A systematic review and meta-analysis on transvaginal ultrasonography in the diagnosis of deep invasive endometriosis

Yuanhua Xiang, Gang Wang, Lingjuan Zhou, Qin Wang, Qinjie Yang

Department of Ultrasound Medicine, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Taizhou, China
Contributions: (I) Conception and design: Y Xiang, Q Yang; (II) Administrative support: G Wang; (III) Provision of study materials or patients: Y Xiang, G Wang, L Zhou, Q Wang; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: L Zhou, Q Wang, Q Yang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.
Correspondence to: Qinjie Yang. Department of Ultrasound Medicine, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, 150 Ximen Street, Taizhou 317000, China. Email: yangqj@enzemed.com.

Background: The uterus is prone to many diseases, including endometriosis. Transvaginal ultrasonography (TVS) has become a frequently used detection method for deep invasive endometriosis (DIE).

Methods: The combinations of relevant keywords and medical topic title terms were searched in the databases of PubMed, Medline, and Embase, and the Cochrane Central Register of Controlled Trials (endometriosis) from their inception to June 2021. Based on the descriptive terms of endometriosis, deep infiltrating endometrium, heterotopia, deep endometriosis, ultrasound, and TVS, the full texts of the target articles were obtained and subjected to a manual search. Meta-analysis was performed using RevMan 5.3 software provided by the Cochrane collaboration.

Results: A total of 12 articles were included in this study, involving 1,707 patients overall. The sensitivity range was 0.57 to 0.98, and the specificity range was 0.87 to 1.00. The positive likelihood ratio (PLR) was 6.2282 [95% confidence interval (CI): 3.774 to 8.932]; the negative likelihood ratio (NLR) was 0.0664 (95% CI: 0.03 to 0.09); and the duration of remission was 1,174.7 (95% CI: 683.8 to 1,793.4). The sensitivity, specificity, PLR, NLR, and diagnostic odds ratio (DOR) of χ^2 were 36.10 (P=0.021), 27.00 (P=0.035), 53.11 (P=0.001), 55.22 (P=0.001), and 63.89 (P=0.001), respectively, and the differences were statistically significant.

Discussion: A total of 12 articles were included in this meta-analysis, and the results were basically stable. Diagnosis with TVS showed high sensitivity (98%) and specificity (nearly 100%), indicating that it is a reasonable detection method for DIE, improving the disease status of patients.

Keywords: Transvaginal ultrasonography (TVS); deep invasive endometriosis (DIE); meta-analysis

Introduction

Endometriosis is a common uterine disease (1) involving the presence of endometrial outside of the uterine cavity, in other body parts. The general symptoms of endometriosis include menstrual difficulties, infertility, delayed menstruation, and physical pain (2,3). Deep invasive endometriosis (DIE) is the developmental result of ectopic endometriosis, which mainly refers to the depth of endometrial infiltration of the vagina, with associated symptoms including dysmenorrhea, cramps, chronic pelvic pain, and other clinical symptoms (4-6). There is currently no internationally unified staging method for the classification of endometriosis; the commonly used American Fertility Society (AFS) endometriosis modified staging method is mainly used to score the degree of ovarian and fallopian tube adhesions and ovarian endometriotic cysts, but it does not describe the location of the lesion and

282

Xiang et al. Meta-analysis of TVS in DIE

the depth of infiltration (7).

At present, ultrasound and magnetic resonance imaging (MRI) are the main imaging methods for the diagnosis of endometriosis. Clinical studies have shown that both of these two methods have high diagnostic value for endometriosis. However, when there are more fibrostromal components in the lesion and less or no bleeding lesion, MRI has certain difficulty in identifying DIE lesion and MRI (8). Due to the high cost of check, time consuming, and restricted clinical usage, in general, ultrasonic has more application value in diagnosis of DIE. Ultrasound generally fall into the abdominal ultrasound, transvaginal ultrasound by transrectal ultrasound and endoscopic ultrasonography examination (9-14). Transvaginal ultrasound is considered as a kind of convenient safe noninvasive way to check. In addition, with the continuous progress of science, ultrasound equipment and technology are constantly innovating, transvaginal ultrasound examination can well evaluate the situation of various positions and organs in the pelvic cavity, which is the preferred imaging examination method for the diagnosis of death in patients of reproductive age (15).

At present, the diagnostic sensitivity of various auxiliary examinations of DIE reported in international literature is mainly obtained through retrospective case analyses, and to date, there has been no true case study of randomized controlled analysis, and no statistical analysis. In this study, we adopted meta-analysis to analyze the sensitivity of various diagnostic methods of DIE, conduct a quantitative and comprehensive analysis of Chinese and international articles related to DIE, and explore the sensitivity of transvaginal ultrasonography (TVS) diagnostic methods, aiming to provide a reliable reference for clinical examination of DIE. We present the following article in accordance with the PRISMA reporting checklist (available at https://apm.amegroups.com/article/view/10.21037/apm-21-3761/rc).

