Diagnostic accuracy of transvaginal ultrasound examination for assigning a specific diagnosis to adnexal masses: A meta-analysis

XIYA ZHANG[1], XUAN MENG[2], TING DOU[3] and HUI SUN[4]

[1]Department of Special Inspection, Shandong Qingdao Hospital of Integrated Traditional and Western Medicine, Qingdao, Shandong 266002; [2]Department of Ultrasound, Qingdao Fuwai Cardiovascular Hospital, Qingdao, Shandong 266021; [3]Department of Ultrasound Medicine, The Third People's Hospital of Qingdao, Qingdao, Shandong 266041; [4]Department of Special Inspection, Qingdao Women and Children's Hospital, Qingdao, Shandong 266000, P.R. China

Received February 24, 2020; Accepted September 18, 2020

DOI: 10.3892/etm.2020.9395

Abstract. Transvaginal ultrasound (TVUS) is a standard imaging modality for differentiating patients with benign or malignant suspected adnexal mass. To date, numerous studies have assessed the diagnostic accuracy of TVUS in various settings but with variable results. Therefore, the purpose of the present study was to perform a meta-analysis to evaluate the diagnostic accuracy of TVUS for the differentiation of adnexal masses. An electronic search in the Medline, Scopus, Cochrane and Embase databases from inception till November 2019 was carried out. Meta-analysis was performed to obtain pooled sensitivity and specificity of TVUS to distinguish malignant from benign adnexal masses. The quality assessment of diagnostic accuracy studies-2 tool was used to assess the quality of trials. A total of 41 studies with 18,391 patients were included. The pooled sensitivity and specificity of TVUS was 92% (95% CI: 90-94%) and 89% (95% CI: 85-92%), respectively. The area under the receiver operating characteristic curve was 0.96 (95% CI: 0.84-1.00). There was considerable heterogeneity with a statistically significant chi-square test (P<0.001) and I^2 of 99%. Meta-regression results indicated that index test standards, patient selection bias and study design were potential sources of heterogeneity (P<0.05). The funnel plot was symmetrical and low publication bias was confirmed by an insignificant Deek's test (P=0.90). The present systematic review and meta-analysis indicated that TVUS is useful in differentiating between benign and malignant tumours among patients with suspected adnexal mass with high sensitivity and specificity.

Correspondence to: Dr Hui Sun, Department of Special Inspection, Qingdao Women and Children's Hospital, 6 Tongfu Road, Qingdao, Shandong 266000, P.R. China
E-mail: m1028120753@163.com

Key words: adnexal tumour, meta-analysis, transvaginal ultrasonography, validation studies

Introduction

An adnexal tumour is defined as an enlarged structure within the adnexa of the uterus (1). It represents a spectrum of benign and malignant conditions that may originate from either gynaecological or non-gynaecological sources (2). The pathology is usually an incidental finding diagnosed during a routine clinical examination or may be present in females with any gynaecological complaint (3). Since adnexal masses may present with a wide range of symptoms, it is frequently difficult to differentiate benign tumours from other malignant lesions such as ovarian cancer (2).

Cross-sectional imaging strategies have a major role in managing patients with adnexal tumours, as they are able to consistently differentiate between benign and malignant masses affecting the fallopian tube and ovary. It is also helpful in differentiating uterine and gastrointestinal pathologies from adnexal abnormalities (1). Early and accurate diagnosis of adnexal mass is essential for formulating a treatment plan. The ability of the imaging modality to differentiate between a benign and malignant nature of a lesion further influences the decision for the requirement of expectant management (cases with no symptoms or reproductive dysfunction) or the requirement of surgery (for borderline or invasive tumours) (4). Laparoscopic observation and histopathological examination are considered the gold standard for the specific diagnosis of adnexal mass (5). However, the invasive nature of the procedure is a significant limitation for its use in routine clinical practice.

Despite several advances and technological advancements in the field of radiodiagnosis, simple transvaginal ultrasound (TVUS) has been a standard procedure for the initial diagnosis of patients with adnexal mass (6,7). Several studies have reported that TVUS may also help in discriminating between benign and malignant adnexal masses and also to make a specific diagnosis (6,7). To the best of our knowledge, there have been no systematic efforts to perform a data

synthesis to evaluate the diagnostic accuracy of this method. Therefore, the purpose of the present study was to perform a meta-analysis to evaluate the diagnostic accuracy of TVUS for the differentiation of an adnexal mass as benign or malignant.

Materials and methods

Inclusion criteria. All types of studies examining the diagnostic accuracy of TVUS for a specific diagnosis of an adnexal mass and comparing it with standard laparoscopic or histopathological examination as the reference standard were considered. Studies were to report on sensitivity and specificity or provide data to calculate these values. Only full-text articles were included, while unpublished data were excluded. Studies with a sample size of <10 patients and case reports were also excluded.

Search strategy. An extensive and systematic electronic search was performed in the Medline, Scopus, Cochrane Library and Embase databases. Both medical subject headings along with free text terms were utilized for the literature search. The search terms used were as follows: 'Validation studies', 'adnexal mass', 'pattern recognition', 'transvaginal ultrasonography', 'benign adnexal mass', 'malignant adnexal mass', 'gynaecological disorders', 'sensitivity', 'specificity', 'diagnosis', 'adnexal lesions' and 'diagnostic accuracy studies'. The time limit for the search was from inception to November 2019 without any language restriction. Reference lists of primary studies were hand-searched to find any missed articles for inclusion in the review.

Selection of studies. Primary screening of title, keywords and abstracts was performed by two authors independently (XZ and XM). Full-text articles of the relevant entries were retrieved. These were further screened independently by the two authors (XZ and XM) for final inclusion in the review. Agreement between the two authors in making decisions related to inclusion or exclusion of studies was found to be excellent with a kappa value of 0.82. Disagreements during the selection of studies were resolved by consulting the third author (TD).

Data extraction and management. The primary investigator (XZ) performed data extraction using a data-extraction form. The following details were extracted: Study setting, study design, inclusion and exclusion criteria, reference standards, index test, total participants, comorbidities, mean age, sensitivity and specificity values. The extracted data were entered into STATA software. They were double-checked for correct entry by comparing the data in the review and the study reports. The following outcome measures were analysed in the review: Sensitivity, specificity, diagnostic odds ratio (DOR), likelihood ratio positive (LRP) and likelihood ratio negative (LRN).

Risk of bias assessment. The risk of bias for all of the included studies was assessed by two authors (XZ and XM) independently using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool (8). Studies were rated for patient selection bias, conduct and interpretation of index test and reference standard, as well as time interval (i.e. flow and timing) of the outcome assessments. The studies were graded as having low, high or unclear risk of bias for each domain.

Statistical analysis. The final estimate of sensitivity, specificity, LRN, LRP and DOR for TVUS was obtained using the bivariate meta-analysis method. The summary receiver operator characteristic curve was constructed from which area under the curve (AUC) was obtained. An AUC value closer to 1 was indicative of a better diagnostic value.

