | |
| > 200 | 95 (87–97) | | 90 (78–95) | |
| Reference standard | | .03 | | .04 |
| Pathology | 97 (93–98) | | 83 (65–91) | |
| Pathology or imaging follow-up | 92 (88–95) | | 94 (91–97) | |
| Study design | | .70 | | .96 |
| Prospective | 95 (92–97) | | 91 (87–95) | |
| Retrospective | 94 (86–97) | | 91 (77–97) | |
| Level of analysis | | .89 | | .86 |
| Per person | 95 (91–98) | | 90 (87–97) | |
| Per lesion | 95 (91–97) | | 91 (83–95) | |
| Malignancy rate | | .02 | | .09 |
| < 25% | 91 (87–95) | | 94 (92–97) | |
| ≥ 25% | 97 (93–98) | | 86 (74–93) | |

Note—Data are percentages with 95% CIs in parentheses.

There were limitations to this study. First, MRI is used to evaluate ovarian and adnexal masses that are indeterminate on US, but we did not assess whether articles explicitly reported that the included adnexal lesions met this criterion. Second, O-RADS US and O-RADS MRI were evaluated in distinct patient samples rather than in head-to-head comparison studies. Third, the studies commonly had risks of bias or concerns of applicability related to patient selection, potentially limiting the generalizability of results. Fourth, diagnostic performance was not evaluated in terms of readers' specific experience in use of O-RADS US or O-RADS MRI, both of which were in existence at the time of this study. Finally, diagnostic performance in each study was determined in a single cutoff risk category; summary results were not obtained for individual O-RADS categories.

## Conclusion

O-RADS US and O-RADS MRI both had high sensitivity for malignancy of ovarian and adnexal lesions. O-RADS MRI, but not O-RADS US, also had high specificity. Significant heterogeneity in diagnostic performance among studies was observed for O-RADS US in terms of sensitivity and specificity and for O-RADS MRI in terms of specificity but not sensitivity. In head-to-head comparisons, O-RADS US exhibited no significant difference in diagnostic performance compared with the IOTA simple rules and the ADNEX model. Awareness of these diagnostic performance results for O-RADS US and O-RADS MRI will be helpful as these systems are increasingly implemented into clinical practice.

