

Comparison of physical examination, ultrasound techniques and magnetic resonance imaging for the diagnosis of deep infiltrating endometriosis: A systematic review and meta-analysis of diagnostic accuracy studies

XIAO ZHANG, TAO HE and WEN SHEN

Department of Ultrasound, Hanchuan People's Hospital, Hanchuan, Hubei 431600, P.R. China

Received December 10, 2019; Accepted May 15, 2020

DOI: 10.3892/etm.2020.9043

Abstract. Visual inspection via laparoscopy is considered the gold standard for the diagnosis of deep infiltrating endometriosis. Laparoscopy is an invasive procedure; therefore, it would be beneficial to patients if accurate non-invasive modalities were available for the diagnosis of deep infiltrating endometriosis. The purpose of the current review and meta-analysis was to assess the diagnostic accuracy of clinical examination, transvaginal ultrasound (TVUS), transrectal ultrasound (TRUS) and MRI as alternative methods for diagnosis of deep infiltrating endometriosis. A systematic search of the Medline, Scopus, Embase and Cochrane library databases, between their inception and September 2019, was performed. The quality of trials was assessed using the quality assessment of diagnostic accuracy studies-2 tool. Meta-analyses were conducted to obtain the pooled sensitivity, specificity, positive and negative likelihood ratios and diagnostic odds ratio for each of the three imaging modalities and clinical examination. A total of 30 studies with 4,565 participants were included in the review. Physical examination had a pooled sensitivity of 71% and a specificity of 69%, with an average diagnostic accuracy [area under the curve (AUC) =0.76]. TVUS had a pooled sensitivity of 76% and a specificity of 94%, with higher diagnostic accuracy than physical examination (AUC =0.92). TRUS had a pooled sensitivity of 91% and a specificity of 80% with an AUC of 0.93. MRI had a pooled sensitivity of 82% and a specificity of 87% with higher diagnostic accuracy than physical examination (AUC =0.91). All the imaging modalities had good clinical utility, as indicated by the Fagan plot. The present analysis demonstrates that the imaging modalities TVUS, TRUS and MRI may be highly useful alternatives to

laparoscopy for diagnosis of deep infiltrating endometriosis and that these techniques have a high sensitivity and specificity.

Introduction

Endometriosis is one of the most common benign conditions affecting women. The prevalence of the disease is ~10% in females of reproductive age (15-49 years) and 20-50% in those diagnosed with infertility (1,2). Dysmenorrhea, chronic pelvic pain and infertility are commonly associated with the disorder (3). Endometriosis can be classified as either superficial or deep. Peritoneal infiltration of <5 mm is defined as superficial endometriosis. The presence of endometrial tissue, fibrosis and hyperplasia with >5 mm peritoneal infiltration is defined as deep infiltrating endometriosis (DIE) (3). DIE accounts for 15-30% of all diagnosed endometriosis cases (4,5). While there is no known cure for the disease, available treatment options include expectant management, hormonal therapy and surgical management (6). Treatment is usually personalized, considering the therapeutic goal (subfertility or pain relief) and the site of disease (6).

Pharmacological therapies for endometriosis include progestogens, the combined oral contraceptive pill, gonadotropin-releasing hormone agonists and weak androgens (7). These drugs alleviate the symptoms of chronic pelvic pain and dysmenorrhoea; however, they are associated with adverse events, including irritability, breast discomfort, bone loss and androgenic symptoms (7). Surgical treatment of endometriosis can alleviate pain and improve fertility. However, 5-year recurrence rates following surgery are high (40-50%) (8,9). The potential success and impact of these strategies is dependent on early and accurate diagnosis (10). Laparoscopic observation and biopsy are considered to be the gold standard for diagnosis of DIE (9). However, laparoscopy is invasive and does not allow any time for preoperative planning. Several non-invasive imaging modalities, such as physical examination, transvaginal ultrasonography (TVUS), transrectal sonography (TRUS) and MRI, are also available for the diagnosis of DIE (11). Accuracy of such non-invasive modalities would ensure their use as standard, significantly reducing the risk associated with surgery while facilitating early diagnosis and treatment (11). A number of studies

Correspondence to: Dr Tao He, Department of Ultrasound, Hanchuan People's Hospital, 1 People's Avenue, Hanchuan, Hubei 431600, P.R. China
E-mail: 17771215810@163.com

Key words: deep infiltrating endometriosis, MRI, ultrasonography, validation

have attempted to assess the accuracy of such non-invasive diagnostic modalities for DIE in various settings with different results. However, to the best of our knowledge, there have been no systematic efforts to collate evidence comparing these different diagnostic techniques. Therefore, the purpose of the present study was to systematically search the literature and perform a meta-analysis of diagnostic data to compare the accuracy of physical examination, ultrasound techniques (TVUS & TRUS) and MRI in the diagnosis of DIE.

Materials and methods

Inclusion criteria

Studies. All studies examining the diagnostic accuracy of physical examination, ultrasound techniques (TVUS or TRUS) and MRI for diagnosis of DIE were included where they reported sensitivity and specificity values of any of the aforementioned diagnostic techniques or provided the data required to calculate these rates. Studies published as full-text manuscripts were included while abstracts and case reports were excluded.

Participants. Studies conducted on patients with suspected DIE were included irrespective of the medical co-morbidities suffered by the participants and the setting in which the study was conducted.

Index test. Studies utilizing physical examination, TVUS, TRUS or MRI to diagnose DIE were included.

Reference standards. Studies were included only if the diagnostic accuracy of physical examination, TVUS, TRUS or MRI was compared with standard laparoscopic or histopathological examination. Diagnosis using the reference standard must have been made by specialist doctors or trained researchers.

Search strategy. An extensive electronic search of the Medline (PubMed) (<https://pubmed.ncbi.nlm.nih.gov/>), Scopus (<https://www.scopus.com/>), Embase (<https://www.embase.com/>), and Cochrane library (<https://www.cochranelibrary.com/>) was conducted. A combination of medical subject heading (MeSH) and free-text terms was used to conduct the literature search. The following MeSH terms and free-text terms were used for the literature search: 'Validation studies', 'deep infiltrating endometriosis', 'physical examination', 'transvaginal ultrasonography', 'transrectal ultrasonography', 'magnetic resonance imaging', 'gynaecological disorders', 'sensitivity', 'specificity', 'diagnosis', 'endometriosis' and 'diagnostic accuracy studies'. These search terms were used in various combinations using Boolean operators such as 'AND', 'OR' or 'NOT'. Additional filters related to timeline of search in the database was between their inception and September 2019. The language of the studies was restricted to English.

Selection of studies. The literature search was performed by two investigators independently. Records were screened by their title, abstract and keywords for possible inclusion in the review. Full texts of relevant studies were extracted and screened further, based on the eligibility criteria for final inclusion in the review. The reference list of all full-text articles was searched manually to identify any missed studies. Any disagreements between the two authors during the entire

selection process were resolved either through consensus or consultation with an independent third investigator.

Data extraction and management. The primary investigator extracted the relevant study characteristics for the review from all included studies. The following data were extracted: Author, year of publication, study design, study setting, index test, reference standards, comorbidities, number of participants, mean age, inclusion and exclusion criteria, true positives, true negatives, false positives and false negatives. Data was transferred to STATA software version 14.2 (StataCorp LLC) by the primary investigator. Data entry was double-checked for accuracy by the third investigator.

Risk of bias assessment. The risk of bias in all of the included studies was assessed by two investigators independently using the quality assessment of diagnostic accuracy studies-2 (QUADAS-2) tool (12). The following domains were used for the assessment of the risk of bias: Patient selection, index test, reference standard and flow and timing of assessments. Each domain was graded as high, low or unclear risk of bias. Any discrepancies were resolved by consensus or consultation with the third investigator.

Statistical analysis. Meta-analysis was performed using STATA software version 14.2 (StataCorp LLC) to obtain the pooled estimate and 95% confidence intervals (CI) of sensitivity, specificity, positive likelihood ratio (LRP), negative likelihood ratio (LRN) and summary diagnostic odds ratio (DOR) for each of the diagnostic tests. Summary receiver operator characteristic (sROC) curves were constructed and the area under the curve (AUC) was obtained. Each data point in the sROC curve represents a separate study. Test performance accuracy was classified as follows: AUC 0.5±0.7, low; AUC 0.7±0.9, medium; and AUC >0.9, high (13).