Forest plots were used to graphically represent the study-specific and pooled estimates of sensitivity and specificity. The clinical value of the TVUS was determined by the LR scattergram. The probability of a patient having a benign or malignant adnexal mass was tested using the Fagan plot. Heterogeneity was assessed graphically using bivariate boxplots and tested using the chi-square test and I^2 statistic. The source of heterogeneity was explored with meta-regression using study-related covariates such as the study design, year of publication, sample size, study region and quality-related factors. Publication bias was tested using Deek's test and graphically depicted by a funnel plot. The analysis was performed using the 'metandi' command package in STATA 14.2 software (StataCorp).

Results

Selection of studies. After database screening, a total of 2,442 records were retrieved, of which 927 records were from Medline, 813 from Scopus, 590 from Embase and 112 from the Cochrane library (Fig. 1). After the first stage of screening, 243 relevant studies were retained. The full text of these studies was examined against the eligibility criteria. In total, 41 studies with 18,391 participants satisfying the inclusion criteria were included in the present review (9-49).

Characteristics of included studies. The characteristics of the included studies are described in Table I. Of the included studies, 35 were prospective studies. Most of the studies were performed in high-income European countries such as the United Kingdom, Italy, Belgium and Spain. The average age of the participants ranged from 33.3 to 53.3 years. The sample size of the studies varied from 37 to 2,403 patients. All of the included studies used laparoscopy or laparotomy with histopathology as the reference standard for comparing the diagnostic accuracy of TVUS. The time interval between TVUS and the reference standard varied from 24 h to 12 weeks.

Risk of bias. The assessment of the risk of bias among the included studies is presented in Fig. 2. Of the studies, 90% had a low risk of bias for 'selection bias'. Furthermore, out of the 41 studies, 26 had a low risk of bias for 'conduct and interpretation of index test'. All of the studies had a low risk of bias for the 'conduct of reference standards test and interpretation'. A total of 32 studies had a low risk of bias concerning 'flow and interval between index and reference standard test' among the patients.

Diagnostic performance of TVUS. Analysis of data from the 41 studies provided a pooled sensitivity and specificity of

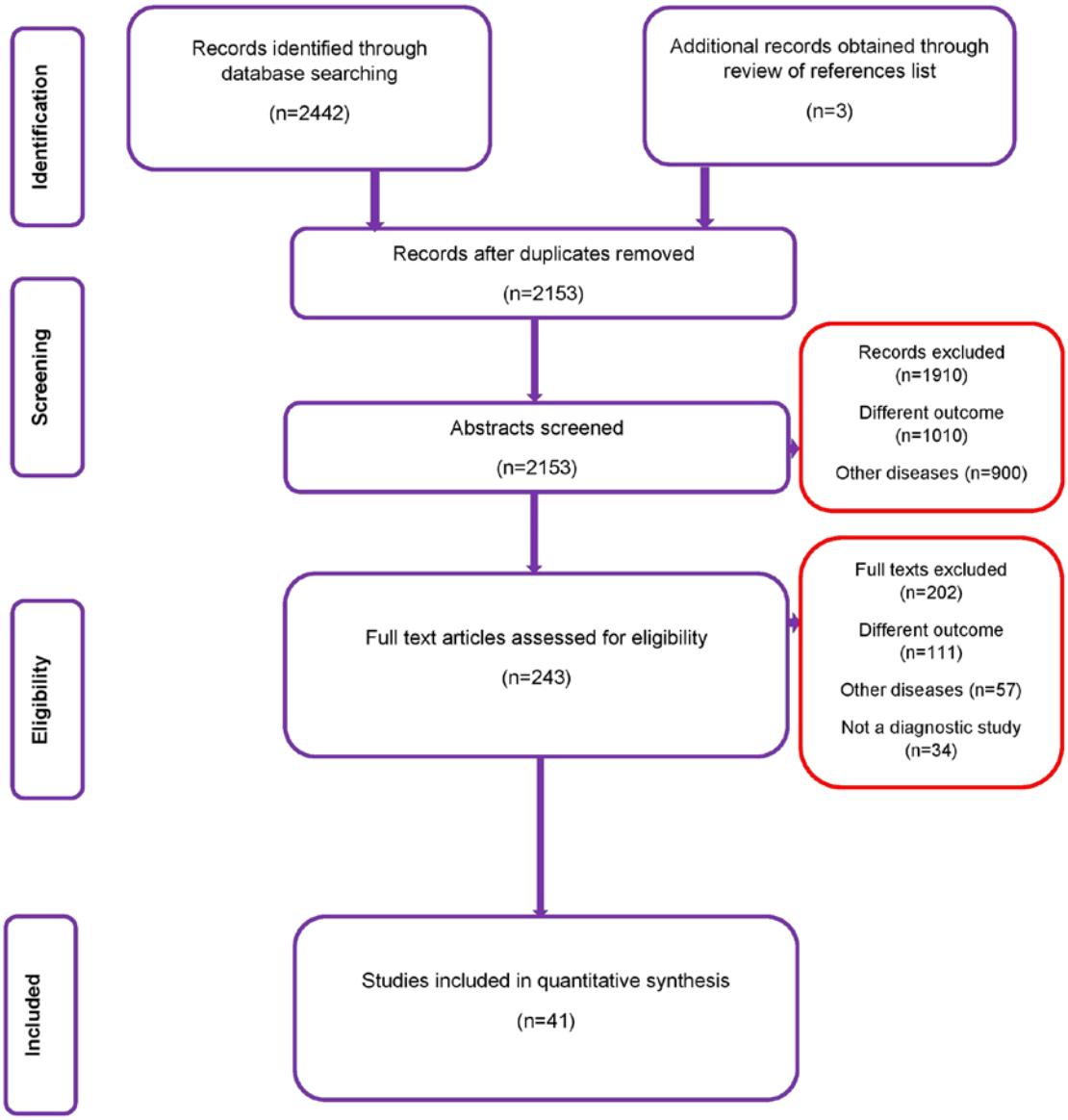

Figure 1. Search strategy of the review.

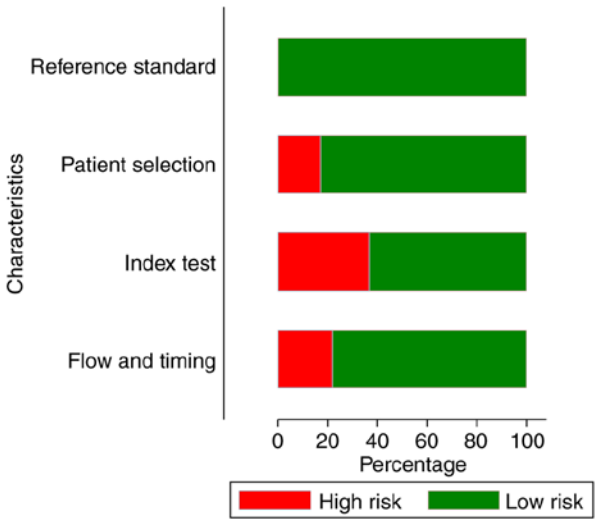

Figure 2. Quality assessment using the quality assessment of diagnostic accuracy studies-2 tool indicating the percentage risk of bias for each characteristic.

