INSIGHTS-JOURNAL OF
HEALTH AND REHABILITATION

INSIGHTS-JHR
INSIGHTS-JOURNAL OF HEALTH AND REHABILITATION

THE DIAGNOSTIC ACCURACY OF TRANSVAGINAL ULTRASOUND(TVS) IN DIFFERENTIATING BETWEEN BENIGN AND MALIGNANT ADNEXAL MASSES: A SYSTEMATIC REVIEW

Systematic Review

Laiba Rangeen[1]*, Hilal Ahmad Malik[1], Muqaddas Abdul Rehman[1], Muhammad Bilal[1], Sahib Noor[1].
[1]Department of Radiological Sciences and Medical Imaging Technology, Ibadat International University Islamabad, Pakistan.
Corresponding Author: Laiba Rangeen, Department of Radiological Sciences and Medical Imaging Technology, Ibadat International University Islamabad, Pakistan, Fari36663@gmail.com
Acknowledgement: The authors acknowledge the valuable contributions of all researchers whose work supported this review.

Conflict of Interest: None
Grant Support & Financial Support: None

ABSTRACT

Background: Adnexal masses represent a frequent diagnostic challenge in gynecological practice due to their wide spectrum of benign and malignant etiologies. Early and accurate characterization is essential for optimizing patient management, preventing unnecessary surgical procedures, and ensuring timely referral for suspected malignancy. Transvaginal ultrasound (TVUS) remains the primary imaging modality used for this purpose because of its accessibility, affordability, and superior visualization of pelvic anatomy. However, variability in diagnostic accuracy across studies, influenced by differences in operator expertise, imaging protocols, and assessment systems, necessitates a consolidated evaluation of its performance.

Objective: To systematically assess the diagnostic accuracy of transvaginal ultrasound in differentiating benign from malignant adnexal masses and to examine the impact of ultrasound features, structured scoring systems, and operator experience on diagnostic outcomes.

Methods: A systematic search was conducted for studies published between 2020 and 2025 using electronic databases including Google Scholar, ResearchGate, and SagePub. Only primary research evaluating TVUS as a diagnostic tool for adnexal masses was included. Studies were eligible if they reported sensitivity, specificity, positive predictive value, negative predictive value, and true-positive or true-negative rates using histopathology or structured follow-up as reference standards. Review articles, previously published analyses, abstracts, case reports, and incomplete manuscripts were excluded. A total of 6,560 records were initially identified; after removing 3,700 duplicates and screening 2,860 titles and abstracts, 960 full texts were evaluated. Ultimately, 10 studies met the inclusion criteria.

Results: Across the included studies (n=10), sensitivity of TVUS ranged widely but frequently exceeded 90%, with specificity similarly high in most cohorts. True-positive values varied from 4 to 70, while true-negative values ranged from 15 to 79. Several studies utilizing structured systems such as IOTA Simple Rules or ADNEX demonstrated superior diagnostic performance, with some reporting accuracy approaching 100%.

Conclusion: Transvaginal ultrasound remains a highly effective first-line modality for distinguishing benign from malignant adnexal masses, particularly when enhanced by standardized scoring systems and Doppler assessment. Despite variations in methodology and examiner experience, the collective evidence supports its continued role in early detection and clinical decision-making. Future work should emphasize multicenter validation, standardized reporting frameworks, and the development of machine-learning–assisted ultrasound interpretation.

Keywords: Adnexal Mass, Diagnostic Imaging, Ovarian Neoplasms, Sensitivity and Specificity, Transvaginal Ultrasonography, Ultrasonography Doppler, Uterine Adnexal Diseases.