Forest plots were used to graphically represent the study-specific and pooled estimates. A Fagan plot was constructed to estimate how much the result of a diagnostic test changes the probability that a patient has DIE. Between-study heterogeneity was assessed using bivariate box plots, χ^2 tests and I^2 statistics. I^2 values of <25% indicated mild heterogeneity, 25-75% indicated moderate heterogeneity and >75% indicated substantial heterogeneity (14). Publication bias was assessed graphically via funnel plots and the asymmetry of the plot was evaluated using Deek's test.

Results

Selection of studies. In total, 1,210 citations were found, of which 445 were retrieved from Medline, 411 from Scopus, 302 from Embase and 52 from the Cochrane library. After the first stage of screening (title, abstract and keywords), 98 relevant studies were retrieved. The full text of these studies was reviewed based on the eligibility criteria. Bibliographies of the retrieved articles were reviewed, and 7 additional studies were identified. In the final analysis 30 studies with 4,565 participants satisfying the inclusion criteria were included (Fig. 1) (15-44).

Characteristics of included studies. Characteristics of the studies are described in Table I. With the exception of

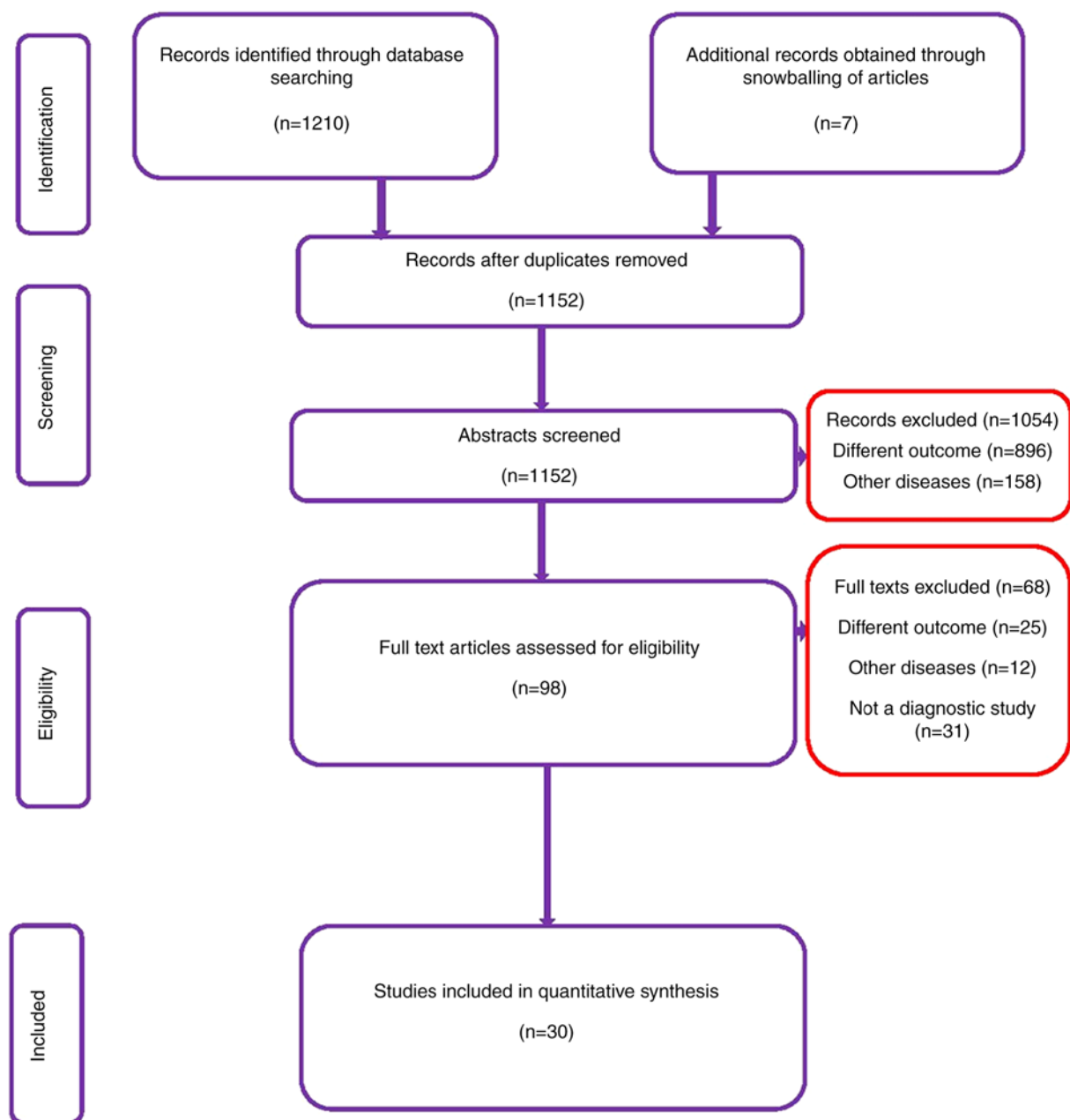


Figure 1. Flow chart of search strategy.

Chapron *et al* (19), all included studies were prospective studies. All the studies were conducted in high-income countries as per World Bank income group country classification (45). The majority of the studies (50%) were conducted in Italy followed by the United Kingdom, France, the United States of America and Austria. The mean age of the participants ranged from 28-36 years. In total, 4,565 participants were assessed in the included studies, with 469 participants undergoing physical examination, 3,847 undergoing TVUS, 728 undergoing TRUS and 1,298 undergoing MRI. The total sample size of studies varied from 31-1,440 participants. Among the 30 studies included, 5 studies reported on the diagnostic accuracy of physical examination, 21 on TVUS, 6 on TRUS and 13 on MRI.

Methodological quality of the included studies. The risk of bias assessment of the included studies is presented in Table II

and Fig. S1. A total of 13 studies had a high risk of bias for patient selection, while 5 studies had a high risk of bias for the conduct and interpretation of the index test domain. A total of 12 studies had a high risk of bias due to the conduct and interpretation of reference standards and 12 studies had a high risk of bias for the flow of patients and the time interval between index test and reference standard.

Diagnostic performance of physical examination. In total, five studies assessed the diagnostic accuracy of physical examination for DIE. The pooled sensitivity and specificity of physical examination for the diagnosis of DIE were 71% (95% CI, 60-80%) and 69% (95% CI, 54-82%), respectively. The DOR was 5 (95% CI, 3-12) as indicated by the Forest plot (Fig. 2A). LRP was 2.3 (95% CI, 1.5-3.6) and LRN was 0.42 (95% CI, 0.29-0.61). Fig. 3A depicts the sROC curve for

Table I. Characteristics of the included studies (n=30).