TVUS for differentiating benign and malignant adnexal mass of 92% (95% CI: 90-94%) and 89% (95% CI: 85-92%), respectively (Fig. 3). The DOR was 97 (95% CI: 65-147). The LRP was 8.3 (95% CI: 6.1-11.3) and the LRN was 0.09 (0.06-0.12). The upper right quadrant in the LR scatter diagram was occupied by these values, indicating that the TVUS may be used for confirmation only (Fig. 4). The AUC was 0.96 (95% CI: 0.84-1.00) (Fig. 5), indicating a highdiagnostic value. TVUS for adnexal mass had a good clinical value, as Fagan's nomogram had a significantly different post-test probability (positive, 80%; negative, 4%) compared to the pre-test probability (28%) (Fig. 6).

There was considerable heterogeneity with a statistically significant chi-square test result ($P<0.001$) and an I^2 value of 99%. As indicated in the bivariate box plot (Fig. 7), 4 studies were outside the circle, demonstrating a possibility of inter-study heterogeneity. Meta-regression for assessing the source of heterogeneity suggested that the selection domain, standards of index test conduct and study design

Table I. Characteristics of the included studies (n=41).

First author and year	Country	Study design	Sample size	Type of diagnostic modality	Gold standard comparator	Time interval between index test and reference standard	Mean age (years)	(Refs.)
Alcázar 2013	Spain	Prospective	340	Simple TVUS-based rules	Histopathology	3 weeks	42.1	(9)
Daemen 2011	19 Ultrasound centers in eight countries (Belgium, Canada, China, Czech Republic, Italy, Poland, Sweden, UK)	Prospective	1,938	Subjective assessment by grayscale TVUS	Histopathology	Not specified	Not specified	(10)
Ruiz de Gauna 2015	Spain	Prospective	247	Simple TVUS-based rules	Histopathology	3 weeks	43.6	(11)
Fathallah 2011	France	Prospective	122	Simple TVUS-based rules	Histopathology	Not specified	Not specified	(12)
Granberg 1990	Sweden	Prospective	180	Subjective assessment by	Histopathology grayscale TVUS	1 week to 1 month	Not specified	(13)
Guerriero 2010	Italy	Prospective	2,148	Subjective assessment by	Histopathology grayscale TVUS	Not specified	42	(14)
Hartman 2012	Brazil	Cross-sectional	110	Subjective assessment by	Histopathology grayscale TVUS	Mean time interval=64.4 days	Benign, 46.8 Malignant, 53.4	(15)
Jain 1993	United States of America	Prospective study	37	Endovaginal US	Surgery or laparoscopy	1-4 weeks	41.5	(16)
Jain 1994	United States of America	Prospective study	49	Endovaginal US	Surgery or laparoscopy	1-5 days	45	(17)
Knafel 2013	Poland	Prospective	226	Subjective assessment	Histopathology by TVUS	Not specified	47	(18)
Komatsu 1996	Japan	Retrospective study	82	TVUS	Histologic examination	2 weeks	45.9	(19)
Lucidarme 2010	France	Prospective	255	TVUS	Histologic examination	Not specified	Not specified	(20)
Mancuso 2004	Italy	Retrospective	125	TVUS	Histologic examination	Not specified	Not specified	(21)
Moszynski 2013	Poland	Retrospective	318	TVUS	Histologic examination	Not specified	Not specified	(22)
Nunes 2012	United Kingdom	Prospective	292	TVUS	Histologic examination	120 days	53.2	(23)
Nunes 2013	United Kingdom	Prospective	303	TVUS	Histologic examination	120 days	51	(24)
Radosa 2014	Germany	Retrospective	1,320	Pattern recognition by TVUS	Histopathology	Not specified	33.3	(25)
Romagnolo 2006	Italy	Prospective	221	Subjective assessment by TVUS	Histopathology	Not specified	Not specified	(26)
Roman 1997	USA	Prospective	226	Grayscale TVUS	Histopathology	Not specified	Not specified	(27)

Table I. Continued.

First author and year	Country	Study design	Sample size	Type of diagnostic modality	Gold standard comparator	Time interval between index test and reference standard	Mean age (years)	(Refs.)
Salle 1995	France	Prospective	91	Subjective assessment by TVUS	Histopathology	Not specified	Not specified	(28)
Sayasneh 2013	United Kingdom	Prospective multicentric study	255	2D grayscale TVUS	Histopathology	120 days	46	(29)
Sayasneh 2015	United Kingdom	Prospective multicentric study	313	2D grayscale TVUS	Histopathology	Not specified	47	(30)
Shetty 2017	India	Prospective	136	Pattern recognition by TVUS	Histopathology	Not specified	40.5	(31)
Shetty 2019	India	Prospective	183	IOTA Simple rules using TVUS	Histopathology	12 weeks	37.5	(32)
Silvestre 2015	Brazil	Prospective	75	IOTA Simple rules using TVUS	Histopathology	7 days	Not specified	(33)
Sohaib 2005	United Kingdom	Prospective	72	Subjective assessment by grayscale TVUS	Histopathology	Not specified	53	(34)
Sokalska 2009	Nine European centers	Retrospective	860	Grayscale TVUS	Histopathology	Not specified	37	(35)
Stein 1995	USA	Prospective	160	Grayscale TVUS	Histopathology	Not specified	114 patients were premenopausal (mean, 33 years; range, 13-53), 39 were perimenopausal or postmenopausal (mean, 57; range, 44-80) and eight had undergone hysterectomy (mean, 44 years; range, 33-61)	(36)
Strigini 1996	Italy	Prospective	109	TVUS	Laparotomy and histopathology	1 week	Median, 43	(37)
Tantipalakorn 2014	Thailand	Prospective	398	IOTA simple rules using TVUS	Pathological or operative findings	24 h	42.4	(38)

Table I. Continued.

First author and year	Country	Study design	Sample size	Type of diagnostic modality	Gold standard comparator	Time interval between index test and reference standard	Mean age (years)	(Refs.)
Testa 2014	18 centres in six countries (Sweden, Belgium, Italy, Poland, Spain and Czech Republic)	Prospective	2,403	IOTA Logistic regression model using TVUS	Histopathology	120 days	Not specified	(39)
Timmerman 1999	Belgium	Prospective	300	TVUS	Histopathology	Not specified	Premenopausal (mean age, 40; range, 22-57); postmenopausal (mean age, 65; range, 47-93)	(40)
Timmerman 2010	19 centres in eight European countries	Prospective	1,501	IOTA Simple rules using TVUS	Histopathology	120 days	46	(41)
Utrilla-Layna 2015	Spain	Prospective	367	Pattern recognition by TVUS	Histopathology	Not specified	46.5	(42)
Valentin 1999	Sweden	Prospective	173	TVUS	Histopathology	8 days	98 were premenopausal (median age, 37.5; range, 18-54), 70 were postmenopausal (median age, 66; range, 51-88; median 15 years past menopause with a range of 1-44 years), four had undergone hysterectomy (median age, 51.5; range, 44-66)	(43)
Valentin 2009	Nine European US centres	Prospective	534	TVUS	Histopathology	120 days	48.8	(44)
Van Calster 2007	Nine European US centres	Prospective	809	TVUS	Histopathology	Not specified	49	(45)

Table I. Continued.