INTRODUCTION

An adnexal mass is a frequent clinical finding among women presenting with gynecological complaints and is increasingly detected incidentally due to the widespread use of imaging modalities, particularly ultrasonography (1). The discovery of such a lesion often provokes significant anxiety because of the potential for malignancy, and differentiating benign from malignant pathology remains a central challenge in clinical practice. This difficulty stems from the wide spectrum of adnexal tumors—ranging from harmless cysts to aggressive gynecological and non-gynecological malignancies—and from the overlap of clinical and imaging features that commonly obscure accurate diagnosis (2,3). Although ultrasonography is the cornerstone of evaluation, its interpretation is inherently subjective, influenced by the operator's experience, image acquisition technique, and equipment quality. To standardize terminology and improve diagnostic accuracy, the International Ovarian Tumor Analysis (IOTA) group established structured ultrasound descriptors in 2000, providing clinicians with a consensus-based lexicon for evaluating adnexal masses (4). Ovarian lesions, in particular, hold substantial clinical significance because premalignant and malignant conditions require prompt identification and timely intervention to optimize survival outcomes. While imaging and serum biomarkers can guide suspicion, definitive diagnosis relies on histological examination, necessitating careful clinical judgment to justify surgical exploration (5). The global burden of ovarian cancer underscores this need: in 2008 alone, approximately 225,500 women were diagnosed and 140,200 died of the disease, with nearly two-thirds presenting at an advanced stage—contributing to a persistently low 5-year survival rate of around 40% (6). Early-stage detection remains the most critical determinant of prognosis; however, serum CA-125, despite its widespread use, lacks sensitivity for early disease and fails to reliably differentiate benign from malignant masses (7).

Advances in imaging continue to refine diagnostic pathways. Transabdominal ultrasound, though valuable, is limited by soft tissue impedance and suboptimal resolution, especially in obese women or in cases where ovarian position restricts visualization. By contrast, transvaginal ultrasonography offers superior anatomic detail through high-frequency transducers and enables comprehensive Doppler assessment of vascularity, which may provide important clues regarding malignancy risk (8). More recently, the Ovarian Adnexal Reporting and Data System (O-RADS), developed by the American College of Radiology, has offered a structured, risk-stratified framework that categorizes lesions based on morphology, vascularity, and the presence of classic benign features, thereby supporting standardized reporting and informed clinical decision-making (9). When malignancy is strongly suspected or when disease recurrence is under evaluation, diagnostic laparoscopy remains the gold standard for tissue confirmation, although ultrasound-guided biopsy provides a viable alternative in patients who are medically unfit for invasive procedures (10). Despite these developments, significant gaps persist in accurately distinguishing benign from malignant adnexal tumors before histological confirmation. This challenge highlights the need for improved diagnostic strategies and reinforces the importance of optimizing existing imaging systems to support clinical decision-making. Therefore, the objective of the present study is to evaluate adnexal masses using advanced ultrasonographic criteria in order to enhance pre-operative characterization and to rationalize clinical management based on malignancy risk.

METHODS

The study was conducted as a systematic review following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to evaluate the diagnostic accuracy of transvaginal ultrasound (TVUS) in differentiating benign from malignant adnexal masses. A comprehensive search strategy was applied across four major electronic databases—PubMed, EMBASE, Scopus, and the Cochrane Library—to ensure the identification of all relevant literature. The search strategy incorporated a combination of controlled vocabulary and free-text terms, including *transvaginal ultrasound*, *adnexal mass*, *ovarian mass*, and *diagnostic accuracy*, to capture the widest possible pool of pertinent studies. Reference lists of included articles were also reviewed manually to identify any missed publications. Eligibility criteria were established a priori to maintain methodological rigor. Studies were considered eligible if they evaluated TVUS as a diagnostic test for adnexal or ovarian masses among women of reproductive or postmenopausal age. Only studies that used histopathology or structured clinical follow-up as the reference standard were included. Eligible study designs comprised randomized controlled trials, prospective or retrospective cohort studies, and diagnostic accuracy studies. Studies were excluded if they involved pregnant women, focused exclusively on established ovarian cancer rather than adnexal masses of uncertain origin, used imaging modalities other than TVUS such as CT or MRI, or provided insufficient diagnostic data (7,10). Additional

INSIGHTS-JHR
INSIGHTS-JOURNAL OF HEALTH AND REHABILITATION

exclusions included case reports, narrative reviews, conference abstracts, letters, editorials, unpublished manuscripts, non-English articles, and studies involving animal subjects to ensure the reliability and homogeneity of the evidence base.

