Ultrasound Obstet Gynecol 2023; **61**: 310–324
Published online 11 February 2023 in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.24976.

Diagnostic accuracy of ultrasound signs for detecting adnexal torsion: systematic review and meta-analysis

I. GARDE[1], C. PAREDES[2], L. VENTURA[3], M. A. PASCUAL[4], S. AJOSSA[5], S. GUERRIERO[5], J. VARA[6], M. LINARES[7] and J. L. ALCÁZAR[6]

[1]*Department of Obstetrics and Gynecology, Hospital Universitario de Cruces, Bilbao, Spain;* [2]*Department of Obstetrics and Gynecology, Hospital Materno-Infantil, Badajoz, Spain;* [3]*School of Medicine, University of Navarra, Pamplona, Spain;* [4]*Department of Obstetrics, Gynecology, and Reproduction, Institut Universitary Dexeus, Barcelona, Spain;* [5]*Centro Integrato di Procreazione Medicalmente Assistita (PMA) e Diagnostica Ostetrico-Ginecologica, Azienda Ospedaliero Universitaria-Policlinico Duilio Casula, Monserrato, University of Cagliari, Cagliari, Italy;* [6]*Department of Obstetrics and Gynecology, Clínica Universidad de Navarra, Pamplona, Spain;* [7]*Department of Obstetrics and Gynecology, Hospital Universitario Puerta del Mar, Cadiz, Spain*

KEYWORDS: adnexa; Doppler; ovary; torsion; ultrasound

CONTRIBUTION

What are the novel findings of this work?
This systematic review and meta-analysis shows that the presence of an adnexal mass and pelvic fluid have poor diagnostic accuracy as ultrasound signs of adnexal torsion, while ovarian edema, the whirlpool sign and decreased or absent ovarian Doppler flow show good specificity but moderate sensitivity for detecting adnexal torsion. The quality of studies assessing the role of different ultrasound signs for diagnosing adnexal torsion was moderate to good. There was significant heterogeneity among studies, which may limit the reliability of our findings regarding the diagnostic capability of sonographic signs in the diagnosis of adnexal torsion.

What are the clinical implications of this work?
We present up-to-date evidence on the role of ultrasound in diagnosing adnexal torsion. Most of the classic ultrasound signs are specific but have moderate sensitivity for diagnosing this condition.

ABSTRACT

Objective *To evaluate the diagnostic accuracy of different ultrasound signs for diagnosing adnexal torsion, using surgery as the reference standard.*

Methods *This was a systematic review and meta-analysis of studies published between January 1990 and November 2021 evaluating ovarian edema, adnexal mass, ovarian Doppler flow findings, the whirlpool sign and pelvic fluid as ultrasound signs (index tests) for detecting adnexal torsion, using surgical findings as the reference standard. The search for studies was performed in PubMed/MEDLINE, CINAHL, Scopus, The Cochrane Library, ClinicalTrials.gov and Web of Science databases. The Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool was used to evaluate the quality of the studies. Pooled sensitivity, specificity, and positive and negative likelihood ratios were calculated separately, and the post-test probability of adnexal torsion following a positive or negative test was also determined.*

Results *The search identified 1267 citations after excluding duplicates. Eighteen studies were ultimately included in the qualitative and quantitative syntheses. Eight studies (809 patients) analyzed the presence of ovarian edema, eight studies (1044 patients) analyzed the presence of an adnexal mass, 14 studies (1742 patients) analyzed ovarian Doppler flow, six studies (545 patients) analyzed the whirlpool sign and seven studies (981 patients) analyzed the presence of pelvic fluid as ultrasound signs of adnexal torsion. Overall, the quality of most studies was considered to be moderate or good. However, there was a high risk of bias in the patient-selection and index-text domains*

Correspondence to: Dr J. L. Alcázar, Department of Obstetrics and Gynecology, Clínica Universidad de Navarra, Avenida Pío XII 36, 3110 Pamplona, Spain (e-mail: jlalcazar@unav.es)

Accepted: 10 June 2022

(with the exception of the whirlpool sign) in a significant proportion of studies. Pooled sensitivity, specificity, and positive and negative likelihood ratios of each ultrasound sign were 58%, 86%, 4.0 and 0.49 for ovarian edema, 69%, 46%, 1.3 and 0.67 for adnexal mass, 65%, 91%, 7.6 and 0.38 for the whirlpool sign, 53%, 95%, 11.0 and 0.49 for ovarian Doppler findings and 55%, 69%, 1.7 and 0.66 for pelvic fluid. Heterogeneity was high for all analyses.

Conclusions *The presence of an adnexal mass or pelvic fluid have poor diagnostic accuracy as ultrasound signs of adnexal torsion, while the presence of ovarian edema, the whirlpool sign and decreased or absent ovarian Doppler flow have good specificity but moderate sensitivity for detecting adnexal torsion. © 2022 The Authors. Ultrasound in Obstetrics & Gynecology published by John Wiley & Sons Ltd on behalf of International Society of Ultrasound in Obstetrics and Gynecology.*

INTRODUCTION

Adnexal torsion is a relatively common problem in clinical practice, accounting for about 3% of all gynecological emergencies[1]. It involves an abnormal rotation of the ovary and/or the Fallopian tube on their supporting ligaments around the vascular axis. Four pathological patterns have been described: tubo-ovarian torsion, ovarian torsion, tubal torsion and mesenterotubal torsion[2]. It can occur in female patients of any age but is more frequent in the reproductive period and is rare in postmenopausal women[3]. The main concern of this condition is that adnexal torsion may lead to the loss of the adnexa, more specifically the ovary.

The diagnosis of adnexal torsion is mostly based on clinical symptoms, and it should be suspected in cases of acute unilateral lower abdominal pain associated with nausea and/or vomiting and several laboratory findings, especially leukocytosis[1]. However, these symptoms and signs are quite non-specific. The rotation of the adnexa on its pedicle indicates compromised blood supply, which can be detected on Doppler ultrasound by identifying the decrease or absence of arterial and/or venous blood flow[4]. This situation causes a series of histological reactive changes in the ovary that are visible on ultrasound, such as enlarged ovaries with hyperechogenic stroma and follicles arranged on the periphery, also known as ovarian edema[5]. Adnexal torsion is more frequent in patients with ovarian cysts or masses, as well as in cases of ovarian hyperstimulation[3].