First author and year	Country	Study design	Sample size	Type of diagnostic modality	Gold standard comparator	Time interval between index test and reference standard	Mean age (years)	(Refs.)
Van Gorp 2012	Belgium	Prospective	374	Subjective assessment by TVUS	Histopathology	Not specified	Patients with benign disease: Mean age, 46.2; patients with malignant disease: Mean age, 57.7	(46)
Van Holsbeke 2009	Belgium	Prospective	507	IOTA rules	Histopathology	Not specified using TVUS	40	(47)
Van Trappen 2007	United Kingdom	Prospective	142	TVUS	Histopathology	Not specified	50	(48)
Yamashita 1995	Japan	Prospective	80	TVUS	Histopathology	14 days	43	(49)

IOTA, International Ovarian Tumour Analysis; TVUS, transvaginal ultrasound; 2D, 2-dimensional.

were statistically significant sources of heterogeneity (P<0.05; Fig. 8). The funnel plot for assessing the publication bias was symmetrical and the low publication bias was confirmed by non-significant Deek's test (P=0.90 Fig. 9).

Discussion

Several imaging modalities are available for making a specific diagnosis among patients with adnexal mass (50). However, these modalities cannot replace histopathology or biopsy as the gold standard for diagnosis. Imaging modalities still have a major role in clinical practice as these are non-invasive and are able to significantly reduce the diagnostic delay and complications associated with invasive diagnostic techniques (51). Since TVUS is a widely used imaging tool for adnexal masses, it is important to evaluate the diagnostic accuracy of this modality in differentiating between benign and malignant adnexal mass.

In total, 41 studies with 18,391 participants met the eligibility criteria of the review. The majority of the included studies were prospective studies. Most of them were performed in high-income countries such as the United Kingdom, the USA, Italy and Sweden. The overall quality of evidence was high, as most of the studies had a low risk of bias for all of the four domains of the QUADAS tool.

The diagnostic accuracy of TVUS for differentiating benign and malignant adnexal masses has not been evaluated in any previous reviews, to the best of our knowledge. In the present first meta-analysis, the pooled estimate of the sensitivity of TVUS was 92% and the pooled specificity was 89% with a high diagnostic performance (AUC=0.96). This diagnostic accuracy almost reached that of other biomarkers and algorithms such as CA-125, human epididymis protein 4, Risk of Malignancy Index and the Risk of Ovarian Malignancy Algorithm (52-56).

In the LR scatter diagram, LRP and LRN occupied the left lower quadrant, indicating that the TVUS should be used as the test for confirmation only and not for exclusion. The clinical value of TVUS for adnexal mass was also good, as Fagan's nomogram indicated a significant increase in the post-test probability compared to the pre-test probability. However, while inferring these results, one must consider the quality and differences in methodology of the included studies, which may have influenced the study results. Hence, an analysis of inter-study heterogeneity amongst the included studies was also performed. The present analysis indicated significant inter-study heterogeneity with a significant chi-square test result and I^2 statistic. On further exploration of the source of heterogeneity via meta-regression, it was indicated that the study design, publication year and quality-associated characteristics had a significant influence on the inter-study variability. Deek's test and the funnel plot indicated that there was no significant publication bias among the studies reporting on the diagnostic accuracy of TVUS.

The present study has the following strengths. A comprehensive review was performed by including 41 studies with 18,391 patients to evaluate the diagnostic accuracy of TVUS in differentiating adnexal masses. To the best of our knowledge, the present study was the first to provide pooled estimates for the specific diagnosis of adnexal mass using TVUS. Furthermore, publication bias was determined to be