Author, year	Country	Study design	Population size	Type of diagnostic modality	Gold standard comparator	Mean patient age, years	Time interval between index test and standards reference	(Refs.)
Abrao <i>et al.</i> , 2007	Brazil	Prospective cross-sectional study	104	Clinical examination, TVUS, MRI	Laparoscopy + histopathology	33.8	3 months	(15)
Alborzi <i>et al.</i> , 2018	Iran	Prospective longitudinal study	317	TVUS, TRUS, MRI	Laparoscopy + histopathology	31.0	Not specified	(16)
Bazot <i>et al.</i> , 2009	France	Longitudinal	92	Clinical examination, TVUS, TRUS, MRI	Laparoscopy, laparotomy + histopathology	31.8	<12 months	(17)
Bergamini <i>et al.</i> , 2010	Italy	Prospective	61	TRUS	Laparoscopy, laparotomy + histopathology	33.1	Not specified	(18)
Chapron <i>et al.</i> , 2004	France	Retrospective	81	TRUS, MRI	Laparoscopy or laparotomy	31.9	Not specified	(19)
Charmie <i>et al.</i> , 2009	Brazil	Prospective	92	MRI	Laparoscopy + histopathology	33.0	Not specified	(20)
Dessole <i>et al.</i> , 2003	Italy	Prospective	46	TVUS	Laparoscopy, laparotomy + histopathology	30.3	Not specified	(21)
Eskenazi <i>et al.</i> , 2001	Italy	Prospective (study sample)	90	Clinical examination, TVUS	Laparoscopy, laparotomy + histopathology	35.7	34 days	(22)
Fedele <i>et al.</i> , 1998	Italy	Retrospective (test sample)	140	TRUS	Laparoscopy, laparotomy + histopathology	30.2	1 week	(23)
Grasso <i>et al.</i> , 2009	Italy	Prospective	33	TVUS, MRI	Laparoscopy + histopathology	35.0	1-4 weeks	(24)
Guerriero <i>et al.</i> , 2007	Italy	Prospective	50	TVUS	Laparoscopy + histopathology	33.0	Within 7 days	(25)
Guerriero <i>et al.</i> , 2008	Italy	Prospective	88	TVUS	Laparoscopy + histopathology	33.0	Within 7 days	(26)
Guerriero <i>et al.</i> , 2014	Italy	Prospective	202	TVUS	Laparoscopy + histopathology	34.0	Within 1 month	(27)
Guerriero <i>et al.</i> , 2017	Italy	Prospective	159	TVUS, MRI	Laparoscopy + histopathology	33.0	Surgery was performed over 5 years of analysis	(28)
Ha <i>et al.</i> , 1994	Korea	Prospective	186	MRI	Laparoscopy + histopathology	35.0	Within 2 weeks	(29)
Holland <i>et al.</i> , 2010	London, United Kingdom	Prospective	201	TVUS	Laparoscopy + histopathology	34.9	Mean duration =37.5 days	(30)
Holland <i>et al.</i> , 2013	London, United Kingdom	Prospective	198	TVUS	Laparoscopy + histopathology	35.0	Mean duration =36.8 days	(31)
Hudelist <i>et al.</i> , 2011	United Kingdom	Prospective	129	Clinical examination, TVUS	Laparoscopy + histopathology	32.2	Within 3 months	(32)

Table I. Continued.

Author, year	Country	Study design	Population size	Type of diagnostic modality	Gold standard comparator	Mean patient age, years	Time interval between index test and standards reference	(Refs.)
Hudelist <i>et al</i> , 2013	Austria	Prospective	117	TVUS	Laparoscopy + histopathology	31.6	Within 2 months	(33)
Manganaro <i>et al</i> , 2013	Italy	Prospective	42	MRI	Laparoscopy + histopathology	28.0	Within 3 months	(34)
Piessens <i>et al</i> , 2014	Australia	Prospective	85	TVUS	Laparoscopy + histopathology	Age range given (18-48)	Within 12 months	(35)
Ribeiro <i>et al</i> , 2008	United Kingdom	Prospective	37	TRUS	Laparoscopy + histopathology	35.8	1-3 months	(36)
Saba <i>et al</i> , 2011	Italy	Prospective	59	TVUS, MRI	Laparoscopy, laparotomy + histopathology	Not given	Not specified	(37)
Saccardi <i>et al</i> , 2012	Italy	Prospective	54	Clinical examination, TVUS, MRI	Laparoscopy + histopathology	32.3	Not specified	(38)
Savelli <i>et al</i> , 2011	Italy	Prospective	69	TVUS	Laparoscopy + histopathology	33.6	Within 1 month	(39)
Stratton <i>et al</i> , 2003	USA	Prospective	48	MRI	Laparoscopy + histopathology	Age range (20-44)	Within 1 month	(40)
Takeuchi <i>et al</i> , 2005	Japan	Prospective	31	MRI	Laparoscopy + histopathology	32.1	Not specified	(41)
Ubaldi <i>et al</i> , 1998	Belgium	Prospective	133	TVUS	Laparoscopy + histopathology	Age range (21-41)	1 day	(42)
Valzeno Menada <i>et al</i> , 2008	Italy	Prospective	181	TVUS	Laparoscopy + histopathology	32.0	Not specified	(43)
Vimercati <i>et al</i> , 2012	Italy	Prospective	1440	TVUS	Laparoscopy + histopathology	34.0	4-8 weeks	(44)

TVUS, transvaginal ultrasound; TRUS, transrectal ultrasound.

Table II. Risk of bias assessment for the included studies (n=30).

Author, year	Selection bias	Index test standards	Reference test standards	Flow and timing of Index and Reference standard	(Refs.)
Abrao <i>et al</i> , 2007	Low	Low	Low	Low	(15)
Alborzi <i>et al</i> , 2018	Low	Low	High	Low	(16)
Bazot <i>et al</i> , 2009	Low	Low	Low	Low	(17)
Bergamini <i>et al</i> , 2010	Low	High	High	High	(18)
Chapron <i>et al</i> , 2004	High	High	High	High	(19)
Charmie <i>et al</i> , 2009	High	Low	High	High	(20)
Dessole <i>et al</i> , 2003	Low	High	High	High	(21)
Eskenazi <i>et al</i> , 2001	High	Low	Low	Low	(22)
Fedele <i>et al</i> , 1998	Low	Low	Low	High	(23)
Grasso <i>et al</i> , 2009	High	Low	Low	Low	(24)
Guerrero <i>et al</i> , 2007	Low	Low	Low	Low	(25)
Guerrero <i>et al</i> , 2008	Low	Low	Low	Low	(26)
Guerrero <i>et al</i> , 2014	Low	Low	Low	Low	(27)
Guerrero <i>et al</i> , 2017	Low	Low	Low	Low	(28)
Ha <i>et al</i> , 1994	High	Low	High	Low	(29)
Holland <i>et al</i> , 2010	Low	Low	Low	Low	(30)
Holland <i>et al</i> , 2013	Low	Low	Low	Low	(31)
Hudelist <i>et al</i> , 2011	Low	Low	Low	Low	(32)
Hudelist <i>et al</i> , 2013	Low	Low	Low	Low	(33)
Manganaro <i>et al</i> , 2013	High	Low	Low	Low	(34)
Piessens <i>et al</i> , 2014	Low	Low	High	Low	(35)
Ribeiro <i>et al</i> , 2008	Low	Low	High	Low	(36)
Saba <i>et al</i> , 2011	Low	Low	Low	High	(37)
Saccardi <i>et al</i> , 2012	Low	Low	Low	High	(38)
Savelli <i>et al</i> , 2011	Low	High	Low	Low	(39)
Stratton <i>et al</i> , 2003	High	Low	Low	Low	(40)
Takeuchi <i>et al</i> , 2005	High	Low	Low	Low	(41)
Ubaldi <i>et al</i> , 1998	Low	Low	Low	Low	(42)
Valenzano Menada <i>et al</i> , 2008	High	Low	Low	High	(43)
Vimercati <i>et al</i> , 2012	Low	Low	Low	Low	(44)

physical examination. The AUC was 0.76 (95% CI, 0.66-0.83), which indicates that physical examination has an intermediate level of diagnostic value.

The clinical utility of physical examination for DIE was average and Fagan's nomogram (Fig. 4A) indicated that post-test probability (positive, 37%; negative, 10%) differed significantly from pre-test probability (20%). There was substantial heterogeneity, with an I^2 value of 75%, and χ^2 test for heterogeneity was statistically significant ($P=0.009$). A bivariate box plot (Fig. 5A) indicated that there was a study out of the circle, illustrating the heterogeneity between the included studies. As there were <10 studies, publication bias was not tested.

Diagnostic performance of TVUS. In total, 21 studies assessed the diagnostic accuracy of TVUS. The pooled sensitivity and

specificity of TVUS for diagnosis of DIE were 76% (95% CI, 67-83%) and 94% (95% CI, 88-97%), respectively. The DOR was 47 (95% CI, 21-104) as displayed in the Forest plot (Fig. 2B). LRP was 12.0 (95% CI, 6.3-22.7) and LRN was 0.25 (95% CI, 0.18-0.36). Fig. 3B depicts the sROC curve for TVUS. The AUC was 0.92 (95% CI, 0.86-0.96), which indicates the high diagnostic value of TVUS.