For all these reasons, ultrasound plays an important role in the correct diagnosis of adnexal torsion and is considered to be the imaging modality of choice in affected patients[6]. Many studies have been carried out to date to evaluate the diagnostic performance of ultrasound for diagnosing this pathology. Several meta-analyses analyzing the diagnostic performance of ultrasound in cases of adnexal torsion have also been published[7–9]. However, with the exception of one meta-analysis reporting on the whirlpool sign[8], none of them performed an independent analysis for each of the classic ultrasound signs.

We aimed to perform a systematic review and meta-analysis on the diagnostic accuracy of ultrasound signs for detecting adnexal torsion.

METHODS

Protocol and registration

This systematic review and meta-analysis was performed according to the PRISMA statement[10] and SEDATE guidelines[11]. All methods regarding inclusion/exclusion criteria, data extraction and quality assessment were defined *a priori* (Appendix S1). The methodology was registered in PROSPERO (registration number: pending; provisional ID: 312976) before the initiation of the study. No amendments were made after registration. Institutional review board approval was waived because of the nature and design of the study.

Data sources and search

Three of the authors (I.G., C.P. and J.L.A.) searched six electronic databases (PubMed/MEDLINE, CINAHL, Scopus, The Cochrane Library, ClinicalTrials.gov and Web of Science) to identify potentially eligible studies published between January 1990 and November 2021. The search terms were as follows: 'ultrasound', 'adnexa', 'ovary', 'torsion' and 'Doppler'. Therefore, the following Boolean operators were used: ultrasound AND Doppler AND adnexa OR ovary AND torsion. Only articles published in English, Spanish or French were included.

Study selection

Three authors (I.G., C.P. and L.V.) screened the title and abstract of the identified articles in order to exclude those that were irrelevant, including duplicates, studies not strictly related to the topic of review, case reports, reviews, meta-analyses and letters to the editor. The full texts of relevant articles were then obtained, and the reviewers (I.G., C.P. and J.L.A.) applied independently the inclusion criteria described below.

Eligible studies were prospective or retrospective cohort or case–control studies with at least 20 women included (sample size was set arbitrarily). Participants included girls (premenarchal), adolescents (postmenarchal females under the age of 20), premenopausal non-pregnant and pregnant women, and postmenopausal women with clinical suspicion of adnexal torsion. The index test was ultrasound assessment performed via the transvaginal, abdominal or transrectal route for at least one of the following ultrasound signs related to adnexal torsion[12]: ovarian edema ('enlarged ovary' was considered as ovarian edema), adnexal mass (presence of an adnexal mass distinct from an enlarged ovary, such as a dermoid, simple or hemorrhagic cyst), the whirlpool sign, Doppler assessment of ovarian blood flow (absent *vs* present

ovarian color-map or pulsed-Doppler assessment for venous and/or arterial blood flow for detecting decreased or absent flow), intrafollicular fluid-debris level, the follicular ring sign and fluid in the pelvis. Surgery with or without pathological correlation was used as the reference standard. The minimum data requirement was sufficient data reported to construct a 2×2 table of diagnostic performance. Studies that assessed isolated tubal torsion were not considered for this meta-analysis. Studies including fetuses and/or neonates were also excluded.

The 'snowball strategy' was used to identify potentially relevant papers from the reference lists of papers selected for full-text assessment. In the case of missing relevant data, we sought to contact the authors to request this information.

Data collection process

As stated above, seven ultrasound signs related to the presence of adnexal torsion were selected for this meta-analysis and included ovarian edema, follicular ring, intrafollicular fluid-debris level, the presence of an adnexal mass, ovarian Doppler flow findings, the whirlpool sign and the presence of pelvic free fluid.

The following data were extracted from each of the studies included: first author's name, year of publication, study design (prospective or retrospective cohort or case–control study), population (girls, adolescents, premenopausal non-pregnant and pregnant women or postmenopausal women), recruitment period, sample size, index test (ultrasound route (transvaginal or transabdominal), Doppler settings used, ultrasound sign evaluated), number and experience of examiners, blinding of examiners to clinical presentation or surgical outcome, surgical approach and total cases of adnexal torsion confirmed by surgical findings.

Risk of bias assessment

Quality assessment was carried out using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool adapted for this meta-analysis to assess the risk of bias in individual studies[13]. The QUADAS-2 tool includes four domains: patient selection, index test, reference standard and flow and timing. For each domain, the risk of bias and concerns regarding applicability were classified as high, low or unclear. The results of quality assessment were used for descriptive purposes to evaluate the overall quality of the included studies and to investigate potential sources of heterogeneity. Two authors (J.V. and M.L.) independently assessed the methodological quality using a standard form with quality assessment criteria.

Disagreements were resolved by discussion moderated by a third author (J.L.A.) to reach a consensus. The authors determined the risk of selection bias based on the description of inclusion and exclusion criteria of the studies. For the index-text domain, we assessed whether the ultrasound-sign definition used in the study was clear enough to be replicated in a different study. For evaluation of the reference-standard domain, the method used in the study to determine the presence of adnexal torsion was assessed. For evaluation of the flow-and-timing domain, the description of the time elapsed between the index-test assessment and the reference-standard result was evaluated.

Statistical analysis

Data on the diagnostic performance of ultrasound signs assessed in this meta-analysis were extracted or derived from the included studies. The test result was considered positive when the sign assessed was visualized during the ultrasound examination. Consequently, the test result was considered negative when the sign was not visualized or not specifically mentioned as having been visualized in the evaluated study. In the latter case, it was assumed that the sign was searched for and was not found. The reference standard was adnexal torsion found at surgery.

The primary outcome was pooled sensitivity, specificity and positive and negative likelihood ratios (LR+ and LR−) as well as the diagnostic odd ratio (DOR) of each ultrasound sign in the detection of adnexal torsion. True-positive, true-negative, false-positive and false-negative values were obtained from each study. Post-test probabilities were calculated and plotted on Fagan nomograms.

The presence of heterogeneity in sensitivity and specificity was assessed graphically, by plotting forest plots, as well as by using Cochran's Q statistic and the I^2 index. Tests for heterogeneity examine the null hypothesis that all studies are evaluating the same effect; $P < 0.1$ indicates heterogeneity. According to Higgins *et al.*, I^2 values of 25%, 50% and 75% are considered to indicate low, moderate and high heterogeneity, respectively[14]. A summary receiver-operating-characteristics (sROC) curve was plotted to illustrate the relationship between sensitivity and false-positive rate. If heterogeneity was observed, metaregression was performed using the following as covariables: year of publication, sample size, prevalence of adnexal torsion, study design and population studied.