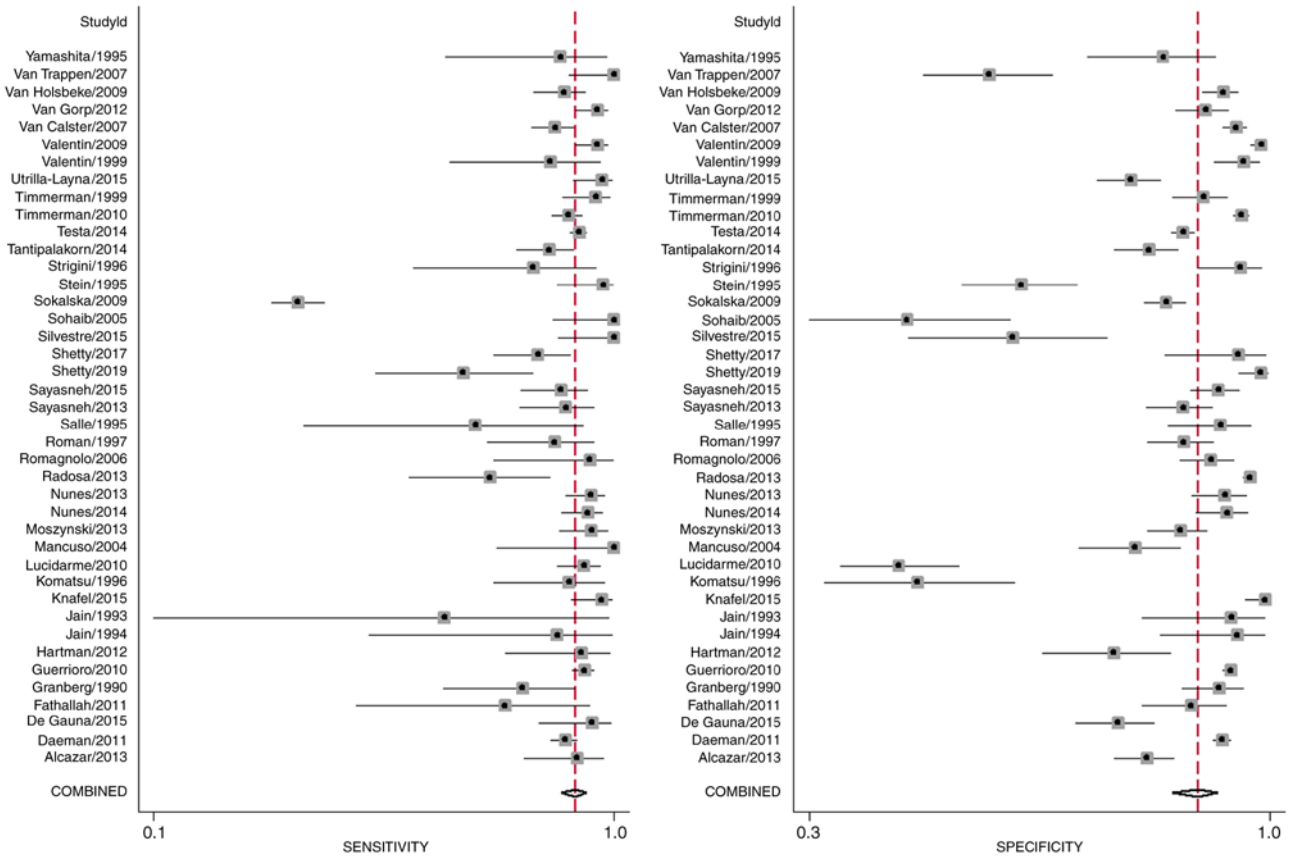

Figure 3. Forest plot indicating the pooled sensitivity and specificity for transvaginal ultrasound. A point estimate and 95% CI of each individual study is presented by a square and horizontal lines, respectively. Diamonds indicate combined sensitivity and specificity with the red line indicating the combined point estimate.

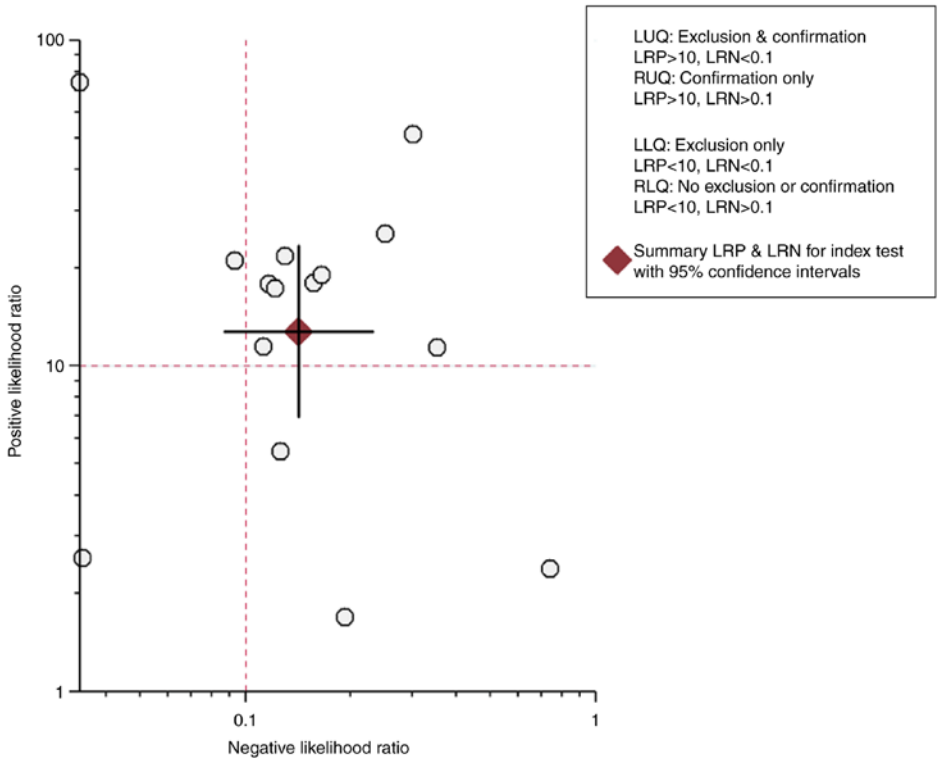

Figure 4. Likelihood ratio scatter diagram providing a summary matrix of the positive and negative likelihood ratio of TVUS for diagnosing adnexal masses. The scatter diagram suggested that TVUS is useful only for confirming the diagnosis of adnexal masses (when positive). LRN, likelihood ratio negative; LRP, likelihood ratio positive; LLQ, left lower quadrant; LUQ, left upper quadrant; RLQ, right lower quadrant; RUQ, right upper quadrant; TVUS, transvaginal ultrasound.

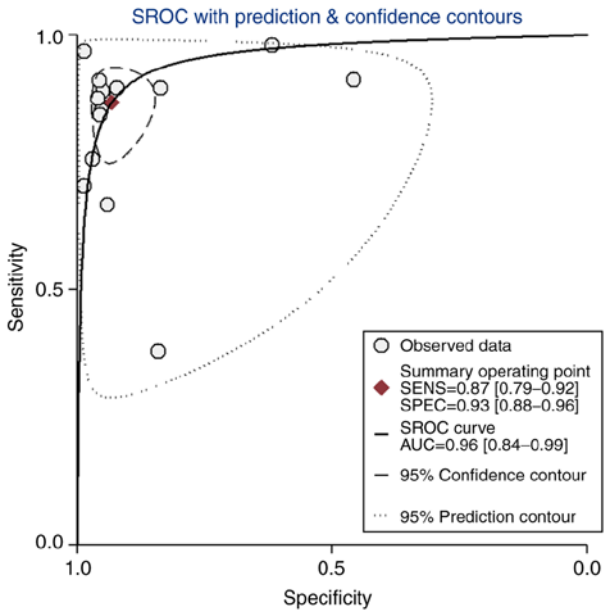

Figure 5. SROC with pooled estimates of SENS and SPEC and AUC for transvaginal ultrasound in the diagnosis of adnexal masses. An AUC value closer to 1 is indicative of a better diagnostic value. SROC, summary receiver operator characteristic curve; AUC, area under the curve; SENS, sensitivity; SPEC, specificity.

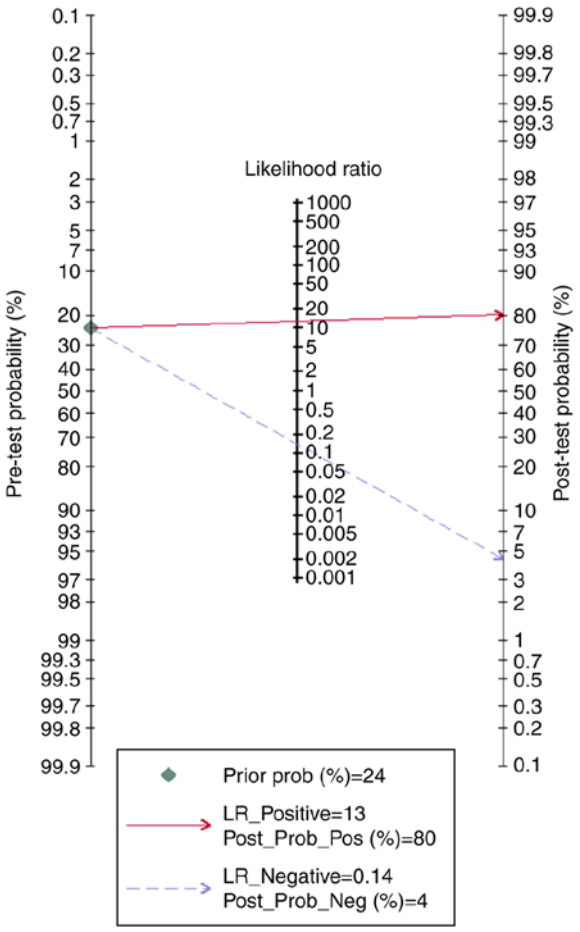

Figure 6. Fagan nomogram for calculating post-test probabilities of the disease from the LR of the test result. The straight line drawn from the patient's pre-test probability of the disease through the LR of the test result points to the post-test probability of the disease. LR, likelihood ratio; Prob, probability; Pos, positive; Neg, negative.