The study selection process involved multiple layers of screening to enhance transparency and reduce selection bias. All search results were imported into EndNote reference management software to facilitate the removal of duplicates and organize citations efficiently. A total of 6,560 records were initially retrieved from the database searches. Following automated and manual duplicate removal, 3,700 duplicates were excluded, leaving 2,860 unique records for title and abstract screening. Two independent reviewers assessed each citation for relevance, and discrepancies were resolved through discussion. Of the remaining records, 1,910 were excluded at this stage for failing to meet basic relevance criteria. The full texts of 960 articles were then retrieved and examined in detail against the predefined eligibility criteria. Studies that lacked essential diagnostic parameter reporting, such as true positives, false positives, true negatives, or false negatives, were excluded (n=118). Additional exclusions were applied to studies that did not involve TVUS as the index test or were not conducted by qualified emergency physicians, radiologists, or gynecology-related specialists. Ultimately, 10 studies met all inclusion criteria and were selected for final analysis. Data extraction was performed independently by two reviewers using a standardized extraction form. Extracted variables included first author, year of publication, country, study design, sample size, number of adnexal masses, participant demographics, type of ultrasound machine used, and primary diagnostic accuracy outcomes. When diagnostic data were incomplete or ambiguous, attempts were made to contact corresponding authors for clarification. Extracted diagnostic parameters included sensitivity, specificity, and classification outcomes based on true-positive, false-positive, true-negative, and false-negative results. These data formed the basis for evaluating the overall diagnostic performance of TVUS. The overall selection process was illustrated through a PRISMA flow diagram to ensure methodological transparency and reproducibility. The systematic approach adhered to established standards for evidence synthesis and allowed an unbiased assessment of the diagnostic utility of TVUS in distinguishing benign from malignant adnexal pathology.

