The diagnostic performance of the Gynecologic Imaging Reporting and Data System (GI-RADS) in adnexal masses

Wen Guo1,2#, Xiuhe Zou3#, Hanyue Xu1,2#, Tao Zhang1,2#, Yunuo Zhao1,2#, Lu Gao4,5#, Wenyue Duan4,5#, Xuelei Ma1#, Ling Zhang1#

1Department of Biotherapy, State Key Laboratory of Biotherapy, West China Hospital, Cancer Center, Sichuan University, Chengdu, China; 2West China Hospital, West China School of Medicine, Sichuan University, Chengdu, China; 3Department of Thyroid Surgery, West China Hospital of Sichuan University, Chengdu, China; 4Department of Oncology, The General Hospital of Western Theater Command, Chengdu, China; 5College of Medicine, Southwest Jiaotong University, Chengdu, China

Contributions: (I) Conception and design: L Zhang, X Ma, W Guo; (II) Administrative support: L Zhang, X Ma; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: W Guo, X Zou, H Xu, T Zhang; (V) Data analysis and interpretation: W Guo X Zou, Y Zhao; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

#These authors contributed equally to this work as co-first authors.

Background: Adnexal masses, mostly benign, are common in the female genital system. However, adnexal masses are the leading cause of death among women with gynecologic cancer. Ultrasound is a common imaging method for diagnosing adnexal masses. Gynecologic Imaging Reporting and Data System (GI-RADS) is a useful diagnostic tool based on objective ultrasound features to diagnose the malignancy of the female genital system. Therefore, we conducted a meta-analysis to evaluate the ability of GI-RADS to differentiate adnexal masses.

Methods: Published articles were searched in PubMed, Medline, and Embase from 1990 to February 2020. Pooled sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), diagnostic odds ratio, and area under the curve (AUC) were estimated via the extracted data from the selected studies.

Results: Ten studies and 2,474 patients were included in this meta-analysis. The pooled sensitivity of selected studies was 0.95 [95% confidence intervals (CI): 0.94–0.97], and the pooled specificity was 0.86 (95% CI: 0.84–0.88). The pooled NLR and PLR were 0.06 (95% CI: 0.04–0.10), and 8.30 (95% CI: 4.93–13.97), respectively. Moreover, the pooled diagnostic odds ratio for GI-RADS was 174.59 (95% CI: 76.70–397.42), and the AUC was 0.9806.

Conclusions: This research indicates that GI-RADS might be a valuable tool to distinguish malignancies from adnexal masses.

Keywords: Gynecologic Imaging Reporting and Data System (GI-RADS); diagnosis; adnexal mass; ovarian cancer

Submitted Jul 06, 2020. Accepted for publication Dec 01, 2020.
doi: 10.21037/atm-20-5170

View this article at: http://dx.doi.org/10.21037/atm-20-5170

^ ORCID: Wen Guo, 0000-0001-9717-8566; Xiuhe Zou, 0000-0001-8176-4579; Hanyue Xu, 0000-0002-9905-5874; Tao Zhang, 0000-0003-0562-6865; Yunuo Zhao, 0000-0002-5818-9278; Lu Gao, 0000-0003-1528-589X; Wenyue Duan, 0000-0003-2858-6261; Ling Zhang, 0000-0001-7297-8532; Xuelei Ma, 0000-0002-9148-5001.
Introduction

Adnexal masses are common in the female genital system, and most of them are benign. Nevertheless, ovarian cancer is the third most common gynecologic malignancy (1.6%) and the leading cause of death among women with gynecologic cancer (1,2). Therefore, it is significant to find early and accurate diagnostic methods to recognize the malignant adnexal masses. Now, as recommended by the guidelines of ovarian cancer, differential diagnostic methods include ultrasound, Computerized Tomography (CT), Magnetic Resonance Imaging (MRI), and serum cancer antigen tests (3-5).

Ultrasound examination is used in investigating pelvic tumors. However, because of the experiential discrepancy among inspectors, the results of evaluating adnexal masses are not objective enough. According to a multicentre clinical trial, the false positive rate of investigating ovarian cancer is 24% (6). Meanwhile, problems also exist in the interpretation of the information from ultrasound doctors to clinicians (6).

Similar to Breast Imaging Reporting and Data System (BI-RADS), designed by the American College of Radiology to reduce the influence of subjective judgment on ultrasonic examination results of breasts, Gynecologic Imaging Reporting and Data System (GI-RADS) was introduced by Amor. F to standardize the ultrasound report of the gynecologic system (7,8). This system included several ultrasonic morphological and clinical features and they were laterality, maximum diameter, echogenicity, wall thickness, cystic content, solid areas, septations, solid papillary projections, the presence of ascites, the doppler evaluation. The classification of GI-RADS includes GI-RADS 1 to GI-RADS 5, and a higher level indicates that the mass is more likely to be malignant. The kappa coefficients of GI-RADS categorization were 0.896 (95% CI: 0.775 to 1.000), according to a previous study (9). The diagnostic performance of GI-RADS in different studies is highly heterogeneous, and no clinical trial of large samples had investigated such discrepancy. Thus, this present study aimed to design a meta-analysis to evaluate and review the effect of the GI-RADS. Further, we also compared different ovarian cancer diagnostic methods with GI-RADS. We presented this article following the PRISMA reporting checklist (available at http://dx.doi.org/10.21037/atm-20-5170).