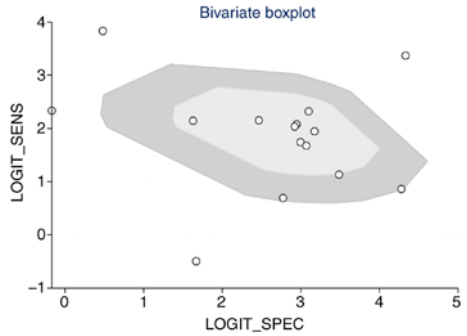

Figure 7. Bivariate boxplot of the sensitivity and specificity in the included studies. The inner oval region represents the median distribution of the data points and the outer oval represents the 95% confidence boundary. Studies outside this grey area are considered as outliers. LOGIT_SENS, logit sensitivity; LOGIT_SPEC, logit specificity.

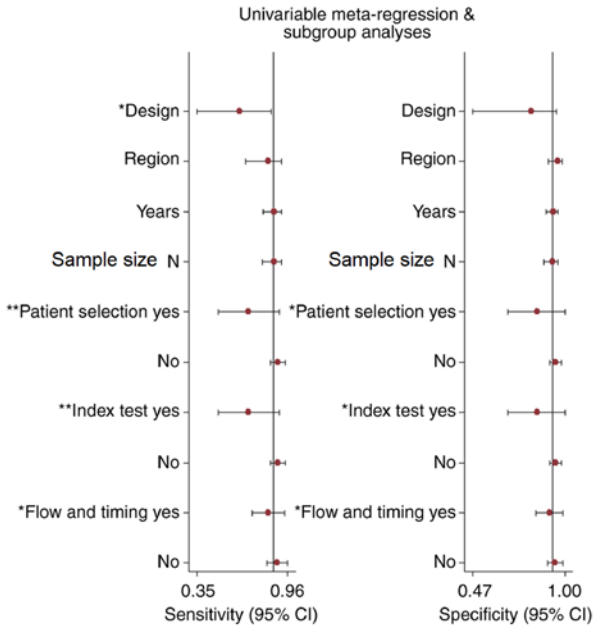

Figure 8. Meta-regression plot for different variables to explore sources of heterogeneity in the meta-analysis. Point estimates are depicted by red circles and 95% Cis are depicted by horizontal lines for each variable. Statistically significant results are marked by asterisks. *P<0.05; **P<0.01.

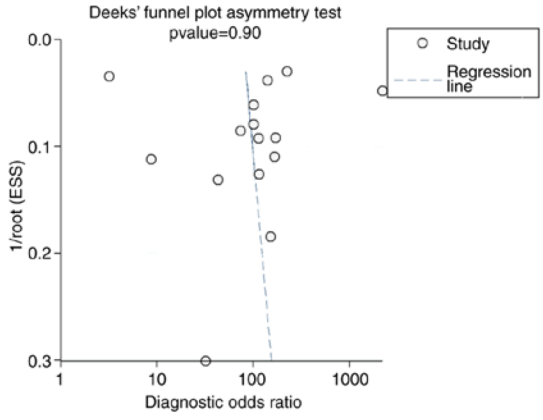

Figure 9. Deek's funnel plot for assessing publication bias in the included studies with super-imposed regression line. The statistically insignificant P-value (0.90) for the slope coefficient suggests symmetry in the data and a low likelihood of publication bias. ESS, effective sample size.

insignificant, which adds credibility to the results obtained in the present review. However, the present study also has certain limitations. First, certain studies had a high risk of bias, which may have influenced the pooled estimates. Furthermore, there was significant inter-study heterogeneity in the review. This limits the study's ability to interpret the pooled results. However, it was attempted to overcome this limitation by exploring the potential source of heterogeneity among the included studies by a meta-regression analysis.

Despite these limitations, the present study provided valuable insight regarding the diagnostic accuracy of non-invasive techniques for differentiating benign and malignant adnexal masses. While TVUS had good sensitivity and specificity, it can only almost reach the SnNout triage test criteria for sensitivity. It cannot meet the SpPin criteria for the specificity of a diagnostic test (57). This means that TVUS can rule out a adnexal mass to be free from malignancy but cannot differentiate benign and malignant with utmost certainty based on radiological evidence. These results are in line with the international guidelines for the diagnosis of adnexal masses, which suggests TVUS as a first-line imaging modality to rule out malignancies such as ovarian cancer (6). However, it is not a replacement for laparoscopic surgery and biopsy, which is still the gold standard for the differentiation of adnexal masses.

In conclusion, the present study indicated that TVUS may be a useful imaging modality for differentiating between benign and malignant tumour among patients with adnexal mass with high sensitivity and specificity. TVUS may be employed as an efficient and rapid screening tool for suspected adnexal masses to rule out malignancy.

Acknowledgements

Not applicable.