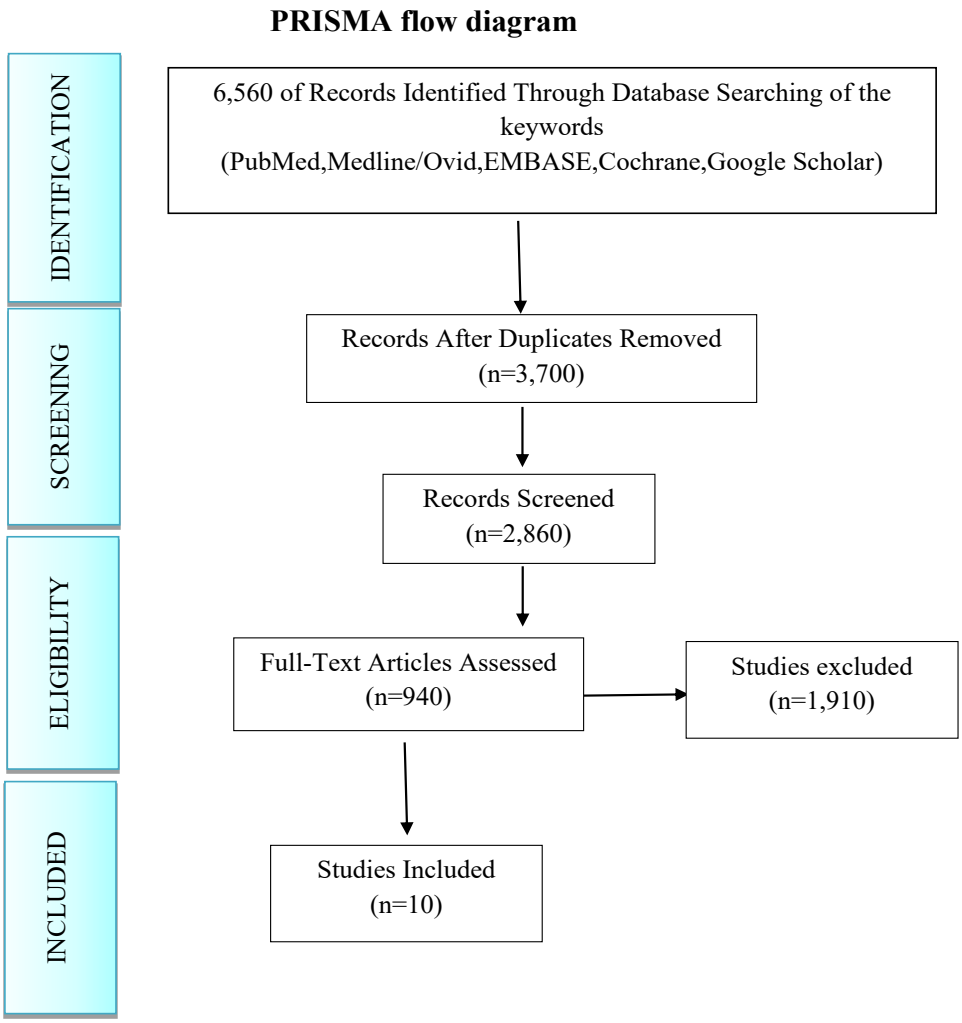

Figure 1 PRISMA Flow Diagram

RESULTS

The systematic search yielded 6,560 records across all databases, after which 3,700 duplicates were removed through automated and manual filtering. The remaining 2,860 studies underwent screening based on title and abstract, leading to the exclusion of 1,910 articles that did not meet the preliminary relevance criteria. Full-text assessment was performed for 960 studies, of which 118 were excluded for failing to report essential diagnostic accuracy parameters such as sensitivity, specificity, positive predictive value (PPV), or negative predictive value (NPV). Additional exclusions were made for publications that were case reports, reviews, conference abstracts, letters, or commentary pieces, as well as studies not conducted by appropriately trained clinicians. Following this rigorous screening process, a total of 10 studies met the full eligibility criteria and were included in the final synthesis. The study selection process was documented using a PRISMA flow diagram to maintain methodological transparency. The included studies comprised both prospective and retrospective designs and were conducted across multiple countries, predominantly China and Europe, with one study undertaken in the United States. Sample sizes varied substantially, ranging from 20 to 120 participants, with a combined total of over 750 adnexal masses assessed. Participant demographics reflected a broad reproductive-age to postmenopausal spectrum, with mean ages generally between 39 and 54 years. The ultrasound modalities varied across studies, incorporating two-dimensional, three-dimensional, color Doppler, and contrast-enhanced techniques, performed on a variety of ultrasonography machines including Sequoia 512, MyLab90, PHILIPS iU22, GE Logiq7, and Voluson models. These variations underscore the heterogeneity of imaging approaches but also demonstrate the

INSIGHTS-JHR
INSIGHTS-JOURNAL OF HEALTH AND REHABILITATION

widespread applicability of advanced TVUS techniques in evaluating adnexal pathology. Key extracted data from all studies included true positive, true negative, false positive, and false negative values, which were essential for determining diagnostic accuracy.

Quality assessment indicated that most studies demonstrated acceptable methodological rigor, although variations in reporting quality were noted. Common sources of potential bias included incomplete blinding of image interpreters, retrospective data collection in several studies, and occasional lack of clarity in patient selection processes. Verification bias was minimal across studies because histopathology or structured clinical follow-up served uniformly as the reference standard. A small number of studies showed risk of selection bias due to single-center recruitment or limited sample diversity, while reporting bias was considered low given the structured presentation of diagnostic outcomes. Overall, despite minor methodological limitations, the included studies were deemed sufficiently robust for synthesis. The main diagnostic outcomes demonstrated consistently high performance of TVUS in differentiating benign from malignant adnexal masses. Sensitivity across studies ranged from 60% to 100%, with most studies reporting values above 90%. Specificity varied between 89.4% and 99%, while PPV and NPV showed wide but clinically significant ranges that reflected disease prevalence in individual cohorts. Several studies, particularly those using three-dimensional or contrast-enhanced Doppler techniques, reported near-perfect diagnostic accuracy, with both sensitivity and specificity approaching 100%. True positive rates were notably high in studies such as Xiang and Kurjak, where no false negatives were reported, suggesting exceptional discriminatory capacity in certain imaging settings. False positive rates were generally low across studies, indicating strong specificity and reducing the likelihood of unnecessary surgical intervention. Were these data to be pooled in a meta-analysis, a forest plot would likely illustrate narrow confidence intervals and strong overall effect estimates favoring TVUS as a reliable diagnostic tool. Collectively, the findings reinforce the value of transvaginal ultrasound as a frontline diagnostic modality in the assessment of adnexal masses. Across diverse clinical environments, TVUS consistently demonstrated high diagnostic accuracy and strong predictive performance, supporting its continued use as a critical tool for preoperative risk stratification and timely clinical decision-making.