Methods

Literature retrieval strategy

Combinations of relevant keywords and medical topic title terms were searched in in the databases of PubMed, Medline, and Embase, and the Cochrane Central Register of Controlled Trials (endometriosis). The search time was from the date of establishment of the database to June 2021. Based on the descriptive terms of endometriosis, deep infiltrating endometrium, heterotopia, deep endometriosis, ultrasound, and TVS, the full texts of the target articles were obtained and manual searched.

Inclusion and exclusion criteria of the articles

Articles meeting the following criteria were eligible for inclusion: participants who had been diagnosed with DIE by histopathological examination; participants who experienced symptoms of endometriosis; randomized controlled trials (RCT); participants who had been diagnosed by transvaginal ultrasound (TVU); articles containing the relevant keywords; clinical trials; participants who were over 18 years old; and articles which had been published successfully.

Articles meeting the following criteria were excluded: high similarity to another original text; did not contain the required indicators; non-clinical trials; no application of TVU (such as rectal ultrasound); participants with other reproductive diseases; Chinese articles; other meta-analysis articles; and articles without clear outcomes and incomplete medical data of patients.

Screening and research methods of articles

Articles were screened and data were extracted independently, and cross-check was performed. If there were differing opinions, experts were consulted with to decide the data selection. The selected articles were read carefully to record the sensitivity of the various endometriosis diagnostic methods in each article. The diagnostic sensitivity was the ratio of the discovered cases to the actual total cases, it was a count data, so the relative ratio (RR) of sensitivity of various diagnostic methods could be taken as the effect size. The heterogeneity of meta-analysis comprises 2 main parts: one is the heterogeneity of the quality of the original article. The articles included in this study were all D grades of the random allocation scheme, so the homogeneity was better according to this aspect. Another important source is the consistency or tendency of the various research results included in the analysis, which can be expressed by the test statistic Q, conforming to the χ^2 distribution of v=K–1. According to the heterogeneity test results, the method of effect combination is determined. If there was no obvious heterogeneity among the articles, the fixed effects model (FEM) was selected; while if the heterogeneity among the articles was obvious, the random effects model (REM) was selected.

Table 1 The quality of included 12 articles

Variables	QUADAS items													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Yes (n)	10	12	12	3	10	2	11	11	11	11	12	9	11	12
No (n)	2	0	0	9	1	9	1	1	0	1	0	2	0	0
Unknown (n)	0	0	0	0	1	1	0	0	1	0	0	1	1	0
Yes (%)	83	100	100	25	83	16	92	92	92	92	100	75	92	100

QUADAS, Quality Assessment of Diagnostic Accuracy Studies.

Data extraction

The data were extracted independently by 2 researchers. An Excel table was drawn based on the basic information of the articles, the characteristics of the participants, intervention measures, outcome indicators, and bias evaluation results. After the data extraction was completed, cross-checking was conducted. If there were differences of opinion during the data extraction process, a third researcher was consulted for their opinion and arbitration. The data extracted for this study included basic information of the articles (title, first author, year of publication, author information, and source), basic participant characteristics of the subjects (gender, age, research sample size, and baseline comparability), research methods, research plan design, intervention measures of experimental group and control group, outcome evaluation indicators, and outcome data.

Quality assessment

The quality of the articles was assessed according to the "bias risk assessment" recommended in version 5.3 of the Cochrane Handbook for Systematic Reviews of Interventions (The Cochrane Collaboration, Copenhagen, Denmark). The evaluation contents included the following 7 items: which random method was used; whether allocation concealment was conducted; the implementation of blinding method between patients and researchers; assessment of the effect of blinding method; whether the results were complete; whether the survey results were credible; and other biases. Regarding the RCTs in item 7 above, "satisfied" meant that the bias was relatively low; "not satisfied" meant that the bias was high, and the research failed to give sufficient detailed reports; if it was not mentioned, the risk was unknown. The evaluation included random sequence generation, allocation concealment, blinding, and tracking/

exit. A total of 1–3 points indicated low quality, and 4–7 points indicated high quality. The Quality Assessment of Diagnostic Accuracy Studies (QUADAS) recommended by the Cochrane Collaboration was used as a quality evaluation tool for diagnostic experiments to evaluate the risk of bias in the articles (16). As shown in *Table 1*, if the quality of the articles was 100%, it showed a high score and the highest credibility; while if the quality of the articles was 5–10%, the bias risk was high.