Statistical analysis was performed using Meta-analytical Integration of Diagnostic Accuracy Studies (MIDAS) and METANDI commands in Stata version 12 for Windows (Stata Corp., College Station, TX, USA); $P < 0.05$ was considered to indicate statistical significance.

RESULTS

Search results

A flowchart summarizing literature identification and selection of studies is shown in Figure 1. The electronic search identified 1949 citations (853 in PubMed/MEDLINE, 16 in CINAHL, 796 in Scopus, none in The Cochrane Library, one in ClinicalTrials.gov and 283 in Web of Science). After removal of 682 duplicate

records, 1267 citations remained. Of these, 1064 papers were excluded after reading the title and another 100 were excluded after reading the abstract.

We examined the full text of the remaining 103 articles. Eighty-five studies were excluded because of the following reasons: sample size < 20 cases, overlapping patient populations, case report, the index test was not ultrasound, the reference standard was not available or it was not surgical findings only, or there were insufficient data to construct a 2 × 2 table (for example, retrospective studies in which all cases had adnexal torsion and therefore the false-positive and true-negative cases were zero by definition) (Appendix S2). Only one study assessed intrafollicular fluid-debris level[15] and only three studies assessed the follicular ring sign[16–18]. We decided to exclude these signs from the meta-analysis because the number of identified studies was insufficient to perform quantitative synthesis.

Eighteen studies were ultimately included in the analysis[17–34]. The studies analyzed the accuracy of ovarian edema, adnexal mass, the whirlpool sign, ovarian Doppler flow and pelvic fluid for detecting adnexal torsion in patients with clinical suspicion of the condition. Most studies assessed more than one ultrasound sign. The analysis was performed separately for each ultrasound sign. There was no need to contact the authors of any of the studies, as all relevant data to perform the meta-analysis were available.

Characteristics of included studies

The characteristics of the selected studies are shown in Table 1. The studies were published between 1998 and 2022 and reported on 2101 patients. Among these 2101 patients, 870 had adnexal torsion at surgery (by laparoscopic or laparotomy access).

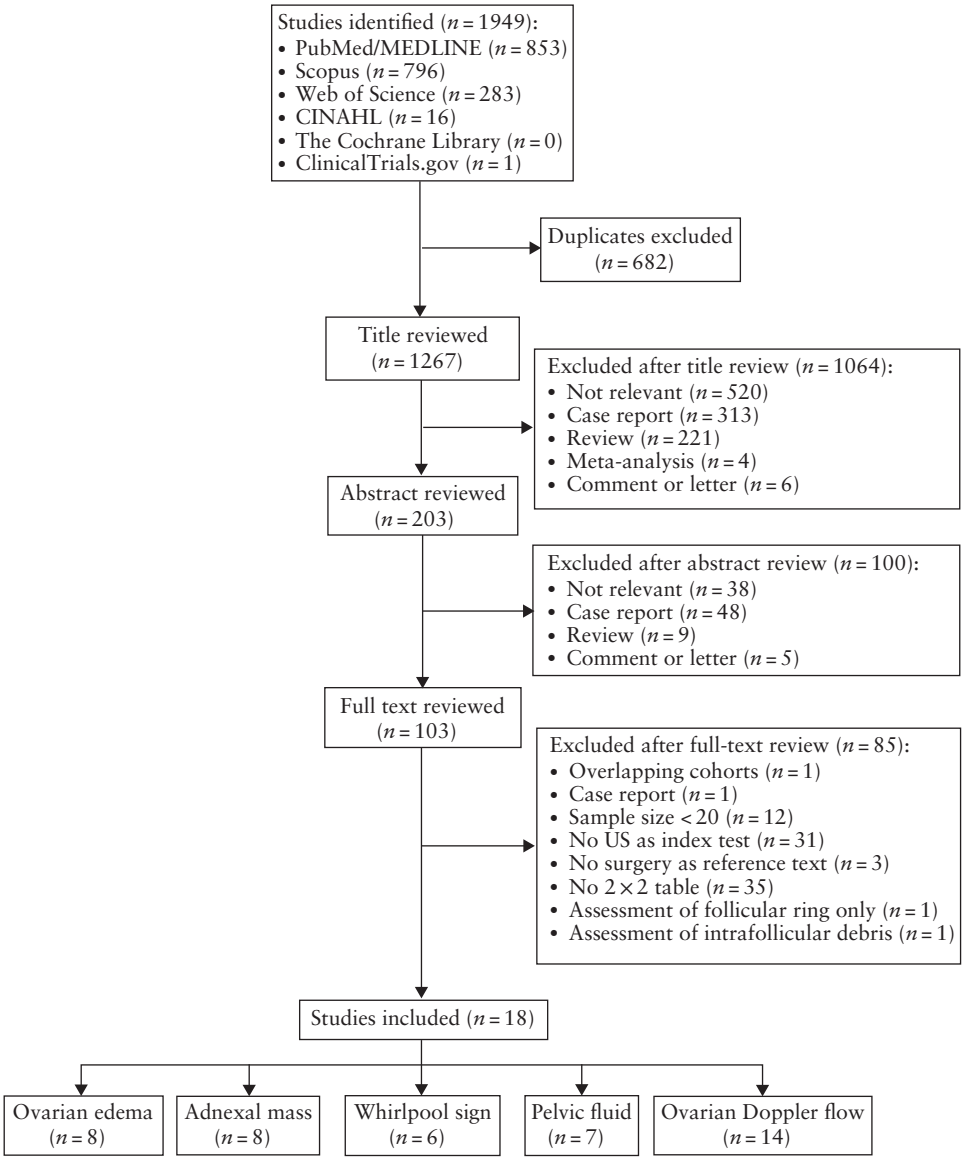


Figure 1 Flowchart summarizing selection of studies evaluating diagnostic accuracy of different ultrasound signs for adnexal torsion.