The clinical utility of TVUS for DIE was good and Fagan's nomogram (Fig. 4B) revealed that post-test probability (positive, 75%; negative, 6%) differed significantly from pre-test probability (20%). There was substantial heterogeneity, with an I^2 value of 99%, and χ^2 test for heterogeneity was statistically significant ($P<0.001$). A bivariate box plot (Fig. 5B) indicated 4 studies out of the circle illustrating heterogeneity between the included studies. The

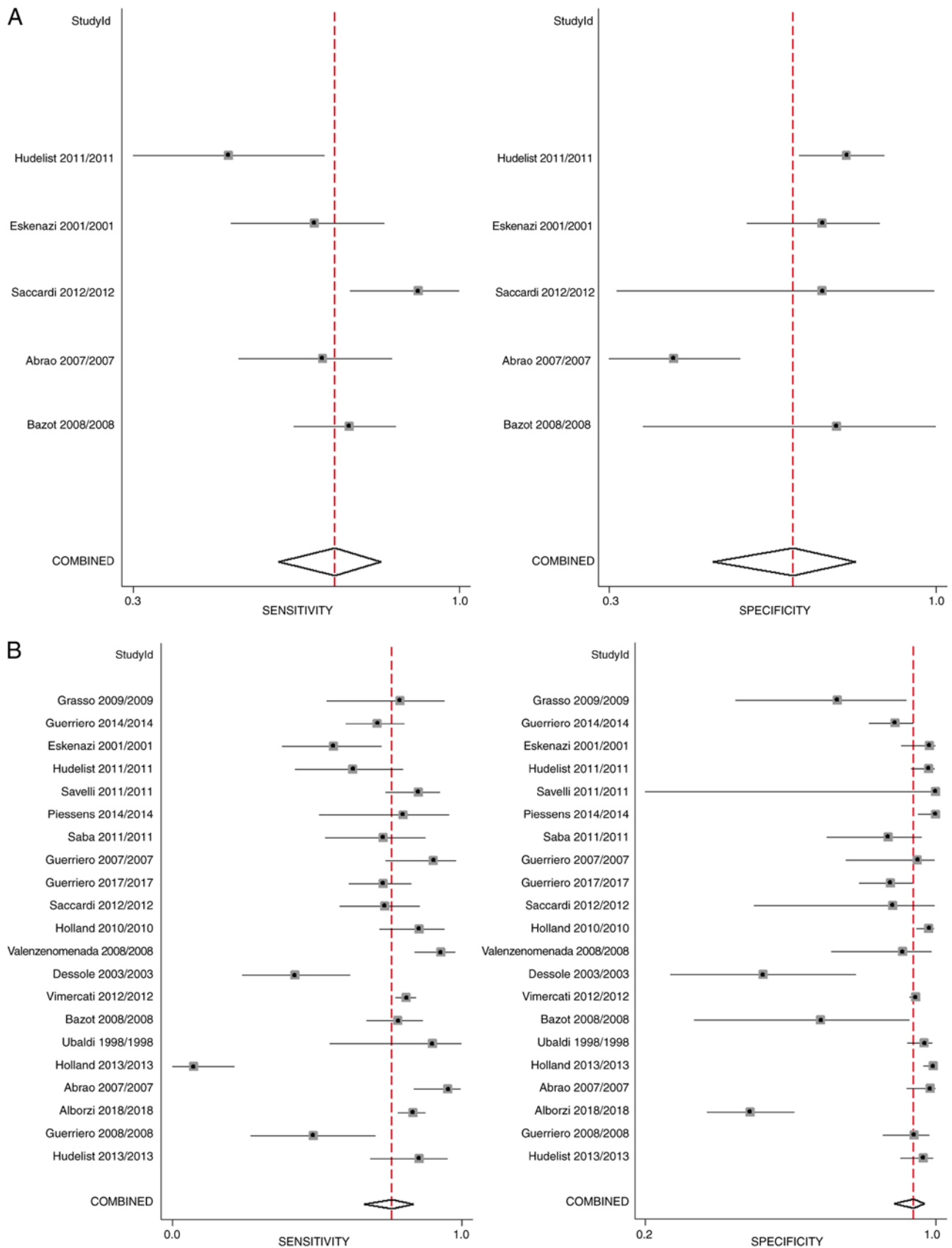


Figure 2. Forest plot exhibiting pooled sensitivity and specificity for non-invasive modalities in diagnosing deep infiltrating endometriosis. (A) Physical examination. (B) Transvaginal ultrasound.

Funnel plot (Fig. S2) for publication bias was symmetrical with non-significant Deek's test (P=0.60).

Diagnostic performance of TRUS. In total, six studies assessed the diagnostic accuracy of TRUS. The pooled sensitivity and

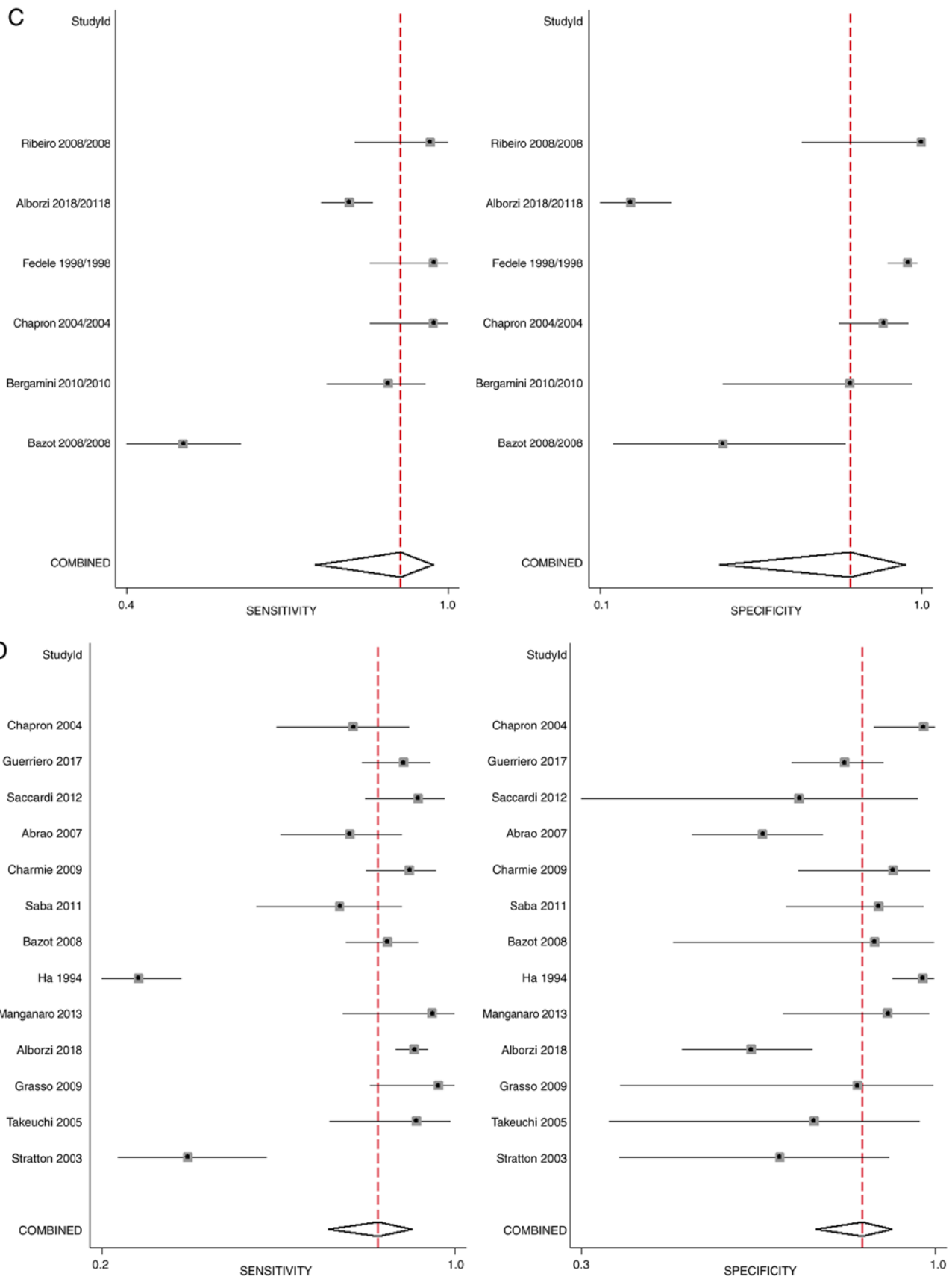


Figure 2. Continued. Forest plot exhibiting pooled sensitivity and specificity for non-invasive modalities in diagnosing deep infiltrating endometriosis. (C) Transrectal ultrasound. (D) Magnetic resonance imaging.