Data extraction

The forest map clearly showed the research results of each article and matched articles with corresponding confidence intervals (CI). If there was no overlap among CI of the articles, it indicated that there was statistical inhomogeneity among the articles. In the case of acceptable inhomogeneity, the REM and FEM could be combined for further analysis. The sensitivity of the research results was analyzed by investigating whether a single study affected the overall results of the combination. Each article included in the study was removed at one time, the results of the remaining articles were combined, and the combined results of each article were compared with the respective results to confirm whether the results were the same. Generally speaking, it is believed that it will have an impact on comprehensive research in the following 2 situations. Firstly, if an article is deleted, the estimated value of the comprehensive combined effect is other than 95% of the combined effect. When an article is deleted, the result will be significantly different. If there is little difference in the results of an article that affects the whole, it indicates the sensitivity of the combined results and the results obtained are unstable. On the contrary, when the results show sensitivity and stability, the conclusion is verified as correct.

Statistical analysis

The data of each document were extracted and summarized according to the 4-column table, and RevMan 5.3 software (The Nordic Cochrane Center, Copenhagen, Denmark, 2014) was used for heterogeneity analysis. The Spearman correlation coefficient of the logarithm of sensitivity and specificity was calculated to detect the heterogeneity caused by the threshold effect. The Q test, χ^2 test, and I^2 test were adopted to detect heterogeneity caused by non-threshold effects. The REM was used to combine sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), and diagnostic odds ratio (DOR) to draw a summary receiver operating characteristic (SROC) curve. The χ^2 test and I^2 test were performed on the combined sensitivity and combined specificity of the 3 inspection methods. A P value <0.05 indicated that the difference was statistically significant.

Results

Results of articles retrieval

A total of 133 related articles were retrieved in this study, of which 66 were retrieved from databases and 67 were retrieved from registers. Then, 26 duplicate articles were eliminated, and 73 articles that obviously did not meet the inclusion criteria were excluded after reading of their titles and abstracts, and finally 12 articles (17-28) that satisfied the inclusion criteria were included. The specific article selection process is illustrated in *Figure 1*. *Table 2* shows the author, year of publication, number of participants, true positive results (TP), false positive results (FP), false negative results (FN), and true negative results (TN).

Risk bias assessment of the included articles

The Cochrane Handbook version 5.3 systematic review writing manual was adopted to evaluate the risk of bias among the 10 included articles, and output the risk of bias chart, as shown in *Figures 2,3*. In addition, the RevMan 5.3 software was applied to express the risk of bias in quality evaluation of the articles. The risk of case selection bias was mainly due to the lack of emphasis on whether the participants were included consecutively or not. A total of 12 articles were evaluated for offset risk. Except for 2 articles, the remaining articles were all low-risk. The risk of deviation of the diagnostic gold standard mainly refers to whether the description of the diagnostic gold standard

is clear. Five articles regarded the results of laparoscopy as the gold standard, rather than the results of pathological examination. A total of 7 articles did not clearly indicate whether there is a difference between the part to be tested and the standard part. The applicability of the cases in the 10 articles was better, and the applicability of the part to be tested and the gold standard part were better.

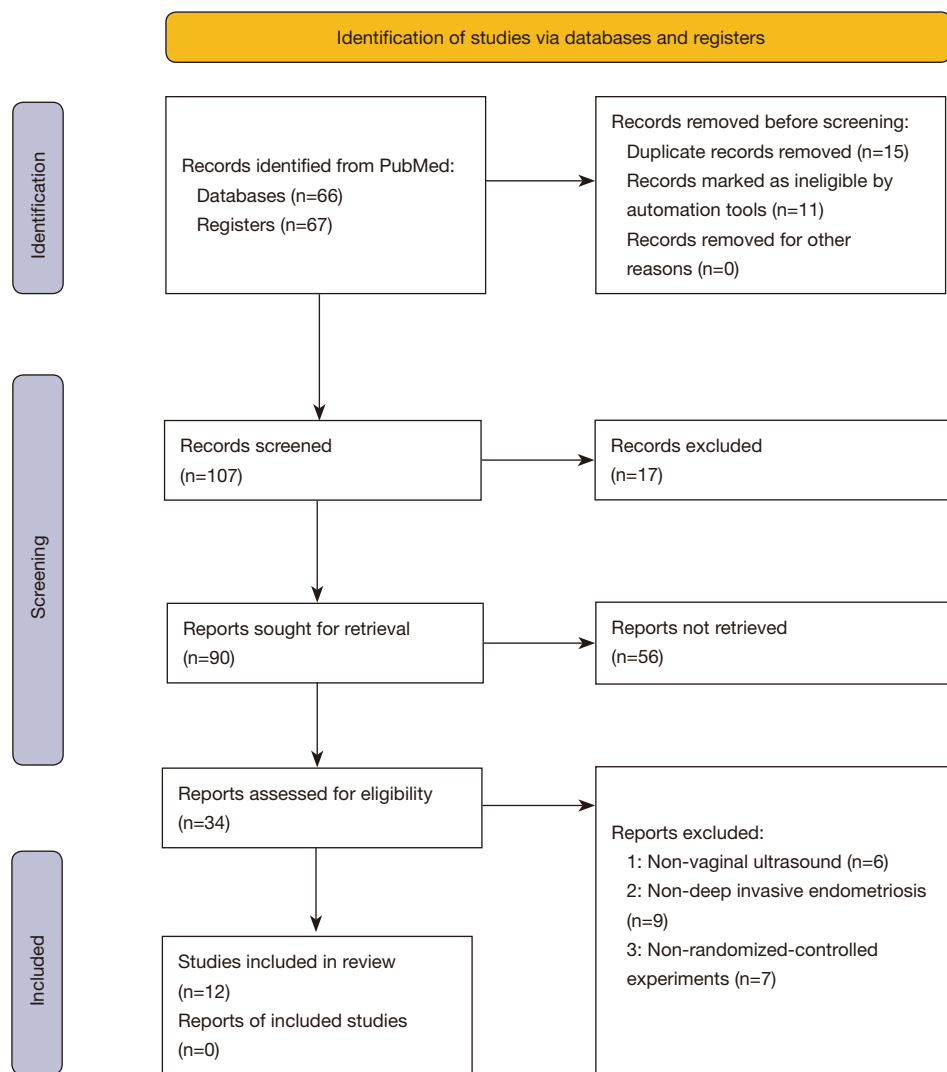### Meta-analysis on TVU detection of DIE