Table 1 Characteristics of studies included in systematic review and meta-analysis

Study	*Study period*	*Study design*	*Patients* (n)	*Population*	*AT cases* (n)	*Examiners* (n)	*Examiner blinded*	*Time from US to surgery*	*Index test*	*US route*
Carugno (2022)[22]	2014–2018	Retro	63*	Non-pregnant premenopausal women and postmenopausal women	47*	NA	No	Median, 10–16 h	Adnexal mass, Doppler flow	TAS/TVS
Meyer (2022)[30]	2011–2020	Retro	120	Pediatric patients, pregnant and non-pregnant premenopausal women and postmenopausal women	83	NA	Yes	NA	Ovarian edema, adnexal mass, whirlpool sign, Doppler flow, pelvic fluid	TAS/TVS
Meyer (2021)[29]	2011– 2020	Retro	115	Pediatric patients, pregnant and non-pregnant premenopausal women and postmenopausal women	86	NA	NA	NA	Ovarian edema, adnexal mass, whirlpool sign, Doppler flow, pelvic fluid	TAS/TVS
Yatsenko (2021)[18]	NA	Retro	129	Non-pregnant premenopausal and postmenopausal women	106	1	Yes	< 24 h	Ovarian edema, adnexal mass, whirlpool sign, Doppler flow, pelvic fluid	TAS/TVS
Otjen (2020)[17]	2004–2015	Case–control	430	Pediatric patients	99	12	No	NA	Adnexal mass, Doppler flow, pelvic fluid	TAS
Budhram (2019)[21]	2000–2014	Case–control	184†	Pediatric patients, pregnant and non-pregnant premenopausal women and postmenopausal women	92†	NA	No	NA	Ovarian edema, Doppler flow	TAS/TVS
Ghulmiyyah (2019)[23]	2009–2015	Retro	37	Non-pregnant premenopausal or postmenopausal women	10	1	Yes	NA	Ovarian edema, adnexal mass, Doppler flow, pelvic fluid	TAS/TVS
Gu (2018)[24]	2012–2017	Case–control	54	Non-pregnant premenopausal women	28	2	Yes	NA	Whirlpool sign	TAS/TVS
Melcer (2018)[28]	2009–2016	Retro	87	Pediatric patients	53	NA	NA	< 16 h	Ovarian edema, adnexal mass, pelvic fluid	TAS
Rostamzadeh (2014)[32]	2011–2012	Prosp	323	Non-pregnant premenopausal women	43	1	Yes	< 6 h	Doppler flow	TAS
Swenson (2014)[33]	2005–2010	Case–control	40	Non-pregnant premenopausal women and postmenopausal women	15	2	Yes	< 48 h	Doppler flow	TVS
Naiditch (2013)[31]	2007–2011	Retro	113	Pediatric patients	14	12	NA	< 12 h	Doppler flow	TAS
Mashiach (2011)[27]	2002–2008	Retro	63	Pregnant and non-pregnant premenopausal women	47	5	NA	NA	Ovarian edema, adnexal mass, Doppler flow, pelvic fluid	NA
Bar-On (2010)[19]	2006–2008	Retro	77‡	Pregnant and non-pregnant premenopausal women	36‡	NA	NA	< 60 h	Doppler flow	TVS
Valsky (2010)[34]	2006–2009	Retro	80	Pregnant and non-pregnant premenopausal women	18	NA	Yes	< 24 h	Whirlpool sign	TAS/TVS
Linam (2007)[26]	1998–2005	Case–control	74§	Pediatric patients	40§	NA	NA	NA	Ovarian edema, Doppler flow	TAS
Ben-Ami (2002)[20]	NA	Prosp	65	Non-pregnant premenopausal women	15	NA	Yes	NA	Doppler flow	TVS
Lee (1998)[25]	NA	Prosp	47	Pregnant and non-pregnant premenopausal women and postmenopausal women	32	NA	Yes	< 48 h	Whirlpool sign	TAS/TVS

Only first author of each study is given. Reference standard was surgical findings in all studies. *Data for ovarian Doppler flow available from 58 patients (44 with adnexal torsion (AT)). †Data for ovarian Doppler flow available from 166 patients (77 with AT). ‡Data for ovarian Doppler flow available from 40 patients (16 with AT). §Data for ovarian Doppler flow available from 43 patients (24 with AT). NA, not available; Prosp, prospective; Retro, retrospective; TAS, transabdominal sonography; TVS, transvaginal sonography; US, ultrasound.

The series was consecutive in all cohort studies. Three of the studies were prospective in design[20,25,32], 10 studies were retrospective cohort studies[18,19,22,23,27–31,34] and five were retrospective case–control studies[17,21,24,26,33]. Two of the five case–control studies matched the groups by age[21,33], while the remaining studies did not match cases and controls[17,24,26]. Four studies analyzed only pediatric patients[17,26,28,31]. Four studies included both non-pregnant and pregnant women[19,25,27,34], seven studies included only non-pregnant women[18,20,22–24,32,33] and three studies included a mixed cohort of patients (pediatric and non-pregnant/pregnant women)[21,29,30]. Two studies reported by the same group included a different set of patients, with one study including primary cases of adnexal torsion[30] and the other including cases of recurrent adnexal torsion[29]. The number of observers and the information regarding blinding to the reference standard are shown in Table 1.

The ultrasound examination was carried out via the transvaginal or transabdominal route in most studies, depending on the type of study population. In most studies with a mixed population, the proportion of patients explored by these routes was reported poorly or not specified. One study did not describe the route used for ultrasound examination[27]. The length of experience of examiners was reported in seven studies[17,18,24,27,30,31,33]. In all studies, ultrasound examination was performed by an expert examiner. The type of equipment used was reported in eight studies; in all of these studies, the equipment can be considered as high-quality for the time when the study was performed[18,23–26,30,33,34].

Surgical findings were used as the reference standard in all 18 studies. In most studies, surgery was performed laparoscopically. The time between the suspicion of adnexal torsion on ultrasound and surgical intervention was reported in nine studies, varying from 30 min to 60 h[18,19,22,25,28,31–34].

Qualitative synthesis

The results of the evaluation of the risk of bias and concerns regarding applicability of the included studies according to the QUADAS-2 tool are summarized in Table 2. Six studies were considered to have a high risk of patient selection bias because they were case–control studies[17,21,24,26,33] or had inadequate patient exclusion[20].