specificity of TRUS for diagnosis of DIE were 91% (95% CI, 74-97%) and 80% (95% CI, 43-95%), respectively. The DOR

was 39 (95% CI, 35-43) as shown in the Forest plot (Fig. 2C). LRP was 4.5 (95% CI, 1.1-18.4) and LRN was 0.12 (95%

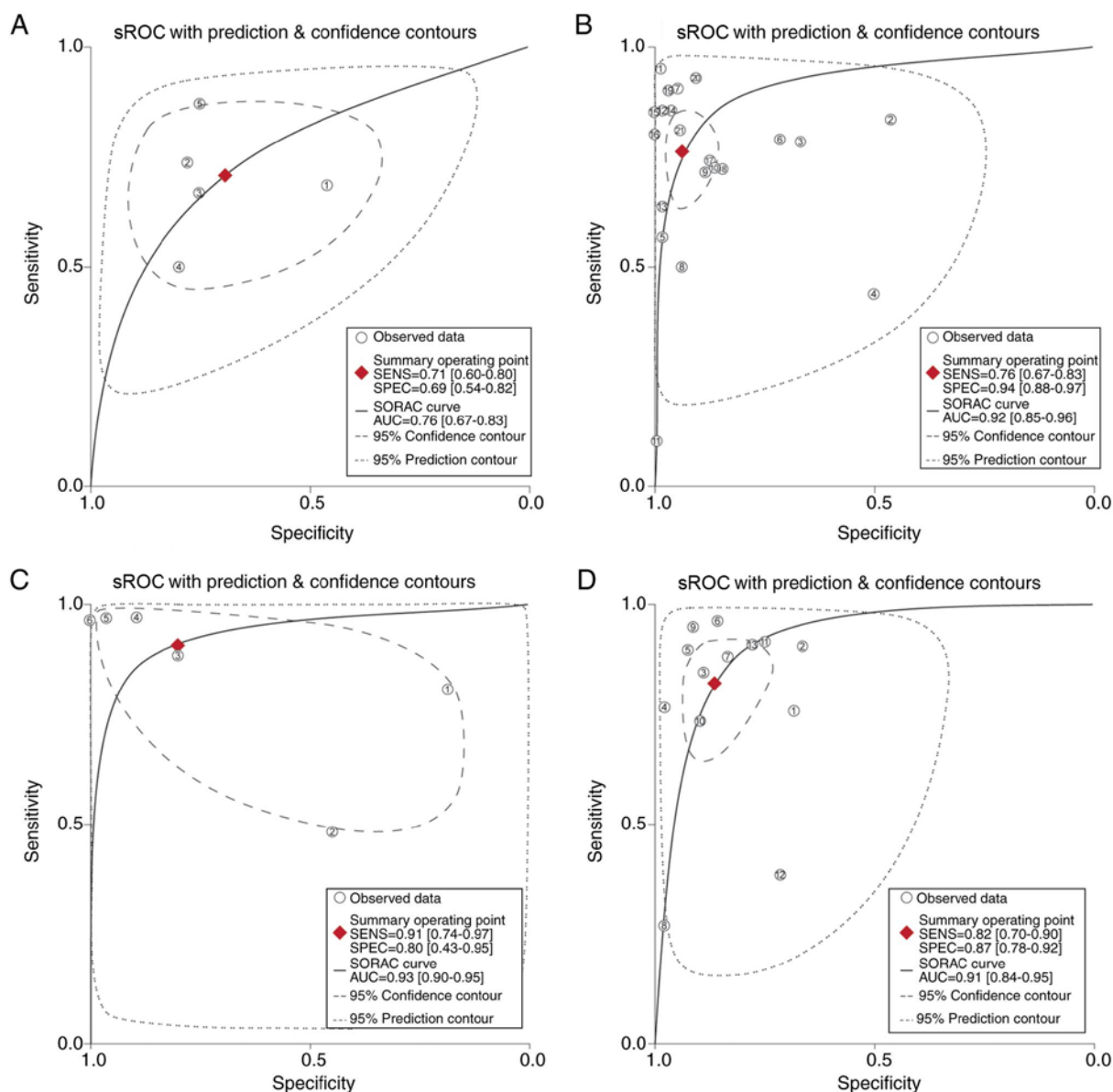


Figure 3. sROC curves for non-invasive modalities in diagnosing deep infiltrating endometriosis. (A) Physical examination. (B) Transvaginal ultrasound. (C) Transrectal ultrasound. (D) Magnetic resonance imaging. sROC, summary receiver operator characteristic.

CI, 0.03-0.45). Fig. 3C depicts the sROC curve for TRUS. The AUC was found to be 0.93 (95% CI, 0.86-0.97), which indicates the high diagnostic value of TRUS.

The clinical utility of TRUS for DIE was good and Fagan's nomogram (Fig. 4C) showed that post-test probability (positive, 53%; negative, 3%) differed significantly from pre-test probability (20%). There was substantial heterogeneity, with an I^2 value of 83%, and the χ^2 test for heterogeneity was statistically significant ($P=0.001$). A bivariate box plot (Fig. 5C) indicated that there was a study out of the circle, illustrating the heterogeneity between the included studies. As there were <10 studies, publication bias was not tested.

Diagnostic performance of MRI. In total, 13 studies assessed the diagnostic accuracy of MRI. The pooled sensitivity and specificity of MRI for diagnosis of DIE were 82% (95% CI, 70-90%) and 87% (95% CI, 78-92%), respectively. The DOR was 30 (95% CI, 14-62) as shown in the Forest plot (Fig. 2D).

LRP was 6.1 (95% CI, 3.7-9.9) and LRN was 0.21 (95% CI, 0.12-0.35). Fig. 3D depicts the sROC curve for MRI. The AUC was found to be 0.91 (95% CI, 0.80-0.97), which indicates a high diagnostic value of MRI.

The clinical utility of MRI for DIE was good and Fagan's nomogram (Fig. 4D) showed that post-test probability (positive, 60%; negative, 5%) differed significantly from pre-test probability (20%). There was substantial heterogeneity, with an I^2 value of 98%, and χ^2 test for heterogeneity was statistically significant ($P<0.001$). A bivariate box plot (Fig. 5D) indicated 2 studies out of the circle, showing the heterogeneity between the included studies. The Funnel plot (Fig. S3) for publication bias was symmetrical with non-significant Deek's test ($P=0.41$).

Discussion

Various imaging modalities are available for diagnosis of DIE. Clinical history and physical examination of the pelvis may