A forest plot of sensitivity and specificity in the diagnosis of DIE by TVU is shown in *Figure 4*. The sensitivity ranged from 0.57 to 0.98 [mean (M) =0.9437; 95% CI: 0.07 to 1.00], and the specificity range was from 0.87 to 1.00 (M =0.8485; 95% CI: 0.16 to 1.00). The PLR was 6.2282 (95% CI: 3.774 to 8.932), NLR was 0.0664 (95% CI: 0.03 to 0.09), and DOR was 1,174.7 (95% CI: 683.8 to 1,793.4). The sensitivity, specificity, PLR, NLR, and DOR values of χ^2 were 36.10 (P=0.021), 27.00 (P=0.035), 53.11 (P=0.001), 55.22 (P=0.001), and 63.89 (P=0.001), respectively, showing that there was obvious heterogeneity among the articles. Heterogeneity analysis threshold effect showed that the Spearman correlation coefficient of the sensitivity logarithm and specificity logarithm of TVU was –0.31 (P=0.22), and the difference was statistically visible. There was heterogeneity caused by non-threshold effects, so the REM was used to combine the statistical indicators. The sensitivity, specificity, PLR, NLR, and DOR are shown in *Table 3*.

SROC for TVU to detect DIE

The SROC shows a global summary of test performance and shows the trade-off between sensitivity and specificity. As illustrated in *Figure 5*, the SROC of the TVU results showed the TP rate and FP rate of individual articles. The data showed that the SROC was located near the ideal upper left corner of the SROC, and the maximum sensitivity and specificity (q value) was 0.69; while the area under the curve (AUC) was 0.97, indicating a higher overall accuracy.

Discussion

In this meta-analysis, there were a total of 12 articles included. In the quality evaluation of these 12 articles, 1 was assessed as quality grade A and 9 as quality grade B. A DIE ultrasound shows irregular hypoechoic feedback located in

Identification of studies via databases and registers

Identification

Records identified from PubMed:
Databases (n=66)
Registers (n=67)

Records removed before screening:
Duplicate records removed (n=15)
Records marked as ineligible by automation tools (n=11)
Records removed for other reasons (n=0)

Screening

Records screened
(n=107)

Records excluded
(n=17)

Reports sought for retrieval
(n=90)

Reports not retrieved
(n=56)

Reports assessed for eligibility
(n=34)

Reports excluded:
1: Non-vaginal ultrasound (n=6)
2: Non-deep invasive endometriosis (n=9)
3: Non-randomized-controlled experiments (n=7)

Included

Studies included in review
(n=12)
Reports of included studies
(n=0)

Figure 1 The entire literature retrieval and screening process.

the muscle layer or elsewhere. The diagnosis and evaluation of DIE ultrasound information can guide the selection of reasonable treatment methods during the treatment process. Diagnosis of DIE is a pathological examination, and the non-invasive forms include ultrasound and MRI. However, the study by Leonardi *et al.* (29) showed that MRI was not as effective as ultrasound for detecting DIE. However, when there are more fibrous interstitial components in the lesion and less or no bleeding lesion, MRI has some difficulty in identifying DIE lesion. Moreover, MRI examination is expensive and time-consuming, and its wide clinical application is limited. Therefore, compared with MRI, ultrasound in the diagnosis of DIE is not accurate,

convenient, fast, cost-effective, and easy to operate (30,31).