**Provenance and review:** Not solicited; externally peer reviewed.

**Peer reviewers:** All peer reviewers chose not to disclose their identities.

## References

1. American Cancer Society website. Key statistics for ovarian cancer. www.cancer.org/cancer/ovarian-cancer/about/key-statistics.html. Accessed Jan 10, 2023
2. Lheureux S, Braunstein M, Oza AM. Epithelial ovarian cancer: evolution of management in the era of precision medicine. *CA Cancer J Clin* 2019; 69:280–304
3. Narod S. Can advanced-stage ovarian cancer be cured? *Nat Rev Clin Oncol* 2016; 13:255–261
4. Glanc P, Benacerraf B, Bourne T, et al. First international consensus report on adnexal masses: management recommendations. *J Ultrasound Med* 2017; 36:849–863
5. Timmerman D, Valentin L, Bourne TH, Collins WP, Verrelst H, Vergote I; International Ovarian Tumor Analysis (IOTA) Group. Terms, definitions and measurements to describe the sonographic features of adnexal tumors: a consensus opinion from the International Ovarian Tumor Analysis (IOTA) Group. *Ultrasound Obstet Gynecol* 2000; 16:500–505
6. Van Calster B, Van Hoorde K, Valentin L, et al.; International Ovarian Tumour Analysis Group. 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* 2014; 349:g5920
7. Amor F, Alcázar JL, Vaccaro H, León M, Iturra A. GI-RADS reporting system for ultrasound evaluation of adnexal masses in clinical practice: a prospective multicenter study. *Ultrasound Obstet Gynecol* 2011; 38:450–455
8. Andreotti RF, Timmerman D, Strachowski LM, 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
9. Meys EMJ, Kaijser J, Kruitwagen RFPM, et al. Subjective assessment versus ultrasound models to diagnose ovarian cancer: a systematic review and meta-analysis. *Eur J Cancer* 2016; 58:17–29
10. Froyman W, Landolfo C, De Cock B, et al. Risk of complications in patients with conservatively managed ovarian tumours (IOTA5): a 2-year interim analysis of a multicentre, prospective, cohort study. *Lancet Oncol* 2019; 20:448–458
11. Sadowski EA, Paroder V, Patel-Lippmann K, et al. Indeterminate adnexal cysts at US: prevalence and characteristics of ovarian cancer. *Radiology* 2018; 287:1041–1049
12. Hricak H, Chen M, Coakley FV, et al. Complex adnexal masses: detection and characterization with MR imaging—multivariate analysis. *Radiology* 2000; 214:39–46
13. Sohaib SAA, Sahdev A, Van Trappen P, Jacobs IJ, Reznek RH. Characterization of adnexal mass lesions on MR imaging. *AJR* 2003; 180:1297–1304
14. Thomassin-Naggara I, Aubert E, Rockall A, et al. Adnexal masses: development and preliminary validation of an MR imaging scoring system. *Radiology* 2013; 267:432–443
15. Thomassin-Naggara I, Poncelet E, Jalaguier-Coudray A, et al. Ovarian-Adnexal Reporting Data System magnetic resonance imaging (O-RADS MRI) score for risk stratification of sonographically indeterminate adnexal masses. *JAMA Netw Open* 2020; 3:e1919896
16. Wengert GJ, Dabi Y, Kermarrec E, et al.; EURAD Study Group. O-RADS MRI classification of indeterminate adnexal lesions: time-intensity curve analysis is better than visual assessment. *Radiology* 2022; 303:566–575
17. Vara J, Manzour N, Chacón E, et al. Ovarian Adnexal Reporting Data System (O-RADS) for classifying adnexal masses: a systematic review and meta-analysis. *Cancers (Basel)* 2022; 14:3151
18. Rizzo S, Cozzi A, Dolciami M, et al. O-RADS MRI: a systematic review and meta-analysis of diagnostic performance and category-wise malignancy rates. *Radiology* 2022 Nov 22 [published online]
19. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009; 339:b2700
20. Whiting PF, Rutjes AW, Westwood ME, et al.; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155:529–536
21. Yang B, Mallett S, Takwoingi Y, et al.; QUADAS-C Group. QUADAS-C: a tool for assessing risk of bias in comparative diagnostic accuracy studies. *Ann Intern Med* 2021; 174:1592–1599
22. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; 58:982–990
23. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001; 20:2865–2884
24. Deeks JJ. Systematic reviews in health care: systematic reviews of evaluations of diagnostic and screening tests. *BMJ* 2001; 323:157–162
25. Higgins JPT, Altman DG, Gøtzsche PC, et al.; Cochrane Bias Methods Group; Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011; 343:d5928
26. Basha MAA, Metwally MI, Gamil SA, 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
27. Cao L, Wei M, Liu Y, 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
28. Chen GY, Hsu TF, Chan IS, et al. Comparison of the O-RADS and ADNEX models regarding malignancy rate and validity in evaluating adnexal lesions. *Eur Radiol* 2022; 32:7854–7864