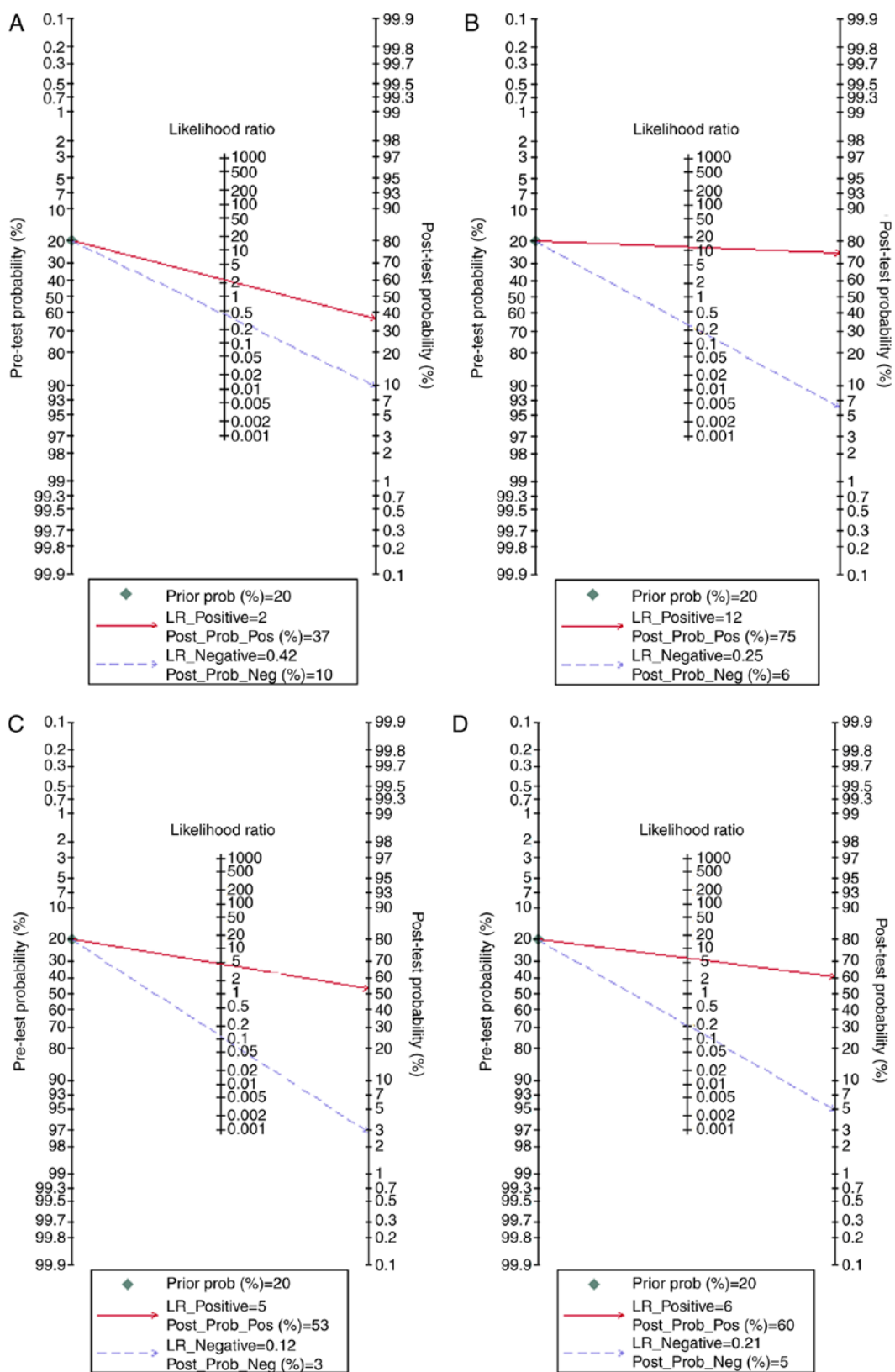


Figure 4. Fagan nomogram evaluating the overall value of non-invasive modalities in diagnosing deep infiltrating endometriosis. (A) Physical examination. (B) Transvaginal ultrasound (C) Transrectal ultrasound. (D) Magnetic resonance imaging.

suggest endometriosis; however, the heterogeneity of clinical presentation, the prevalence of asymptomatic endometriosis (2-50%), and a lack of association between disease severity and presenting symptoms results in significant difficulty

in obtaining an accurate diagnosis based solely on patient history and physical examination (4,5,8,9). Several advanced ultrasonographic techniques have been described to identify endometriosis, such as pelvic organ mobility, sliding

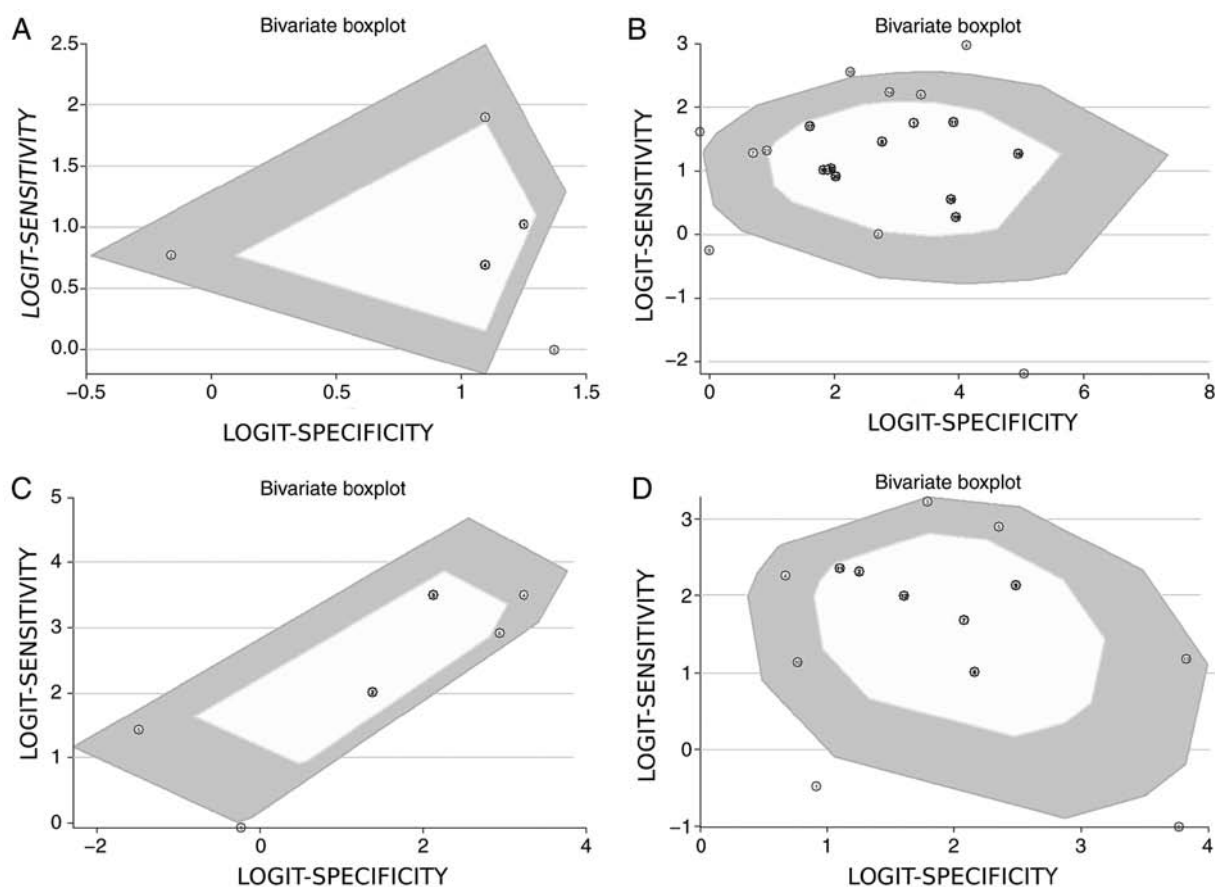


Figure 5. Bivariate box plot of the sensitivity and specificity in the included studies. (A) Physical examination. (B) Transvaginal ultrasound. (C) Transrectal ultrasound. (D) Magnetic resonance imaging.

sign, rectal water contrast, tenderness-guided ultrasound and bowel preparation (25,33,43). These techniques can significantly improve the diagnostic performance of TVUS for endometriosis. Furthermore, the MRI 'jelly method' appears to have improved diagnostic accuracy compared to older MRI modalities (41). Although these modalities cannot replace histopathology or biopsy as the gold standard for diagnosis, a high diagnostic accuracy of such non-invasive methods may reduce the diagnostic delay and complications associated with invasive diagnostic methods. Hence, it is important to precisely determine the diagnostic accuracy of each of these modalities. In the present review of 30 studies with 4,565 participants, the results suggested that while physical examination has intermediate diagnostic accuracy, TRUS, TVUS and MRI have high diagnostic accuracy for DIE.

To the best of our knowledge, diagnostic accuracy of physical examination for DIE has not been assessed in any previous reviews. The present study results indicated the pooled sensitivity and specificity of physical examination to be 71 and 69%, respectively, with average diagnostic accuracy (AUC =0.76). TVUS had a pooled sensitivity and specificity of 76 and 94% with high diagnostic accuracy (AUC =0.92). These findings are in agreement with earlier reviews on the diagnostic accuracy of TVUS, which also indicated the high specificity and high diagnostic accuracy of this technique (AUC >0.90) (11,46-48). However, compared with the previous reviews, a greater number of studies were included in the present analysis.

In the present analysis, TRUS had a pooled sensitivity of 91% and a specificity of 80% with an AUC of 0.93, which is similar to the diagnostic accuracy of TVUS. To the best of our knowledge, no previous reviews have compared the accuracy of TRUS and TVUS for the diagnosis of DIE. The present study revealed MRI to have lower sensitivity (82%) and specificity (87%) for diagnosis of DIE when compared with the previous reviews of Moura *et al* (46) and Nisenblat *et al* (11). Their studies demonstrated ~90% sensitivity and specificity of MRI for diagnosis of DIE. However, compared to these previous reviews, a greater number of studies were included in the present analysis. The results of the present study also suggested that TVUS may be used to exclude the diagnosis of endometriosis with certainty, which was not seen in the previous reviews.