This study showed that among the sensitivity, specificity, PLR, NLR, DOR, and AUC of SROC, PLR and NLR were more stable than sensitivity and specificity. The larger the PLR and the smaller the NLR, the higher the diagnostic accuracy. The higher the DOR value, the higher the diagnostic accuracy. Both PLR and NLR are related to the selection of diagnostic threshold points. The AUC eliminates the influence of the diagnosis boundary value and can better reflect the diagnosis efficiency. The larger the AUC, the higher the diagnostic accuracy. Therefore, in this study, the combination of the SROC curve of the TVU diagnostic value and the DOR had a higher value,

286

Xiang et al. Meta-analysis of TVS in DIE

Table 2 The basic characteristics of the included articles

First author	Year of publication	Number of cases	TP	FP	FN	TN	Score
Abrao (17)	2007	104	53	0	1	50	8
Bazot (18)	2009	92	59	0	4	29	7
Exacoustos (19)	2014	104	67	5	4	28	7
Fratelli (20)	2013	420	126	10	64	220	9
Holland (21)	2013	198	3	2	6	187	7
Hudelist (22)	2011	129	28	1	3	97	5
Hudelist (23)	2009	200	46	3	2	149	8
Piketty (24)	2009	134	68	2	7	57	9
Savelli (25)	2011	58	51	0	5	2	8
Scioscia (26)	2020	88	8	1	5	9	7
Valenzano (27)	2008	90	13	5	10	62	8
Vimercati (28)	2012	90	14	4	4	68	8

TP, true positive; FP, false positive; FN, false negative; TN, true negative.

indicating that the TVU has a good display rate and is beneficial to detecting lesions. The PLR was 6.2282 (95% CI: 3.774 to 8.932), and the NLR was 0.0664 (95% CI: 0.03 to 0.09). The larger the PLR and the smaller the NLR, the higher the quality of the results. The duration of response was related to PLR and NLR, and the value was larger (DOR =1,174.7, 95% CI: 683.8 to 1,793.4), indicating that the TVU detection of endometriosis is more effective. The results all showed that there was significant heterogeneity among the included articles. Heterogeneity analysis threshold effect showed that the Spearman correlation coefficient of the sensitivity logarithm and specificity logarithm of TVU was –0.31 (P=0.22), and the difference was statistically significant. The results indicate that TVU diagnosis is effective in the detection of DIE and can be used in routine examinations.

This study has some limitations. Firstly, the lack of meeting records and unpublished articles among the searches resulted in a small sample size. Secondly, the language of the articles was limited to English, and the number of articles included would have been greater if non-English language articles had been permitted. In general, the quality of the included articles was higher than the median level of QUADAS. Many articles did not meet items 4 (disease progression bias), 5 (partial verification), 11 (reference standard review bias), 13 (unexplainable

test results), or 14 (withdrawn). According to the detailed analysis of the QUADAS project and research, erroneous classification deviations may occur. It is recommended to further improve the experimental plan, standardize the specific time, methods, and drugs of periodic intervention, and implement high-quality, large-scale samples, and multi-center randomized controlled testing to obtain more reliable evidence.