For the index-test domain, we analyzed the quality according to the ultrasound sign assessed. The definition of each ultrasound sign used by the authors in each study is shown in Table 3. Regarding the ovarian edema sign, two studies were considered to have a high risk of bias because ovarian edema was defined using only a quantitative criterion[21,26]. In two studies, no definition of ovarian edema was provided[23,28]. Regarding the adnexal mass sign, five studies were considered to have a high risk of bias because they reported only on the presence of an ovarian cyst or mass without taking into consideration its size and mass; therefore, physiological ovarian follicles or corpora lutea cannot be ruled out[17,22,23,27,29]. Regarding ovarian Doppler-flow findings, five studies were considered to have a high risk of bias because the criterion reported was not clearly defined (use of imprecise terms such as 'decreased', 'pathological', 'abnormal' or 'positive')[19,21,23,26,31]. For studies assessing the whirlpool sign, all studies were considered to have a low risk of bias, as all of them described correctly this sign[18,24,25,29,30,34]. Regarding the presence of free fluid in the pelvis, all seven studies were considered to have a high risk of bias, as none of them provided an objective definition for this sign[17,18,23,27–30].

For the reference-test domain, all studies were considered to have a low risk of bias because all of them confirmed the presence or absence of adnexal torsion according to surgical findings.

Concerning the flow-and-timing domain, the time elapsed between the index test and reference standard was reported in nine studies[18,19,22,25,28,31–34]. Eight of them were considered as low risk and one as high risk because of a mean time of more than 48 h[19]. The remaining studies were considered as unclear for risk bias, as they did not specify the time interval.

Regarding applicability, all studies were deemed to include patients who matched the review question. For the index-test domain, all studies were considered to have low concerns regarding applicability. Moreover, all studies presented low concerns regarding the reference-standard domain.

Quantitative synthesis

Table 4 summarizes the quantitative synthesis for all five signs assessed and individual forest plots are presented in Figures 2–6, with sROC curves in Figure 7.

Ovarian edema

Eight studies assessed ovarian edema as an ultrasound sign for adnexal torsion[18,21,23,26–30]. All studies were retrospective, and two of them had a case–control design. The studies in this analysis included 809 patients. The mean prevalence of adnexal torsion at surgery was 62% (range, 27–82%).

The pooled sensitivity, specificity, and LR+ and LR– of ovarian edema in the detection of ovarian torsion were 58% (95% CI, 38–76%), 86% (95% CI, 61–96%), 4.0 (95% CI, 1.3–12.6) and 0.49 (95% CI, 0.30–0.79). The DOR was 8 (95% CI, 2–36). Heterogeneity was high for both sensitivity (Cochran's $Q = 144.21$, $P = 0.001$; $I^2 = 95.2\%$) and specificity (Cochran's $Q = 73.87$, $P = 0.001$; $I^2 = 90.5\%$) (Figure 2). Based on metaregression, differences in prevalence among the assessed studies could explain heterogeneity of specificity.

The sROC curve is shown in Figure 7a. The area under the sROC curve was 0.77 (95% CI, 0.66–0.85). The Fagan nomogram showed that a positive result for ovarian edema on ultrasound increased the post-test probability of adnexal torsion from 62% to 87%, while a negative test decreased the post-test probability only

Table 2 Quality assessment of studies included in systematic review and meta-analysis, according to Quality Assessment of Diagnostic Accuracy Studies-2 tool

	Risk of bias								*Applicability concerns*						
		Index test								*Index test*					
Study	*Patient selection*	*Ovarian edema*	*Adnexal mass*	*Whirlpool sign*	*Ovarian Doppler flow*	*Pelvic fluid*	*Reference standard*	*Flow and timing*	*Patient selection*	*Ovarian edema*	*Adnexal mass*	*Whirlpool sign*	*Ovarian Doppler flow*	*Pelvic fluid*	*Reference standard*
Carugno (2022)[22]	Low	—	High	—	Low	—	Low	Low	Low	—	Low	—	Low	—	Low
Meyer (2022)[30]	Low	Low	Low	Low	Low	High	Low	Unclear	Low	Low	Low	Low	Low	Low	Low
Meyer (2021)[29]	Low	Low	High	Low	Low	High	Low	Unclear	Low	Low	Low	Low	Low	Low	Low
Yatsenko (2021)[18]	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low	Low
Otjen (2020)[17]	High	—	High	—	Low	High	Low	Unclear	Low	—	Low	—	Low	Low	Low
Budhram (2019)[21]	High	High	—	—	High	—	Low	Unclear	Low	Low	—	—	Low	—	Low
Ghulmiyyah (2019)[23]	Low	Unclear	High	—	High	High	Low	Unclear	Low	Low	Low	—	Low	Low	Low
Gu (2018)[24]	High	—	—	Low	—		Low	Unclear	Low	—	—	Low	—	—	Low
Melcer (2018)[28]	Low	Unclear	Low	—	—	High	Low	Low	Low	Low	Low	—	—	Low	Low
Rostamzadeh (2014)[32]	Low	—	—	—	Low	—	Low	Low	Low	—	—	—	Low	—	Low
Swenson (2014)[33]	High	—	—	—	Low	—	Low	Low	Low	—	—	—	Low	—	Low
Naiditch (2013)[31]	Low	—	—	—	High	—	Low	Low	Low	—	—	—	Low	—	Low
Mashiach (2011)[27]	Low	Low	High	—	Low	High	Low	Unclear	Low	Low	Low		Low	Low	Low
Bar-On (2010)[19]	Low	—	—	—	High	—	Low	High	Low	—	—	—	Low	—	Low
Valsky (2010)[34]	Low	—	—	Low	—	—	Low	Low	Low	—	—	Low	—	—	Low
Linam (2007)[26]	High	High	—	—	High	—	Low	Unclear	Low	Low	—	—	Low	—	Low
Ben-Ami (2002)[20]	High	—	—	—	Low	—	Low	Unclear	Low	—	—	—	Low	—	Low
Lee (1998)[25]	Low	—	—	Low	—	—	Low	Low	Low	—	—	Low	—	—	Low

Only first author of each study is given.

slightly, from 62% to 44%. No publication bias was observed ($P = 0.52$).