The results of the present study should be interpreted cautiously, considering the inter-study heterogeneity and quality of the included studies. The high heterogeneity limits any strong conclusions on the diagnostic accuracy of the non-invasive modalities for DIE. There was no significant publication bias among studies reporting diagnostic accuracy of TVUS and MRI. However, publication bias for studies reporting TRUS and physical examination could not be assessed due to the limited number of studies included. The overall quality of studies was good, as most of the studies had a low risk of bias with respect to all the four domains measured using the QUADAS tool.

The present study had certain strengths. A comprehensive review including 30 studies with 4,565 participants to evaluate

the accuracy of various imaging modalities and clinical examination in diagnosing DIE was conducted. To the best of our knowledge, this is the first review to provide pooled estimates for four different diagnostic methods of DIE. The lack of any significant publication bias adds credibility to the current results. Nonetheless, the present study has certain limitations. First, some studies included in the review had a high risk of bias in certain domains, which may have influenced individual study results. Secondly, significant heterogeneity was revealed between the studies included in the review, which limited the interpretation of the pooled outcomes.

Despite these limitations, the present study provided valuable insights regarding the diagnostic accuracy of various non-invasive techniques for DIE. Though all the imaging modalities had good sensitivity and specificity, only TVUS was close to fulfilling the SpPin i.e., specific positive in triage test criteria (95% specificity). None of the modalities met the SnNout criteria i.e. sensitive negative for 95% sensitivity of a diagnostic test (11,49). This suggests that TVUS may be used to diagnose DIE, but cannot confirm a woman to be disease-free based on radiological evidence. These findings are in line with the international guidelines for the diagnosis of DIE, which suggests TVUS as a first-line imaging modality following history and physical examination (50,51). However, TVUS does not replace laparoscopic surgery and biopsy, which is still the gold standard for diagnosing DIE.

To summarize, the present study demonstrated that TRUS, TVUS and MRI have good diagnostic accuracy for DIE and have a high sensitivity and specificity. The diagnostic accuracy was most similar for TVUS, TRUS and MRI, with all the modalities having AUC >0.9. Physical examination was also indicated to have reasonable diagnostic accuracy. TVUS was the most specific imaging modality while TRUS was the most sensitive modality for diagnosis of DIE. However, these findings should be interpreted with caution owing to the significant heterogeneity between the included studies. These results suggest that these imaging modalities can be utilized as efficient and time-saving screening tools for DIE.

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. The relevant data is also available in the referenced studies.

Authors' contributions

XZ conceived and designed the study. XZ, TH and WS collected the data, performed the literature search and analyzed the data. XZ was involved in the writing of the manuscript. All authors have read and approved the final manuscript.

Ethical 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

- Irving JA and Clement PB: Diseases of the Peritoneum. In: Blaustein's Pathology of the Female Genital Tract. Kurman RJ, Ellenson LH and Ronnett BM (eds) Springer US, Boston, MA, pp625-678, 2011.
- Wheeler JM: Epidemiology of endometriosis-associated infertility. *J Reprod Med* 34: 41-46, 1989.
- Foti PV, Farina R, Palmucci S, Vizzini IAA, Libertini N, Coronella M, Spadola S, Caltabiano R, Iraci M, Basile A, *et al*: Endometriosis: Clinical features, MR imaging findings and pathologic correlation. *Insights Imaging* 9: 149-172, 2018.
- Nisolle M and Donnez J: Peritoneal endometriosis, ovarian endometriosis, and adenomyotic nodules of the rectovaginal septum are three different entities. *Fertil Steril* 68: 585-596, 1997.
- Cornillie FJ, Oosterlynck D, Lauweryns JM and Koninckx PR: Deeply infiltrating pelvic endometriosis: Histology and clinical significance. *Fertil Steril* 53: 978-983, 1990.
- Johnson NP and Hummelshoj L; World Endometriosis Society Montpellier Consortium: Consensus on current management of endometriosis. *Hum Reprod* 28: 1552-1568, 2013.
- Rafique S and Decherney AH: Medical management of endometriosis. *Clin Obstet Gynecol* 60: 485-496, 2017.
- Guo SW: Recurrence of endometriosis and its control. *Hum Reprod Update* 15: 441-461, 2009.
- Duffy JMN, Arambage K, Correa FJS, Olive D, Farquhar C, Garry R, Barlow DH and Jacobson TZ: Laparoscopic surgery for endometriosis. *Cochrane Database Syst Rev* 3: CD011031, 2014.
- Dmowski WP, Lesniewicz R, Rana N, Pepping P and Noursalehi M: Changing trends in the diagnosis of endometriosis: A comparative study of women with pelvic endometriosis presenting with chronic pelvic pain or infertility. *Fertil Steril* 67: 238-243, 1997.
- Nisenblat V, Bossuyt PMM, Farquhar C, Johnson N and Hull ML: Imaging modalities for the non-invasive diagnosis of endometriosis. *Cochrane Database Syst Rev* 2: CD009591, 2016.
- Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JBR, Leeflang MMG, Sterne JAC and Bossuyt PMM; QUADAS-2 Group: QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 155: 529-536, 2011.
- Pan B, Zhang X and Liu S: Evaluation of the ROC analysis in SPSS in the test/diagnosis. *Strait J Prev Med* 9: 5, 2003.
- Higgins JPT, Altman DG, Sterne JAC (eds). Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (eds), *Cochrane Handbook for Systematic Reviews of Interventions* version 5.2.0 (updated June 2017), Cochrane, 2017.
- Abrao MS, Gonçalves MO Dias JA Jr, Podgaec S, Chamie LP and Blasbalg R: Comparison between clinical examination, transvaginal sonography and magnetic resonance imaging for the diagnosis of deep endometriosis. *Hum Reprod* 22: 3092-3097, 2007.
- Alborzi S, Rasekhi A, Shomali Z, Madadi G, Alborzi M, Kazemi M and Hosseini Nohandani A: Diagnostic accuracy of magnetic resonance imaging, transvaginal, and transrectal ultrasonography in deep infiltrating endometriosis. *Medicine (Baltimore)* 97: e9536, 2018.
- Bazot M, Lafont C, Rouzier R, Roseau G, Thomassin-Naggara I and Daraï E: Diagnostic accuracy of physical examination, transvaginal sonography, rectal endoscopic sonography, and magnetic resonance imaging to diagnose deep infiltrating endometriosis. *Fertil Steril* 92: 1825-1833, 2009.