Methods

Search strategy

We searched the studies published from 1990 to February 2020 in PubMed, MEDLINE, and EMBASE, and the search terms were “Gynecologic Imaging Reporting and Data System” and “GI-RADS”. After reading the title and abstract of the studies, three reviewers (W Guo, Xiuhe Zou and HY Xu) decided whether the studies met the inclusion criteria independently. One experienced investigator (XL Ma) assessed the quality of the included studies and resolved the conflicts between three reviewers.

Inclusion and exclusion criteria

Inclusion criteria were that: (I) the patients had adnexal masses; (II) the adnexal mass had been verified by golden standards (biopsy or follow-up), and the minimal required follow-up time was 6 weeks; (III) the adnexal mass was evaluated by ultrasound and classified by the GI-RADS only; (IV) the article was written in English. The exclusion criteria were that: (I) studies only had abstracts or lacked sufficient data of the diagnostic effect; (II) the adnexal mass was evaluated by other imaging methods or methods combined with ultrasound; (III) the langue of articles was not English.

Data extraction

Two reviewers extracted these data: study characteristics, demographic characteristics, and outcomes from the selected articles. The details are as follows:

(I) Study characteristics: author, year of the study, country, amount of the patients and masses, ultrasound equipment, probe frequency, the technology of the ultrasound, and the golden standard;

(II) Demographic characteristics: age, the prevalence rate for adnexal malignancy, premenopausal rate;

(III) Outcomes: specificity, sensitivity, positive predictive value, negative predictive value, and value of true positive (TP), false positive (FP), true negative (TN), false negative (FN) for the ability of GI-RADS to recognize the adnexal malignancy (10).

Statistical analysis

We used Meta-Disc version 1.4 statistical software to calculate the pooled specificity, sensitivity, positive likelihood ratios (PLR), negative likelihood ratios (NLR), and diagnostic odds ratio (DOR) with 95% confidence intervals (CIs). Also, the summary receiver operator
characteristics (SROC) was constructed by this software to evaluate the relationship between specificity and sensitivity. We also used the inconsistency index ($I^2$) to detect heterogeneity among the included studies and would use the random effect model in the meta-analysis if $I^2$ was more than 50%. The Deeks’ funnel plot asymmetry test performed in Stata version 16.0 software (StatsCorp., College Station, TX) was used to value publication bias.

**Quality assessment**

To assess the included articles, we used the Quality Assessment of Diagnostic Accuracy Studies version 2 (QUADAS-2). The QUADAS-2 can evaluate the quality of studies from 14 items, four domains, and two aspects (11). According to the assessment results, these ten articles were reliable enough to conclude the ability of GI-RADS to differentiate adnexal masses. The QUADAS-2 was operated in Review Manager version 5.3 (Cochrane Informatics and Knowledge Management Department, Copenhagen, Denmark).

**Results**

**Study selection**

After removing the duplicates, we included 311 articles by the initial publication search and selected 77 studies for further filtration by reading the full text. Then, 67 of these publications failing to meet the inclusion criteria were excluded: 60 articles were not related to the subject; 4 articles did not use the GI-RADS to evaluate the adnexal masses; 3 articles did not use the ultrasound to detect the masses.

Additionally, though written by the same authors, two studies of different sample sizes were included.

Finally, we included a total of 10 articles in this study, and the selection flow chart was shown in Figure 1.

**Eligible study characteristics**

Ten selected studies were published from 2009 to 2019, and the details of these studies were shown in Table 1. These studies were conducted in 7 different countries. The total number of patients was 2,474, and that of masses was 2,723. The mean age of all the patients was 42.75 (age range, 6–89), and the prevalence rate of malignant tumors ranged from 13.45% to 41.83%. The type of the ultrasound equipment in these studies was Voluson 730 (GE Healthcare, Milwaukee, WI, USA), GE E8 EXPERT (GE Medical Systems, Zipf, Austria), GE S8 (GE Medical Systems, Zipf, Austria), iU22 digital scanner (Philips Healthcare, Bothell, WA, USA), Voluson E8 (GE Medical Systems, Zipf, Austria), SonoScape S40 Exp/S40 Pro/S40/S35 (Sonoscape Medical Hong Kong Co., Ltd), DC-7 Mindary (DC-7 Mindary Medical Ltd, China), ACUSON X150 ultrasound system (Siemens Health Care, Mountain View, CA, USA), ProSound Alpha 7 ultrasound (Hitachi Aloka Medical America, Inc. Germany). The golden standard of the differentiation of the gynecologic malignancy was biopsy and clinical symptoms during the follow-up.
Histology diagnosis was used as the golden standard of the differentiation of the gynecologic malignancy in all included studies. What’s more, in Tan Z, Basha MAA, and Behnamfar F’s studies, some of the patients had a spontaneous resolution in followed up by ultrasound study. Behnamfar F’s study defined the follow-up interval was 6 weeks, and the other two studies didn’t mention the specific follow-up time.