29. Chen H, Yang BW, Qian L, et al. Deep learning prediction of ovarian malignancy at US compared with O-RADS and expert assessment. *Radiology* 2022; 304:106–113
30. Guo Y, Zhao B, Zhou S, 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; 41:511–518
31. Hack K, Gandhi N, Bouchard-Fortier G, et al. External validation of O-RADS US risk stratification and management system. *Radiology* 2022; 304:114–120
32. Hiett AK, Sonek JD, Guy M, Reid TJ. Performance of IOTA simple rules, simple rules risk assessment, ADNEX model and O-RADS in differentiating between benign and malignant adnexal lesions in North American women. *Ultrasound Obstet Gynecol* 2022; 58:668–676
33. Xie WT, Wang YQ, Xiang ZS, et al. Efficacy of IOTA simple rules, O-RADS, and CA125 to distinguish benign and malignant adnexal masses. *J Ovarian Res* 2022; 15:15
34. Wang R, Li X, Li S, et al. Clinical value of O-RADS combined with serum CA125 and HE4 for the diagnosis of ovarian tumours. *Acta Radiol* 2022 Mar 15 [published online]
35. Lai HW, Lyu GR, Kang Z, et al. 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
36. Jha P, Gupta A, Baran TM, et al. Diagnostic performance of the Ovarian-Adnexal Reporting and Data System (O-RADS) ultrasound risk score in women in the United States. *JAMA Netw Open* 2022; 5:e2216370
37. Wu M, Wang Q, Zhang M, et al. Does combing O-RADS US and CA-125 improve diagnostic accuracy in assessing adnexal malignancy risk in women with different menopausal status? *J Ultrasound Med* 2022 Jul 26 [published online]
38. Guo Y, Phillips CH, Suarez-Weiss K, et al. Interreader agreement and intermodality concordance of O-RADS US and MRI for assessing large, complex ovarian-adnexal cysts. *Radiol Imaging Cancer* 2022; 4:e220064
39. Solis Cano DG, Cervantes Flores HA, De Los Santos Farrera O, Guzman Martinez NB, Soria Céspedes D. Sensitivity and specificity of ultrasonography using Ovarian-Adnexal Reporting and Data System classification versus pathology findings for ovarian cancer. *Cureus* 2021; 13:e17646
40. Ahmed HEK. The usefulness of the ultrasound diagnosis of suspicious ovarian masses based on the O-RADS classification system. *Al-Azhar Int Med J* 2021; 2:1–6
41. Aslan S, Tosun SA. Diagnostic accuracy and validity of the O-RADS MRI score based on a simplified MRI protocol: a single tertiary center retrospective study. *Acta Radiol* 2023; 64:377–386
42. Hottat NA, Badr DA, Van Pachterbeke C, et al Added value of quantitative analysis of diffusion-weighted imaging in Ovarian-Adnexal Reporting and Data System magnetic resonance imaging. *J Magn Reson Imaging* 2022; 56:158–170
43. Sahin H, Panico C, Ursprung S, et al. Non-contrast MRI can accurately characterize adnexal masses: a retrospective study. *Eur Radiol* 2021; 31:6962–6973
44. Elshetry ASF, Hamed EM, Frere RAF, Zaid NA. Impact of adding mean apparent diffusion coefficient (ADCmean) measurements to O-RADS MRI scoring for adnexal lesions characterization: a combined O-RADS MRI/ADCmean approach. *Acad Radiol* 2023; 30:300–311
45. Basu A, Pame M, Bhuyan RK, et al. Diagnostic performance of O-RADS MRI scoring system for the assessment of adnexal masses in routine clinical radiology practice: a single tertiary centre prospective cohort study. *J Clin Diagn Res* 2022; 16:11–16
46. Crestani A, Theodore C, Levaillant J-M, et al. Magnetic resonance and ultrasound fusion imaging to characterise ovarian masses: a feasibility study. *Anticancer Res* 2020; 40:4115–4121
47. Assouline V, Dabi Y, Jalaguier-Coudray A, et al. EURAD study group. How to improve O-RADS MRI score for rating adnexal masses with cystic component? *Eur Radiol* 2022; 32:5943–5953
48. Bang JI, Kim JY, Choi MC, Lee HY, Jang SJ. Application of multimodal imaging biomarker in the differential diagnosis of ovarian mass: integration of conventional and molecular imaging. *Clin Nucl Med* 2022; 47:117–122
49. Pereira PN, Yoshida A, Sarian LO, Barros RHO, Jales RM, Derchain S. Assessment of the performance of the O-RADS MRI score for the evaluation of adnexal masses, with technical notes. *Radiol Bras* 2022; 55:137–144