We attempted to perform a subgroup analysis according to the population assessed in the studies. However, this was not possible owing to the small number of studies assessing a specific population. Two studies focused only on pediatric patients[26,28]. Two studies focused only on non-pregnant pre- and/or postmenopausal women[18,23]. One study included a mixed cohort of pregnant and non-pregnant women[27]. In addition, three studies

Table 3 Definitions of ultrasound signs used in studies included in systematic review and meta-analysis

Study	*Ovarian edema*	*Adnexal mass*	*Whirlpool sign*	*Ovarian Doppler flow*	*Pelvic fluid*
Carugno (2022)[22]	—	Ovarian cyst or mass	—	Absent flow	—
Meyer (2022)[30]	Hypoechoic or heterogeneous stroma with small peripheral follicles	Ovarian cyst > 3 cm	Twisted vascular pedicle	Absent flow	Not defined objectively
Meyer (2021)[29]	Hyperechoic or heterogeneous stroma with small peripheral follicles	Adnexal cyst	Twisted vascular pedicle	Absent flow	Not defined objectively
Yatsenko (2021)[18]	Hypoechoic or heterogeneous stroma with small peripheral follicles	Ovarian cyst or mass > 3 cm	Twisted vascular pedicle	Absent venous and/or arterial flow	Not defined objectively
Otjen (2020)[17]	—	Ovarian cyst or mass	—	Absent flow	Not defined objectively
Budhram (2019)[21]	Ovarian maximum diameter of 3 or 5 cm	—	—	Abnormal flow	—
Ghulmiyyah (2019)[23]	Not defined	Ovarian cyst or mass	—	Abnormal flow	Not defined objectively
Gu (2018)[24]	—	—	Twisted vascular pedicle	—	—
Melcer (2018)[28]	Not defined	Ovarian cyst or mass, features described	—	—	Not defined objectively
Rostamzadeh (2014)[32]	—	—	—	Absent venous and/or arterial flow	—
Swenson (2014)[33]	—	—	—	Absent venous and/or arterial flow	—
Naiditch (2013)[31]	—	—	—	Positive or negative	—
Mashiach (2011)[27]	Hypoechoic or heterogeneous stroma with small peripheral follicles	Ovarian cyst or mass	—	Absent venous and/or arterial flow	Not defined objectively
Bar-On (2010)[19]	—	---	—	Pathological or absent flow	—
Valsky (2010)[34]	—	—	Twisted vascular pedicle	—	—
Linam (2007)[26]	Adnexal volume > 20 mL	—	—	Decreased or absent venous flow	—
Ben-Ami (2002)[20]	—	—	—	Absent venous and/or arterial flow	—
Lee (1998)[25]	—	—	Twisted vascular pedicle	—	—

Only first author of each study is given.

Table 4 Diagnostic performance of ultrasound signs for adnexal torsion

Ultrasound sign	*Sensitivity (95% CI) (%)*	*Specificity (95% CI) (%)*	*AUC (95% CI)*	*DOR (95% CI)*
Ovarian edema	58 (38–76)	86 (61–96)	0.77 (0.66–0.85)	8 (2–36)
Adnexal mass	69 (55–81)	46 (22–71)	0.65 (0.52–0.75)	2 (1–5)
Whirlpool sign	65 (12–96)	91 (81–96)	0.92 (0.81–0.97)	20 (2–164)
Ovarian Doppler flow	53 (34–72)	95 (86–98)	0.86 (0.76–0.92)	22 (7–76)
Pelvic fluid	55 (38–71)	69 (54–80)	0.67 (0.54–0.77)	3 (1–6)

AUC, area under the receiver-operating-characteristics curve; DOR, diagnostic odds ratio.

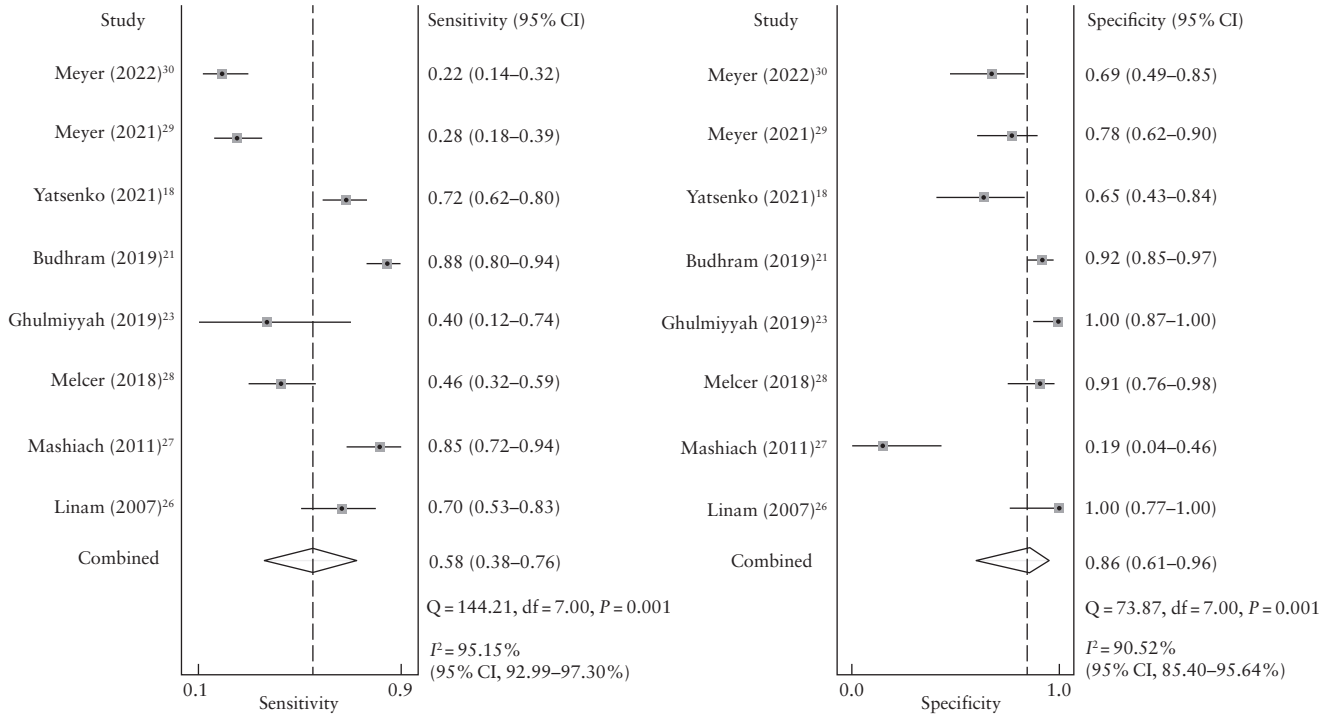


Figure 2 Forest plots of sensitivity and specificity of ovarian edema in the detection of adnexal torsion. Only first author of each study is given.