### Diagnostic value of differentiating malignancy

The pooled sensitivity of selected studies was 0.95 (95% CI: 0.94 to 0.97, I²=61.5%), and the pooled specificity was 0.86 (95% CI: 0.84 to 0.88, I²=94.3%). The pooled NLR and PLR were 0.06 (95% CI: 0.04 to 0.10, I²=59.3%), and 8.30 (95% CI: 4.93 to 13.97, I²=94.2%), respectively. The pooled DOR for GI-RADS was 174.59 (95% CI: 76.70 to 397.42, I²=74.6%). The SROC was shown in Figure 2, and the overall area under the curve (AUC) was 0.9806 (SE, 0.0160).

### Assessment of publication bias

The Deeks’ funnel plot asymmetry test was used to evaluate the publication bias in the included studies (Figure 3). The results showed that the diagnostic OR of GI-RADS was
not significantly biased in the differentiation of benign and malignant adnexal masses (P=0.6).

Quality assessment

The exact assessment results of the eligible studies were shown in Figure 4, and the quality of these studies was sufficient for analysis.

Discussion

This meta-analysis included ten studies from 7 countries to evaluate the diagnostic value of the GI-RADS for adnexal masses. Notably, Amor F. conducted two studies with
Figure 3 Risk of bias and applicability concerns: (A) graph; (B) summary.

different patients in 2009 and 2011. The diagnostic accuracy of ultrasound depended on the operators’ experience, and the difference rate between experts and routines was 15% (95% CI: 0–29%) (8,21-23). Thus, introducing the GI-RADS can strengthen the consistency among the different examiners and improve the diagnostic ability to differentiate the malignancy from the adnexal masses.

Our research systematically analyzed the data of the diagnostic performance of the GI-RADS in these ten studies. The pooled sensitivity and specificity were 0.95 and 0.86, respectively. The pooled NLR was 0.06, and PLR was 8.30. PLR >10 and NLR <0.1 mean high diagnostic accuracy (24). Although the pooled NLR was 0.06 and PLR 8.30 was in our study, the PLR value failed to draw a convincing conclusion of the presence of the adnexal malignancy. Furthermore, the diagnostic accuracy of GI-RADS evaluated by the AUC was 0.9806, indicating a high diagnostic value (25). All of these results showed that the GI-RADS could differentiate gynecological adnexal malignancy from adnexal masses.

There were several limitations to this study. Firstly, since I² of the pooled sensitivity and specificity were over 50%, we should consider the heterogeneity of data. However, lacking enough data in the included articles, we could not perform a sub-analysis. Besides, various types of ultrasound equipment used in different studies might account for heterogeneity among studies. Also, the age range of patients, the prevalence rate for adnexal malignancy, and menopausal proportions varied in these studies, which might increase the heterogeneity of the result. Finally, there were no factors in the included articles that allowed for subgroup analysis, which might hinder finding subgroups...
with higher diagnostic efficacy. Therefore, different ultrasound parameters are desired in subsequent studies to explore better diagnostic performance.

Although we excluded studies using other examine methods such as MRI and contrast-enhanced ultrasound (CEUS), their diagnostic power is worth further discussion. A previous study showed that when combining GI-RADS with 3D-CEUS, the diagnostic sensitivity, specificity, AUC, and kappa value were 94.2%, 95.5%, 0.969, and 0.872, respectively (26). Compared with using GI-RADS alone, the combination of GI-RADS and 3D-CEUS had higher specificity and lower sensitivity and smaller AUC according to this research, partly because the 3D-CEUS could not show real-time hemodynamic variance in the masses. Due to the limited number of studies included, large-scale and multicentre studies are required to compare these two methods.

Furthermore, the detectable rate of malignant adnexal masses detected by the ultrasound was low (8–20%), so MRI was introduced to a reporting data system named Ovarian Adnexal Reporting Data System Magnetic Resonance Imaging (O-RADS MRI). Based on a multicentre cohort study, the sensitivity, specificity, and AUC of O-RADS MRI were 0.93 (95% CI: 0.89–0.96), 0.91 (95% CI: 0.89–0.93), and 0.961 (95% CI: 0.948–0.971) (27-29). The diagnostic performance and accuracy of MRI were similar to that of GI-RADS.

In conclusion, GI-RADS is useful and reliable to improve the diagnostic ability of ultrasound to differentiate adnexal masses and malignant tumors. It is an effective method to settle the inconsistency among different ultrasound operators. With the development of the novel diagnostic tool, combining the GI-RADS with other examinations is a trend. However, further studies are required to confirm the results of our research.

Acknowledgments

We would like to show our deepest thanks to Dr. Zikuan Song and Dr. Yinan Xiao, who have helped us to revise our manuscript.

Funding: None.

Footnote

Reporting Checklist: The authors have completed the PRISMA reporting checklist. Available at http://dx.doi.org/10.21037/atm-20-5170

Peer Review File: Available at http://dx.doi.org/10.21037/atm-20-5170

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.
The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References