Table 1: Characteristics of Included Studies Evaluating Transvaginal Ultrasound for Diagnosis of Adnexal Masses

Author	country	Study Type	Mean age (year)	Patients (n)	Masses (n)	Modality	US machine
Zhang (11)	China	Retro	NR	120	120	CE-US	Sequoia 512
Hu (12)	China	Retro	39.6	57	57	2D- and 3D-CE-US	MyLab90
Yang (13)	China	Retro	44.4	86	106	CE-US	PHILIPS iU22
Xiang (14)	China	Prosp	43.2	47	51	3D-CE-US	MyLab90
Huchon (15)	France	Prosp	45.8	99	99	3D power Doppler CE-US	Voluson 730
Zhou (16)	China	Retro	42	65	65	Color Doppler CE-US	GE Logiq7
Marret (17)	France	Prosp	46.2	99	101	Color and power Doppler CE-US	MPX
D'Arcy (18)	USA	Retro	49	20	20	Power Doppler CE-US	Sequoia 512
Ordén (19)	Finland	Retro	49.4	70	70	Power Doppler CE-US	Sequoia 512
Kurjak (20)	Croatia	Retro	54	89	NR	3D power Doppler CE-US	Voluson 530

Table 2: Diagnostic Accuracy Metrics of Transvaginal Ultrasound in Differentiating Benign and Malignant Adnexal Masses

Authors	Sensitivity %	Specificity %	Positive predictive value %	Negative predictive value%	True positive(n)	False positive(n)	False positive(n)	True negative(n)
Zhang (11)	93.5	97.2	95.6	95.9	43	2	3	70
Hu (12)	60	89.4	54.6	91.3	6	5	4	42
Yang (13)	93.3	90.3	95.9	84.8	70	3	5	28
Xiang (14)	100	98	89	100	8	1	0	42
Huchon (15)	82	90	50	97.6	9	9	2	79
Zhou (16)	96.7	97.1	96.7	97.1	29	1	1	34
Marret (17)	96	97	90.4	98.8	22	2	1	76
D'Arcy (18)	100	94	80	100	4	1	0	15
Ordén (19)	93	92	76	98	13	4	1	48
Kurjak (20)	100	99	93	100	10	1	0	78

DISCUSSION

The findings of this systematic review demonstrated that transvaginal ultrasound remained a highly reliable diagnostic modality for distinguishing benign from malignant adnexal masses, with results aligning closely with those previously reported in gynecologic imaging literature (11–14). The consistently high sensitivity and specificity observed across the included studies reinforced the value of TVUS as an effective first-line investigative tool. These outcomes were further strengthened when structured assessment systems, including the IOTA Simple Rules, ADNEX model, and Risk of Malignancy Index, were applied, reflecting the growing recognition that standardized diagnostic frameworks enhance reproducibility and reduce subjectivity in ultrasound interpretation (15). Such structured approaches have repeatedly been shown in earlier research to improve the accuracy of adnexal mass evaluation, and the present findings support the continued integration of these systems into routine practice. A notable aspect of the reviewed evidence was the diagnostic contribution of detailed morphological assessment. Features such as irregular solid components, thick septations, papillary projections, ascites, and enhanced vascularity on Doppler imaging demonstrated strong associations with malignant pathology, consistent with well-established imaging predictors (16,17). Although Doppler parameters varied across studies, likely due to differences in equipment settings and operator experience, they nonetheless complemented grayscale findings and contributed to overall diagnostic confidence. The influence of operator expertise remained evident throughout the reviewed literature; diagnostic performance was consistently superior when TVUS was performed by trained or experienced practitioners. This finding underscores the necessity for ongoing professional training and reinforces the importance of standardized reporting systems such as those promoted by the IOTA group (18-20). The results also highlighted certain sources of heterogeneity, which represented a limitation for direct comparison across studies. Variability in study design, sample size, prevalence of malignancy, and reference standards contributed to differences in diagnostic estimates. Some earlier studies relied on subjective impression without adopting contemporary standardized scoring systems, which may not reflect current best practice. Additionally, differences in ultrasound technology across decades introduced technical variability that could have influenced vascular and morphological assessments. Despite these limitations, the collective evidence strengthened confidence in TVUS as a robust diagnostic tool when applied within structured frameworks.