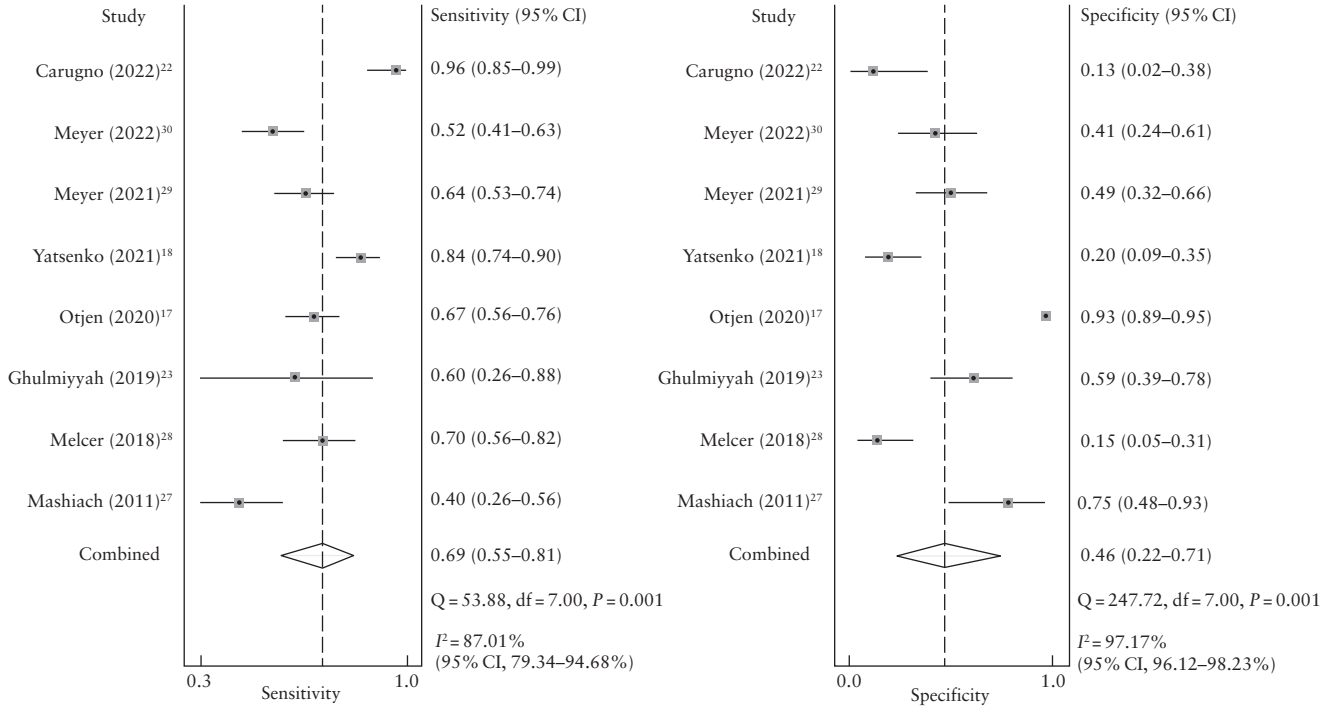


Figure 3 Forest plots of sensitivity and specificity of adnexal mass in the detection of adnexal torsion. Only first author of each study is given.

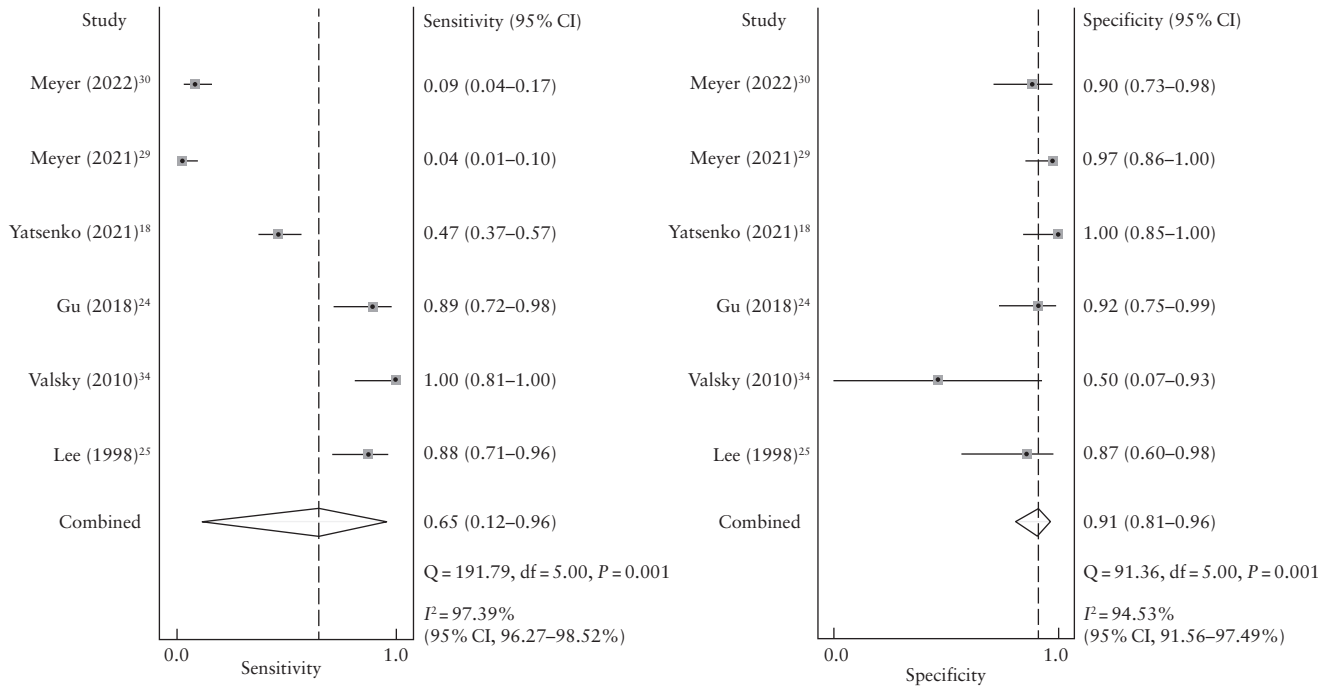


Figure 4 Forest plots of sensitivity and specificity of the whirlpool sign in the detection of adnexal torsion. Only first author of each study is given.