Funding

No funding was received.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

XZ designed the study. XZ, XM, TD and HS were involved in literature search and data interpretation. XM and TD were responsible for the data analysis. XZ prepared the manuscript. HS edited the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Griffin N, Grant LA and Sala E: Adnexal masses: Characterization and imaging strategies. Semin Ultrasound CT MR 31: 330-346, 2010.
2. Givens V, Mitchell GE, Harraway-Smith C, Reddy A and Maness DL: Diagnosis and management of adnexal masses. Am Fam Physician 80: 815-820, 2009.
3. Hakoun AM, AbouAl-Shaar I, Zaza KJ, Abou-Al-Shaar H and A Salloum MN: Adnexal masses in pregnancy: An updated review. Avicenna J Med 7: 153-157, 2017.
4. Fischerova D, Zikan M, Dundr P and Cibula D: Diagnosis, treatment, and follow-up of borderline ovarian tumors. Oncologist 17: 1515-1533, 2012.
5. Curtin JP: Management of the adnexal mass. Gynecol Oncol 55: S42-S46, 1994.
6. American College of Obstetricians and Gynecologists: ACOG practice bulletin. Management of adnexal masses. Obstet Gynecol 110: 201-214, 2007.
7. Coleman BG: Transvaginal sonography of adnexal masses. Radiol Clin North Am 30: 677-691, 1992.
8. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA and Bossuyt PM; QUADAS-2 Group: QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 155: 529-536, 2011.
9. Alcázar JL, Pascual MÁ, Olartecoechea B, Graupera B, Aubá M, Ajossa S, Hereter L, Julve R, Gastón B, Peddes C, *et al*: IOTA simple rules for discriminating between benign and malignant adnexal masses: Prospective external validation. Ultrasound Obstetr Gynecol 42: 467-471, 2013.
10. Daemen A, Valentin L, Fruscio R, Van Holsbeke C, Melis GB, Guerriero S, Czekierdowski A, Jurkovic D, Ombelet W, Rossi A, *et al*: Improving the preoperative classification of adnexal masses as benign or malignant by second-stage tests. Ultrasound Obstet Gynecol 37: 100-106, 2011.
11. Ruiz de Gauna B, Rodriguez D, Olartecoechea B, Aubá M, Jurado M, Gómez Roig MD and Alcázar JL: Diagnostic performance of IOTA simple rules for adnexal masses classification: A comparison between two centers with different ovarian cancer prevalence. Eur J Obstet Gynecol Reprod Biol 191: 10-14, 2015.
12. Fathallah K, Huchon C, Bats AS, Metzger U, Lefrère-Belda MA, Bensaid C and Lécuru F: External validation of simple ultrasound rules of Timmerman on 122 ovarian tumors. Gynecol Obstet Fertil 39: 477-481, 2011 (In French).
13. Granberg S, Norström A and Wikland M: Tumors in the lower pelvis as imaged by vaginal sonography. Gynecol Oncol 37: 224-229, 1990.
14. Guerriero S, Alcazar JL, Ajossa S, Galvan R, Laparte C, García-Manero M, Lopez-Garcia G and Melis GB: Transvaginal color Doppler imaging in the detection of ovarian cancer in a large study population. Int J Gynecol Cancer 20: 781-786, 2010.
15. Hartman CA, Juliato CRT, Sarian LO, Toledo MC, Jales RM, Morais SS, Pitta DD, Marussi EF and Derchain S: Ultrasound criteria and CA 125 as predictive variables of ovarian cancer in women with adnexal tumors. Ultrasound Obstet Gynecol 40: 360-366, 2012.
16. Jain KA, Friedman DL, Pettinger TW, Alagappan R, Jeffrey RB Jr and Sommer FG: Adnexal masses: Comparison of specificity of endovaginal US and pelvic MR imaging. Radiology 186: 697-704, 1993.
17. Jain KA: Prospective evaluation of adnexal masses with endovaginal gray-scale and duplex and color Doppler US: Correlation with pathologic findings. Radiology 191: 63-67, 1994.
18. Knafel A, Nocun A, Banas T, Wiechec M, Jach R, Pietrus M and Pitynski K: IOTA simple ultrasound-based rules; why do we inconclusive results? Int J Gynecol Cancer 2013: 8, 2013.
19. Komatsu T, Konishi I, Mandai M, Togashi K, Kawakami S, Konishi J and Mori T: Adnexal masses: Transvaginal US and gadolinium-enhanced MR imaging assessment of intratumoral structure. Radiology 198: 109-115, 1996.
20. Lucidarme O, Akakpo JP, Granberg S, Sideri M, Levavi H, Schneider A, Autier P, Nir D and Bleiberg H; Ovarian HistoScanning Clinical Study Group: A new computer-aided diagnostic tool for non-invasive characterisation of malignant ovarian masses: Results of a multicentre validation study. Eur Radiol 20: 1822-1830, 2010.