Conclusions

The overall evidence showed that TVU diagnosis was an accurate non-invasive detection method for the detection of DIE in women. Meta-analysis results were basically stable, and TVU diagnosis showed high sensitivity (98%) and specificity (nearly 100%), which may provide a reasonable detection method for DIE and improve the disease status of patients. However, there following limitations remained: the retrieval may be incomplete due to different retrieval mechanisms; there were differences in the evaluation or measurement of risk factors in different studies, which may affect the results. The articles included in the meta-analysis were all published successfully, and failure to include unpublished articles may have potential publication bias. It is recommended that more high-quality, multi-center, large-sample original studies be conducted for verification in the
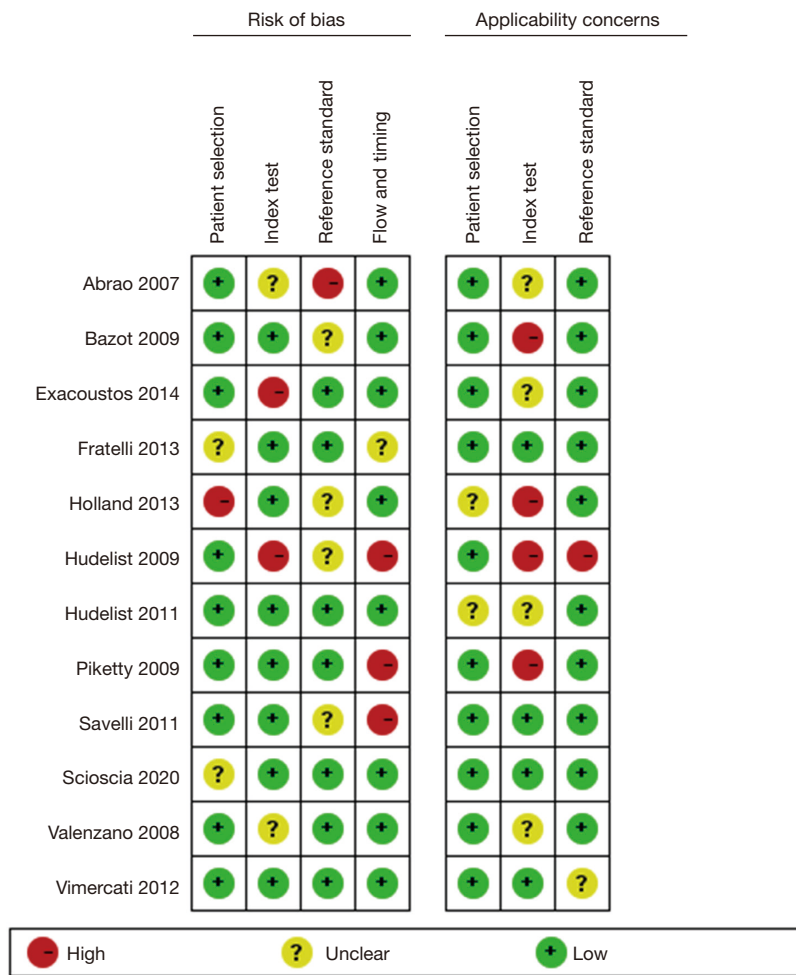
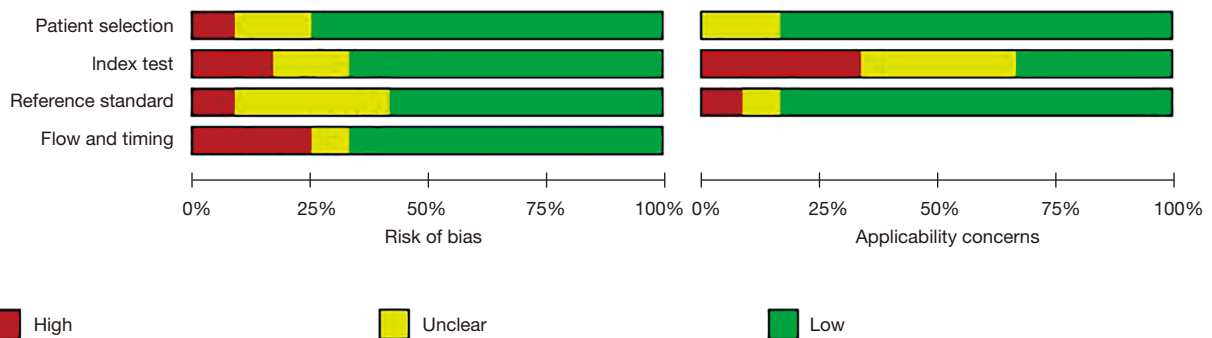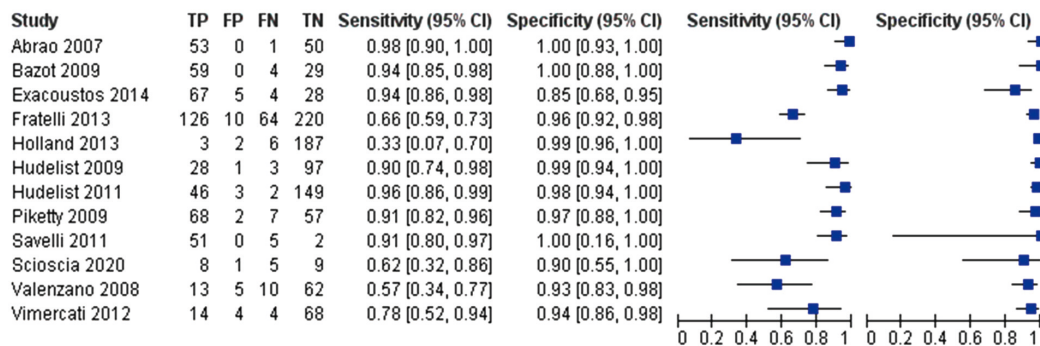**Figure 2** Multiple risk bias evaluation results of included articles.