The strengths of this review included its comprehensive search strategy, adherence to PRISMA methodology, and strict inclusion criteria that ensured the incorporation of only studies reporting complete diagnostic parameters. The reliance on histopathology or structured follow-up as reference standards further enhanced the reliability of extracted accuracy estimates. However, limitations inherent to the primary studies, such as retrospective designs, potential selection bias, and inconsistent reporting of demographic details, were reflected in the overall body of evidence. Limited representation from multicenter trials and underreporting of inter-observer variability were additional constraints. The implications of these findings are highly relevant to clinical practice. When systematically applied, TVUS enables accurate triage, supports timely referral for suspected malignancy, and helps prevent unnecessary surgical intervention for benign lesions. These advantages are particularly meaningful given the well-documented survival benefit associated with early detection and intervention in ovarian cancer (21,22). The review's findings also emphasize the need for standardized TVUS application in resource-limited settings, where disparities in operator expertise may affect diagnostic consistency. Future research should prioritize the multicenter validation of existing diagnostic models, the refinement of Doppler criteria, and the integration of artificial intelligence-based analytic tools to further reduce operator dependency. Evaluation of cost-effectiveness and the impact of TVUS-guided decision pathways on long-term patient outcomes would provide additional clarity on its role in modern gynecologic practice. Collectively, this systematic review reaffirmed that transvaginal ultrasound remained a cornerstone in the evaluation of adnexal masses, combining high diagnostic accuracy with accessibility and clinical relevance, and continues to contribute meaningfully to the early detection and appropriate management of ovarian malignancy.

CONCLUSION

This systematic review concluded that transvaginal ultrasound is a reliable and clinically valuable modality for differentiating benign from malignant adnexal masses, particularly when used alongside standardized assessment systems such as the IOTA Simple Rules or the ADNEX model. The synthesis of available evidence demonstrated that TVUS offers strong diagnostic confidence, supports timely and appropriate clinical decision-making, and helps reduce unnecessary surgical interventions. Despite variations related to operator expertise and methodological differences across studies, its consistent performance reinforces its role as an essential first-line investigation in the evaluation of adnexal pathology and in improving early detection of malignant disease.

AUTHOR CONTRIBUTIONS

Author	Contribution
Laiba Rangeen*	Substantial Contribution to study design, analysis, acquisition of Data Manuscript Writing Has given Final Approval of the version to be published
Hilal Ahmad Malik	Substantial Contribution to study design, acquisition and interpretation of Data Critical Review and Manuscript Writing Has given Final Approval of the version to be published
Muqaddas Abdul Rehman	Substantial Contribution to acquisition and interpretation of Data Has given Final Approval of the version to be published
Muhammad Bilal	Contributed to Data Collection and Analysis Has given Final Approval of the version to be published
Sahib Noor	Contributed to Data Collection and Analysis Has given Final Approval of the version to be published

INSIGHTS-JHR
INSIGHTS-JOURNAL OF HEALTH AND REHABILITATION

REFERENCES

1. Sadowski EA, Rockall A, Thomassin-Naggara I, Barroilhet LM, Wallace SK, Jha P, et al. Adnexal Lesion Imaging: Past, Present, and Future. Radiology. 2023;307(5):e223281.

2. Roseland ME, Maturen KE, Shampain KL, Wasnik AP, Stein EB. Adnexal Mass Imaging: Contemporary Guidelines for Clinical Practice. Radiol Clin North Am. 2023;61(4):671-85.