18. Bergamini V, Ghezzi F, Scarperri S, Raffaelli R, Cromi A and Franchi M: Preoperative assessment of intestinal endometriosis: A comparison of transvaginal sonography with water-contrast in the rectum, transrectal sonography, and barium enema. *Abdom Imaging* 35: 732-736, 2010.
19. Chapron C, Vieira M, Chopin N, Balleyguier C, Barakat H, Dumontier I, Roseau G, Fauconnier A, Foulot H and Dousset B: Accuracy of rectal endoscopic ultrasonography and magnetic resonance imaging in the diagnosis of rectal involvement for patients presenting with deeply infiltrating endometriosis. *Ultrasound Obstet Gynecol* 24: 175-179, 2004.
20. Chamié LP, Blasbalg R, Gonçalves MOC, Carvalho FM, Abrão MS and de Oliveira IS: Accuracy of magnetic resonance imaging for diagnosis and preoperative assessment of deeply infiltrating endometriosis. *Int J Gynaecol Obstet* 106: 198-201, 2009.
21. Dessole S, Farina M, Rubattu G, Cosmi E, Ambrosini G and Nardelli GB: Sonovaginography is a new technique for assessing rectovaginal endometriosis. *Fertil Steril* 79: 1023-1027, 2003.
22. Eskenazi B, Warner M, Bonsignore L, Olive D, Samuels S and Vercellini P: Validation study of nonsurgical diagnosis of endometriosis. *Fertil Steril* 76: 929-935, 2001.
23. Fedele L, Bianchi S, Portuese A, Borruto F and Dorta M: Transrectal ultrasonography in the assessment of rectovaginal endometriosis. *Obstet Gynecol* 91: 444-448, 1998.
24. Grasso RF, Di Giacomo V, Sedati P, Sizzi O, Florio G, Faiella E, Rossetti A, Vescovo RD and Zobel BB: Diagnosis of deep infiltrating endometriosis: accuracy of magnetic resonance imaging and transvaginal 3D ultrasonography. *Abdom Imaging* 35: 716-725, 2010.
25. Guerriero S, Ajossa S, Gerada M, D'Aquila M, Piras B and Melis GB: 'Tenderness-guided' transvaginal ultrasonography: A new method for the detection of deep endometriosis in patients with chronic pelvic pain. *Fertil Steril* 88: 1293-1297, 2007.
26. Guerriero S, Ajossa S, Gerada M, Virgilio B, Angioni S and Melis GB: Diagnostic value of transvaginal 'tenderness-guided' ultrasonography for the prediction of location of deep endometriosis. *Hum Reprod* 23: 2452-2457, 2008.
27. Guerriero S, Saba L, Ajossa S, Peddes C, Angiolucci M, Perniciano M, Melis GB and Alcázar JL: Three-dimensional ultrasonography in the diagnosis of deep endometriosis. *Hum Reprod* 29: 1189-1198, 2014.
28. Guerriero S, Alcázar JL, Pascual MA, Ajossa S, Perniciano M, Piras A, Mais V, Piras B, Schirru F, Benedetto MG and Saba L: Deep infiltrating endometriosis: Comparison between 2-dimensional ultrasonography (US), 3-dimensional US, and magnetic resonance imaging. *J Ultrasound Med* 37: 1511-1521, 2018.
29. Ha HK, Lim YT, Kim HS, Suh TS, Song HH and Kim SJ: Diagnosis of pelvic endometriosis: fat-suppressed T1-weighted vs conventional MR images. *AJR Am J Roentgenol* 163: 127-131, 1994.
30. Holland TK, Yazbek J, Cutner A, Saridogan E, Hoo WL and Jurkovic D: Value of transvaginal ultrasound in assessing severity of pelvic endometriosis. *Ultrasound Obstet Gynecol* 36: 241-248, 2010.
31. Holland TK, Cutner A, Saridogan E, Mavrelou D, Pateman K and Jurkovic D: Ultrasound mapping of pelvic endometriosis: Does the location and number of lesions affect the diagnostic accuracy? A multicentre diagnostic accuracy study. *BMC Womens Health* 13: 43, 2013.
32. Hudelist G, Ballard K, English J, Wright J, Banerjee S, Mastoroudes H, Thomas A, Singer CF and Keckstein J: Transvaginal sonography vs. clinical examination in the preoperative diagnosis of deep infiltrating endometriosis. *Ultrasound Obstet Gynecol* 37: 480-487, 2011.
33. Hudelist G, Fritzer N, Staettner S, Tamma A, Tinelli A, Sparic R and Keckstein J: Uterine sliding sign: A simple sonographic predictor for presence of deep infiltrating endometriosis of the rectum. *Ultrasound Obstet Gynecol* 41: 692-695, 2013.
34. Manganaro L, Vinci V, Bernardo S and Storelli P: The role of 3.0T MRI in the assessment of deep endometriosis located on the uterosacral ligaments. *J Endo* 5: 10-16, 2013.
35. Piessens S, Healey M, Maher P, Tsaltas J and Rombauts L: Can anyone screen for deep infiltrating endometriosis with transvaginal ultrasound? *Aust N Z J Obstet Gynaecol* 54: 462-468, 2014.
36. Ribeiro HSAA, Ribeiro PA, Rossini L, Rodrigues FC, Donadio N and Aoki T: Double-contrast barium enema and transrectal endoscopic ultrasonography in the diagnosis of intestinal deeply infiltrating endometriosis. *J Minim Invasive Gynecol* 15: 315-320, 2008.
37. Saba L, Guerriero S, Sulis R, Piloni M, Ajossa S, Melis G and Mallarini G: Learning curve in the detection of ovarian and deep endometriosis by using magnetic resonance: Comparison with surgical results. *Eur J Radiol* 79: 237-244, 2011.
38. Saccardi C, Cosmi E, Borghero A, Tregnaghi A, Dessole S and Litta P: Comparison between transvaginal sonography, saline contrast sonovaginography and magnetic resonance imaging in the diagnosis of posterior deep infiltrating endometriosis. *Ultrasound Obstet Gynecol* 40: 464-469, 2012.
39. Savelli L, Manuzzi L, Coe M, Mabrouk M, Di Donato N, Venturoli S and Seracchioli R: Comparison of transvaginal sonography and double-contrast barium enema for diagnosing deep infiltrating endometriosis of the posterior compartment. *Ultrasound Obstet Gynecol* 38: 466-471, 2011.
40. Stratton P, Winkel C, Premkumar A, Chow C, Wilson J, Hearn-Stokes R, Heo S, Merino M and Nieman LK: Diagnostic accuracy of laparoscopy, magnetic resonance imaging, and histopathologic examination for the detection of endometriosis. *Fertil Steril* 79: 1078-1085, 2003.
41. Takeuchi H, Kuwatsuru R, Kitade M, Sakurai A, Kikuchi I, Shimanuki H and Kinoshita K: A novel technique using magnetic resonance imaging jelly for evaluation of rectovaginal endometriosis. *Fertil Steril* 83: 442-447, 2005.
42. Ubaldi F, Wisanto A, Camus M, Tournaye H, Clasen K and Devroey P: The role of transvaginal ultrasonography in the detection of pelvic pathologies in the infertility workup. *Hum Reprod* 13: 330-333, 1998.
43. Valenzano Menada M, Remorgida V, Abbamonte LH, Nicoletti A, Ragni N and Ferrero S: Does transvaginal ultrasonography combined with water-contrast in the rectum aid in the diagnosis of rectovaginal endometriosis infiltrating the bowel? *Hum Reprod* 23: 1069-1075, 2008.
44. Vimercati A, Achilarré MT, Scardapane A, Lorusso F, Ceci O, Mangiatori G, Angelelli G, Van Herendael B, Selvaggi L and Bettocchi S: Accuracy of transvaginal sonography and contrast-enhanced magnetic resonance-colonography for the presurgical staging of deep infiltrating endometriosis. *Ultrasound Obstet Gynecol* 40: 592-603, 2012.
45. James FN and Umar S: 2016. The World Bank's classification of countries by income (English). Policy Research working paper; no. WPS 7528. Washington, DC: World Bank Group. Available from: <http://documents.worldbank.org/curated/en/408581467988942234/The-World-Banks-classification-of-countries-by-income>
46. Moura APC, Ribeiro HSAA, Bernardo WM, Simões R, Torres US, D'Ippolito G, Bazot M and Ribeiro PA: Accuracy of transvaginal sonography versus magnetic resonance imaging in the diagnosis of rectosigmoid endometriosis: Systematic review and meta-analysis. *PLoS One* 14: e0214842, 2019.
47. Guerriero S, Saba L, Pascual MA, Ajossa S, Rodriguez I, Mais V and Alcazar JL: Transvaginal ultrasound vs magnetic resonance imaging for diagnosing deep infiltrating endometriosis: Systematic review and meta-analysis. *Ultrasound Obstet Gynecol* 51: 586-595, 2018.
48. Guerriero S, Ajossa S, Orozco R, Perniciano M, Jurado M, Melis GB and Alcazar JL: Accuracy of transvaginal ultrasound for diagnosis of deep endometriosis in the rectosigmoid: Systematic review and meta-analysis. *Ultrasound Obstet Gynecol* 47: 281-289, 2016.
49. Šikić Z: Rules of thumb for positive and negative test results. *J Eval Clin Pract* 28: 1111, 2020.
50. Practice Bulletin no. 114: Management of endometriosis. *Obstet Gynecol* 116: 223-236, 2010.
51. Dunselman GAJ, Vermeulen N, Becker C, Calhaz-Jorge C, D'Hooghe T, Bie BD, Heikinheimo O, Horne AW, Kiesel L, Nap A, et al: ESHRE guideline: Management of women with endometriosis. *Hum Reprod* 29: 400-412, 2014.