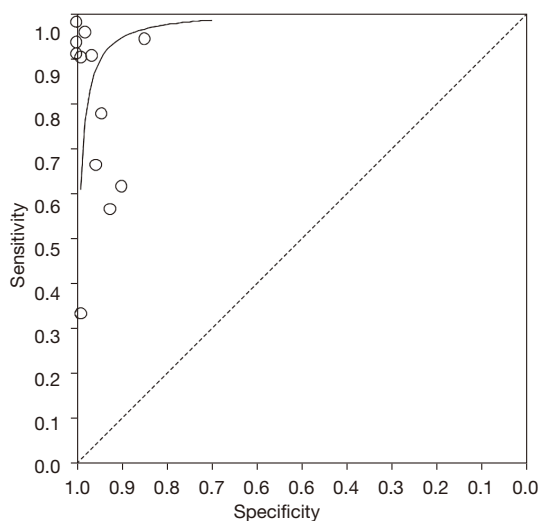**Figure 3** Assessment of risk bias of the included articles.

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Abrao 2007	53	0	1	50	0.98 [0.90, 1.00]	1.00 [0.93, 1.00]
Bazot 2009	59	0	4	29	0.94 [0.85, 0.98]	1.00 [0.88, 1.00]
Exacoustos 2014	67	5	4	28	0.94 [0.86, 0.98]	0.85 [0.68, 0.95]
Fratelli 2013	126	10	64	220	0.66 [0.59, 0.73]	0.96 [0.92, 0.98]
Holland 2013	3	2	6	187	0.33 [0.07, 0.70]	0.99 [0.96, 1.00]
Hudelist 2009	28	1	3	97	0.90 [0.74, 0.98]	0.99 [0.94, 1.00]
Hudelist 2011	46	3	2	149	0.96 [0.86, 0.99]	0.98 [0.94, 1.00]
Piketty 2009	68	2	7	57	0.91 [0.82, 0.96]	0.97 [0.88, 1.00]
Savelli 2011	51	0	5	2	0.91 [0.80, 0.97]	1.00 [0.16, 1.00]
Scioscia 2020	8	1	5	9	0.62 [0.32, 0.86]	0.90 [0.55, 1.00]
Valenzano 2008	13	5	10	62	0.57 [0.34, 0.77]	0.93 [0.83, 0.98]
Vimercati 2012	14	4	4	68	0.78 [0.52, 0.94]	0.94 [0.86, 0.98]

Figure 4 Forest map of the sensitivity and specificity of TVU in detecting deep invasive endometriosis. TVU, transvaginal ultrasound; CI, confidence interval; TP, true positive; FP, false positive; FN, false negative; TN, true negative.

Table 3 The indicators of the TVU

Indicators	Outcome indicators (95% CI)
Sensitivity	0.9437 (0.07–1.00)
Specificity	0.8485 (0.16–1.00)
PLR	6.2282 (3.774–8.932)
NLR	0.0664 (0.03–0.09)
DOR	1,174.7 (683.8–1,793.4)

TVU, transvaginal ultrasound; CI, confidence interval; PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR, diagnostic odds ratio.

future, to provide early warning for the clinic, to detect DIE more quickly and efficiently, and to provide a scientific basis for the treatment of endometriosis.

Figure 5 SROC for TVU to detect deep invasive endometriosis. Each article was represented by each solid circle in the meta-analysis. The size of the solid circle represented the size of each article. The SROC summarized the accuracy of the overall diagnosis. SROC, summary receiver operating characteristic; TVU, transvaginal ultrasound.

licenses/by-nc-nd/4.0/.