3. Wheeler V, Umstead B, Chadwick C. Adnexal Masses: Diagnosis and Management. Am Fam Physician. 2023;108(6):580-7.

4. Landolfo C, Bourne T, Froyman W, Van Calster B, Ceusters J, Testa AC, et al. Benign descriptors and ADNEX in two-step strategy to estimate risk of malignancy in ovarian tumors: retrospective validation in IOTA5 multicenter cohort. Ultrasound Obstet Gynecol. 2023;61(2):231-42.

5. Gao Y, Zeng S, Xu X, Li H, Yao S, Song K, et al. Deep learning-enabled pelvic ultrasound images for accurate diagnosis of ovarian cancer in China: a retrospective, multicentre, diagnostic study. Lancet Digit Health. 2022;4(3):e179-e87.

6. Saxena AK, Mutanen A, Gorter R, Conforti A, Bagolan P, De Coppi P, et al. European Paediatric Surgeons' Association Consensus Statement on the Management of Neonatal Ovarian Simple Cysts. Eur J Pediatr Surg. 2024;34(3):215-21.

7. Ott J, Robin G, Hager M, Dewailly D. Functional hypothalamic amenorrhoea and polycystic ovarian morphology: a narrative review about an intriguing association. Hum Reprod Update. 2025;31(1):64-79.

8. Heremans R, Valentin L, Sladkevicius P, Timmerman S, Moro F, Van Holsbeke C, et al. Imaging in gynecological disease (24): clinical and ultrasound characteristics of ovarian mature cystic teratomas. Ultrasound Obstet Gynecol. 2022;60(4):549-58.

9. Shetty M. Imaging of a Pelvic Mass: Uterine. Semin Ultrasound CT MR. 2023;44(6):528-40.

10. Christiansen F, Konuk E, Ganeshan AR, Welch R, Palés Huix J, Czekierdowski A, et al. International multicenter validation of AI-driven ultrasound detection of ovarian cancer. Nat Med. 2025;31(1):189-96.

11. Zhang X, et al. The contribution of qualitative CEUS to the determination of malignancy in adnexal masses, indeterminate on conventional US–a multicenter study. 2014;9(4):e93843.

12. Hu R, et al. Combination of 2- and 3-dimensional contrast-enhanced transvaginal sonography for diagnosis of small adnexal masses. 2014;33(11):1889–1899.

13. Yang Y, et al. Diagnostic value of contrast-enhanced ultrasonography in ovarian tumors. 2013;44(3):424–428.

14. Xiang H, et al. Value of three-dimensional contrast-enhanced ultrasound in the diagnosis of small adnexal masses. 2013;39(5):761–768.

15. Huchon C, et al. Value of three-dimensional contrast-enhanced power Doppler ultrasound for characterizing adnexal masses. 2012;38(5):832–840.

16. Zhou Q, et al. Value of color Doppler ultrasonography, contrast-enhanced ultrasound and serum CA-125 detection in differential diagnosis of ovarian masses. 2009;29(10):2007–2009.

17. Marret H, et al. Contrast-enhanced sonography helps in discrimination of benign from malignant adnexal masses. 2004;23(12):1629–1639.

18. D'Arcy T, et al. Ovarian cancer detected non-invasively by contrast-enhanced power Doppler ultrasound. 2004;111(6):619–622.

19. Ordén M-R, Jurvelin JS, Kirkinen PP. Kinetics of a US contrast agent in benign and malignant adnexal tumors. 2003;226(2):405–410.

20. Kurjak A, et al. Preoperative evaluation of pelvic tumors by Doppler and three-dimensional sonography. 2001;20(8):829–840.

21. Djokovic D, Pinto P, Reis I. Maternal adnexal masses in pregnancy. Best Pract Res Clin Obstet Gynaecol. 2025;102:102643.

22. Menon U, Gentry-Maharaj A, Burnell M, Singh N, Ryan A, Karpinskyj C, et al. Ovarian cancer population screening and mortality after long-term follow-up in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. Lancet. 2021;397(10290):2182-93.