21. Mancuso A, De Vivo A, Triolo O and Irato S: The role of transvaginal ultrasonography and serum CA 125 assay combined with age and hormonal state in the differential diagnosis of pelvic masses. Eur J Gynaecol Oncol 25: 207-210, 2004.
22. Moszynski R, Szpurek D, Szubert S and Sajdak S: Analysis of false negative results of subjective ultrasonography assessment of adnexal masses. Ginekol Pol 84: 102-107, 2013.
23. Nunes N, Yazbek J, Ambler G, Hoo W, Naftalin J and Jurkovic D: Prospective evaluation of the IOTA logistic regression model LR2 for the diagnosis of ovarian cancer. Ultrasound Obstet Gynecol 40: 355-359, 2012.
24. Nunes N, Ambler G, Hoo WL, Naftalin J, Foo X, Widschwendter M and Jurkovic D: A prospective validation of the IOTA logistic regression models (LR1 and LR2) in comparison to subjective pattern recognition for the diagnosis of ovarian cancer. Int J Gynecol Cancer 23: 1583-1589, 2013.
25. Radosa MP, Vorwergk J, Fitzgerald J, Kaehler C, Schneider U, Camara O, Runnebaum IB and Schleußner E: Sonographic discrimination between benign and malignant adnexal masses in premenopause. Ultraschall Med 35: 339-344, 2014.
26. Romagnolo C, Trivella G, Bonacina M, Fornalè M, Maggino T and Ferrazzi E: Preoperative diagnosis of 221 consecutive ovarian masses: Scoring system and expert evaluation. Eur J Gynaecol Oncol 27: 487-489, 2006.
27. Roman LD, Muderspach LI, Stein SM, Laifer-Narin S, Groshen S and Morrow CP: Pelvic examination, tumor marker level, and gray-scale and Doppler sonography in the prediction of pelvic cancer. Obstet Gynecol 89: 493-500, 1997.
28. Salle B, Gaucherand P, Ecochard R and Rudigoz RC: Role of pulsed color Doppler in the presurgical evaluation of pelvic masses. J Gynecol Obstet Biol Reprod (Paris) 24: 234-240, 1995 (In French).
29. Sayasneh A, Wynants L, Preisler J, Kaijser J, Johnson S, Stalder C, Husicka R, Abdallah Y, Raslan F, Drought A, *et al*: Multicentre external validation of IOTA prediction models and RMI by operators with varied training. Br J Cancer 108: 2448-2454, 2013.
30. Sayasneh A, Kaijser J, Preisler J, Smith AA, Raslan F, Johnson S, Husicka R, Ferrara L, Stalder C, Ghaem-Maghami S, *et al*: Accuracy of ultrasonography performed by examiners with varied training and experience in predicting specific pathology of adnexal masses. Ultrasound Obstet Gynecol 45: 605-612, 2015.
31. Shetty J, Reddy G and Pandey D: Role of sonographic gray-scale pattern recognition in the diagnosis of adnexal masses. J Clin Diagn Res 11: QC12-QC15, 2017.
32. Shetty J, Saradha A, Pandey D, Bhat R, Pratap K and Bharatnur S: IOTA simple ultrasound rules for triage of adnexal mass: Experience from South India. J Obstet Gynecol India 69: 356-362, 2019.
33. Silvestre L, Martins WP and Candido-dos-Reis FJ: Limitations of three-dimensional power Doppler angiography in preoperative evaluation of ovarian tumors. J Ovarian Res 8: 47, 2015.
34. Sohaib SA, Mills TD, Sahdev A, Webb JA, Vantrappen PO, Jacobs IJ and Reznek RH: The role of magnetic resonance imaging and ultrasound in patients with adnexal masses. Clin Radiol 60: 340-348, 2005.
35. Sokalska A, Timmerman D, Testa AC, Van Holsbeke C, Lissoni AA, Leone FP, Jurkovic D and Valentin L: Diagnostic accuracy of transvaginal ultrasound examination for assigning a specific diagnosis to adnexal masses. Ultrasound Obstet Gynecol 34: 462-470, 2009.
36. Stein SM, Laifer-Narin S, Johnson MB, Roman LD, Muderspach LI, Tyszka JM and Ralls PW: Differentiation of benign and malignant adnexal masses: Relative value of gray-scale, color Doppler, and spectral Doppler sonography. AJR Am J Roentgenol 164: 381-386, 1995.
37. Strigini FA, Gadducci A, Del Bravo B, Ferdeghini M and Genazzani AR: Differential diagnosis of adnexal masses with transvaginal sonography, color flow imaging, and serum CA 125 assay in pre- and postmenopausal women. Gynecol Oncol 61: 68-72, 1996.
38. Tantipalakorn C, Wanapirak C, Khunamornpong S, Sukpan K and Tongsong T: IOTA simple rules in differentiating between benign and malignant ovarian tumors. Asian Pac J Cancer Prev 15: 5123-5126, 2014.
39. Testa A, Kaijser J, Wynants L, Fischerova D, Van Holsbeke C, Franchi D, Savelli L, Epstein E, Czekierdowski A, Guerriero S, *et al*: Strategies to diagnose ovarian cancer: New evidence from phase 3 of the multicentre international IOTA study. Br J Cancer 111: 680-688, 2014.
40. Timmerman D, Verrelst H, Bourne TH, De Moor B, Collins WP, Vergote I and Vandewalle J: Artificial neural network models for the preoperative discrimination between malignant and benign adnexal masses. Ultrasound Obstet Gynecol 13: 17-25, 1999.
41. Timmerman D, Ameye L, Fischerova D, Epstein E, Melis GB, Guerriero S, Van Holsbeke C, Savelli L, Fruscio R, Lissoni AA, *et al*: Simple ultrasound rules to distinguish between benign and malignant adnexal masses before surgery: Prospective validation by IOTA group. BMJ 341: c6839, 2010.
42. Utrilla-Layna J, Alcázar JL, Aubá M, Laparte C, Olartecoechea B, Errasti T, Juez L, Mínguez JÁ, Guerriero S and Jurado M: Performance of three-dimensional power Doppler angiography as third-step assessment in differential diagnosis of adnexal masses. Ultrasound Obstet Gynecol 45: 613-617, 2015.
43. Valentin L: Prospective cross-validation of Doppler ultrasound examination and gray-scale ultrasound imaging for discrimination of benign and malignant pelvic masses. Ultrasound Obstet Gynecol 14: 273-283, 1999.
44. Valentin L, Jurkovic D, Van Calster B, Testa A, Van Holsbeke C, Bourne T, Vergote I, Van Huffel S and Timmerman D: Adding a single CA 125 measurement to ultrasound imaging performed by an experienced examiner does not improve preoperative discrimination between benign and malignant adnexal masses. Ultrasound Obstet Gynecol 34: 345-354, 2009.
45. Van Calster B, Timmerman D, Bourne T, Testa AC, Van Holsbeke C, Domali E, Jurkovic D, Neven P, Van Huffel S and Valentin L: Discrimination between benign and malignant adnexal masses by specialist ultrasound examination versus serum CA-125. J Natl Cancer Inst 99: 1706-1714, 2007.
46. Van Gorp T, Veldman J, Van Calster B, Cadron I, Leunen K, Amant F, Timmerman D and Vergote I: Subjective assessment by ultrasound is superior to the risk of malignancy index (RMI) or the risk of ovarian malignancy algorithm (ROMA) in discriminating benign from malignant adnexal masses. Eur J Cancer 48: 1649-1656, 2012.
47. Van Holsbeke C, Van Calster B, Testa AC, Testa AC, Domali E, Lu C, Van Huffel S, Valentin L and Timmerman D: Prospective internal validation of mathematical models to predict malignancy in adnexal masses: Results from the international ovarian tumor analysis study. Clin Cancer Res 15: 684-691, 2009.
48. van Trappen PO, Rufford BD, Mills TD, Sohaib SA, Webb JA, Sahdev A, Carroll MJ, Britton KE, Reznek RH and Jacobs IJ: Differential diagnosis of adnexal masses: Risk of malignancy index, ultrasonography, magnetic resonance imaging, and radioimmunoscintigraphy. Int J Gynecol Cancer 17: 61-67, 2007.
49. Yamashita Y, Torashima M, Hatanaka Y, Harada M, Higashida Y, Takahashi M, Mizutani H, Tashiro H, Iwamasa J, Miyazaki K, *et al*: Adnexal masses: Accuracy of characterization with transvaginal US and precontrast and postcontrast MR imaging. Radiology 194: 557-565, 1995.
50. Karnik A, Tembey RA and Mani S: Value of MRI in characterizing adnexal masses. J Obstet Gynaecol India 65: 259-266, 2015.
51. Kaijser J, Vandecaveye V, Deroose CM, Rockall A, Thomassin-Naggara I, Bourne T and Timmerman D: Imaging techniques for the pre-surgical diagnosis of adnexal tumours. Best Pract Res Clin Obstet Gynaecol 28: 683-695, 2014.
52. Dochez V, Caillon H, Vaucel E, Dimet J, Winer N and Ducarme G: Biomarkers and algorithms for diagnosis of ovarian cancer: CA125, HE4, RMI and ROMA, a review. J Ovarian Res 12: 28, 2019.
53. Ferraro S, Braga F, Lanzoni M, Boracchi P, Biganzoli EM and Panteghini M: Serum human epididymis protein 4 vs carbohydrate antigen 125 for ovarian cancer diagnosis: A systematic review. J Clin Pathol 66: 273-281, 2013.
54. Wang J, Gao J, Yao H, Wu Z, Wang M and Qi J: Diagnostic accuracy of serum HE4, CA125 and ROMA in patients with ovarian cancer: A meta-analysis. Tumour Biol 35: 6127-6138, 2014.
55. Zhen S, Bian LH, Chang LL and Gao X: Comparison of serum human epididymis protein 4 and carbohydrate antigen 125 as markers in ovarian cancer: A meta-analysis. Mol Clin Oncol 2: 559-566, 2014.
56. Li F, Tie R, Chang K, Wang F, Deng S, Lu W, Yu L and Chen M: Does risk for ovarian malignancy algorithm excel human epididymis protein 4 and CA125 in predicting epithelial ovarian cancer: A meta-analysis. BMC Cancer 12: 258, 2012.
57. Pewsner D, Battaglia M, Minder C, Marx A, Bucher HC and Egger M: Ruling a diagnosis in or out with 'SpPIn' and 'SnNOut': A note of caution. BMJ 329: 209-213, 2004.