References

1. Chapron C, Marcellin L, Borghese B, et al. Rethinking mechanisms, diagnosis and management of endometriosis. Nat Rev Endocrinol 2019;15:666-82.
2. Zondervan KT, Becker CM, Missmer SA. Endometriosis. N Engl J Med 2020;382:1244-56.
3. Bulun SE, Yilmaz BD, Sison C, et al. Endometriosis. Endocr Rev 2019;40:1048-79.
4. Broi MGD, Ferriani RA, Navarro PA. Ethiopathogenic mechanisms of endometriosis-related infertility. JBRA Assist Reprod 2019;23:273-80.
5. Králíčková M, Laganà AS, Ghezzi F, et al. Endometriosis and risk of ovarian cancer: what do we know? Arch Gynecol Obstet 2020;301:1-10.
6. Yilmaz BD, Bulun SE. Endometriosis and nuclear receptors. Hum Reprod Update 2019;25:473-85.
7. Koninckx PR, Ussia A, Adamyan L, et al. Pathogenesis of endometriosis: the genetic/epigenetic theory. Fertil Steril 2019;111:327-40.
8. Laganà AS, Garzon S, Götte M, et al. The Pathogenesis of Endometriosis: Molecular and Cell Biology Insights. Int J Mol Sci 2019;20:5615.
9. Leonardi M, Hicks C, El-Assaad F, et al. Endometriosis and the microbiome: a systematic review. BJOG 2020;127:239-49.
10. Sachedina A, Todd N. Dysmenorrhea, Endometriosis and Chronic Pelvic Pain in Adolescents J Clin Res Pediatr Endocrinol 2020;12:7-17.
11. Guerriero S, Saba L, Pascual MA, et al. Transvaginal ultrasound vs magnetic resonance imaging for diagnosing deep infiltrating endometriosis: systematic review and meta-analysis. Ultrasound Obstet Gynecol 2018;51:586-95.
12. Van den Bosch T, Van Schoubroeck D. Ultrasound diagnosis of endometriosis and adenomyosis: State of the art. Best Pract Res Clin Obstet Gynaecol 2018;51:16-24.
13. de Oliveira JGA, Bonfada V, Zanella JFP, et al. Transvaginal ultrasound in deep endometriosis: pictorial essay. Radiol Bras 2019;52:337-41.
14. Zhou Y, Su Y, Liu H, et al. Accuracy of transvaginal ultrasound for diagnosis of deep infiltrating endometriosis in the uterosacral ligaments: Systematic review and meta-analysis. J Gynecol Obstet Hum Reprod 2021;50:101953.
15. Deslandes A, Parange N, Childs JT, et al. Current Status of Transvaginal Ultrasound Accuracy in the Diagnosis of Deep Infiltrating Endometriosis Before Surgery: A Systematic Review of the Literature. J Ultrasound Med 2020;39:1477-90.
16. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529-36.
17. Abrao MS, Gonçalves MO, Dias JA Jr, et al. Comparison between clinical examination, transvaginal sonography and magnetic resonance imaging for the diagnosis of deep endometriosis. Hum Reprod 2007;22:3092-7.
18. Bazot M, Lafont C, Rouzier R, et al. Diagnostic accuracy of physical examination, transvaginal sonography, rectal endoscopic sonography, and magnetic resonance imaging to diagnose deep infiltrating endometriosis. Fertil Steril 2009;92:1825-33.
19. Exacoustos C, Malzoni M, Di Giovanni A, et al. Ultrasound mapping system for the surgical management of deep infiltrating endometriosis. Fertil Steril 2014;102:143-150.e2.
20. Fratelli N, Scioscia M, Bassi E, et al. Transvaginal sonography for preoperative assessment of deep endometriosis. J Clin Ultrasound 2013;41:69-75.
21. Holland TK, Cutner A, Saridogan E, et al. Ultrasound mapping of pelvic endometriosis: does the location and number of lesions affect the diagnostic accuracy? A multicentre diagnostic accuracy study. BMC Womens Health 2013;13:43.
22. Hudelist G, Ballard K, English J, et al. Transvaginal sonography vs. clinical examination in the preoperative diagnosis of deep infiltrating endometriosis. Ultrasound Obstet Gynecol 2011;37:480-7.
23. Hudelist G, Tuttlies F, Rauter G, et al. Can transvaginal sonography predict infiltration depth in patients with deep infiltrating endometriosis of the rectum? Hum Reprod 2009;24:1012-7.
24. Piketty M, Chopin N, Dousset B, et al. Preoperative work-up for patients with deeply infiltrating endometriosis: transvaginal ultrasonography must definitely be the first-line imaging examination. Hum Reprod 2009;24:602-7.
25. Savelli L, Manuzzi L, Coe M, et al. Comparison of transvaginal sonography and double-contrast barium enema for diagnosing deep infiltrating endometriosis of the posterior compartment. Ultrasound Obstet Gynecol 2011;38:466-71.
26. Scioscia M, Zanetti I, Raspanti X, et al. Ultrasound Differential Diagnosis in Deep Infiltrating Endometriosis of the Urinary Tract. J Ultrasound Med 2020;39:2261-75.

27. Valenzano Menada M, Remorgida V, Abbamonte LH, et al. Does transvaginal ultrasonography combined with water-contrast in the rectum aid in the diagnosis of rectovaginal endometriosis infiltrating the bowel? Hum Reprod 2008;23:1069-75.
28. Vimercati A, Achilarre MT, Scardapane A, et al. Accuracy of transvaginal sonography and contrast-enhanced magnetic resonance-colonography for the presurgical staging of deep infiltrating endometriosis. Ultrasound Obstet Gynecol 2012;40:592-603.
29. Leonardi M, Martin E, Reid S, et al. Deep endometriosis transvaginal ultrasound in the workup of patients with signs and symptoms of endometriosis: a cost analysis. BJOG 2019;126:1499-506.
30. Martire FG, Lazzeri L, Conway F, et al. Adolescence and endometriosis: symptoms, ultrasound signs and early diagnosis. Fertil Steril 2020;114:1049-57.
31. Filip L, Duică F, Prădatu A, et al. Endometriosis Associated Infertility: A Critical Review and Analysis on Etiopathogenesis and Therapeutic Approaches. Medicina (Kaunas) 2020;56:460.

(English Language Editor: J. Jones)