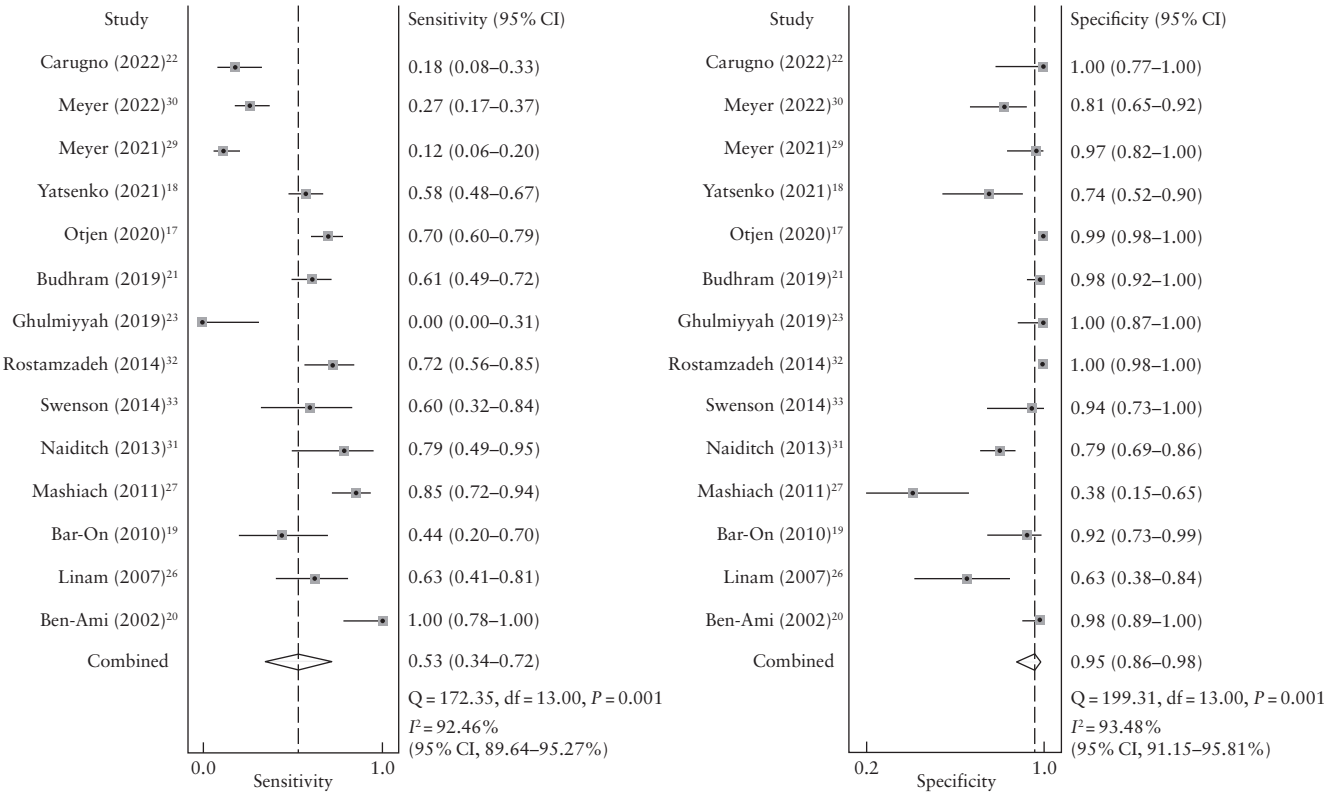


Figure 5 Forest plots of sensitivity and specificity of ovarian Doppler flow in the detection of adnexal torsion. Only first author of each study is given.

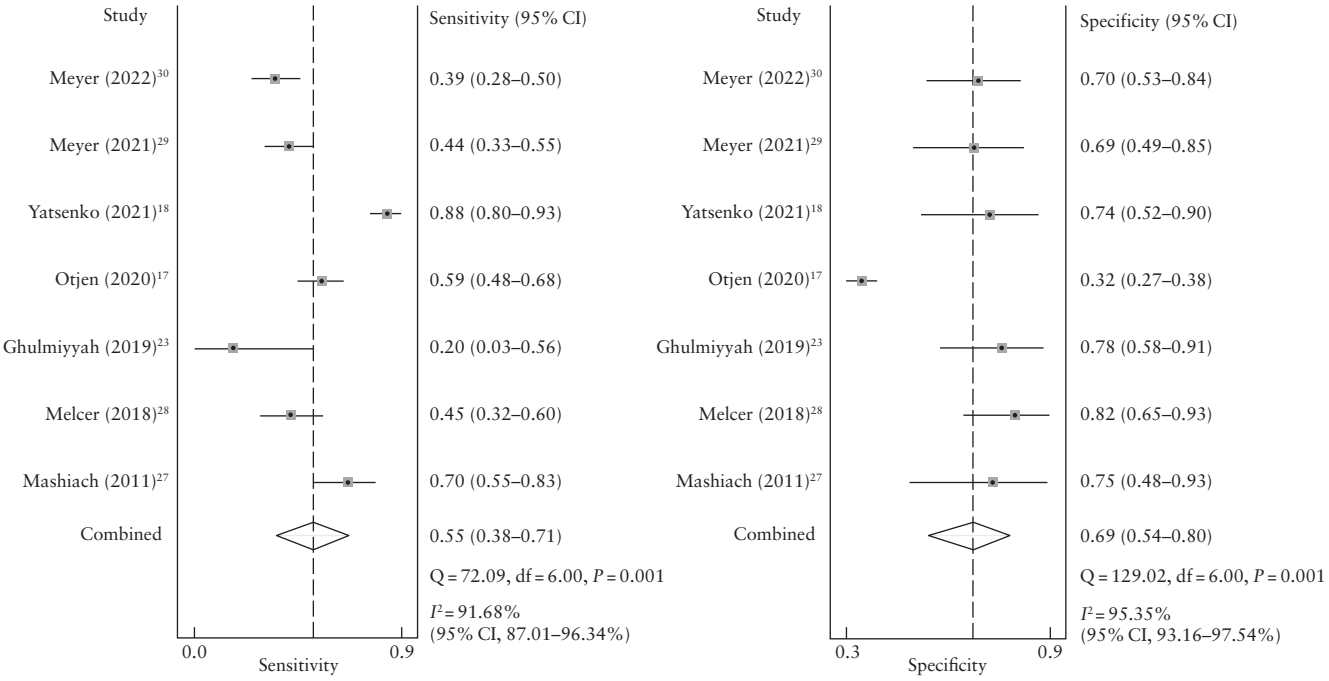


Figure 6 Forest plots of sensitivity and specificity of pelvic fluid in the detection of adnexal torsion. Only first author of each study is given.

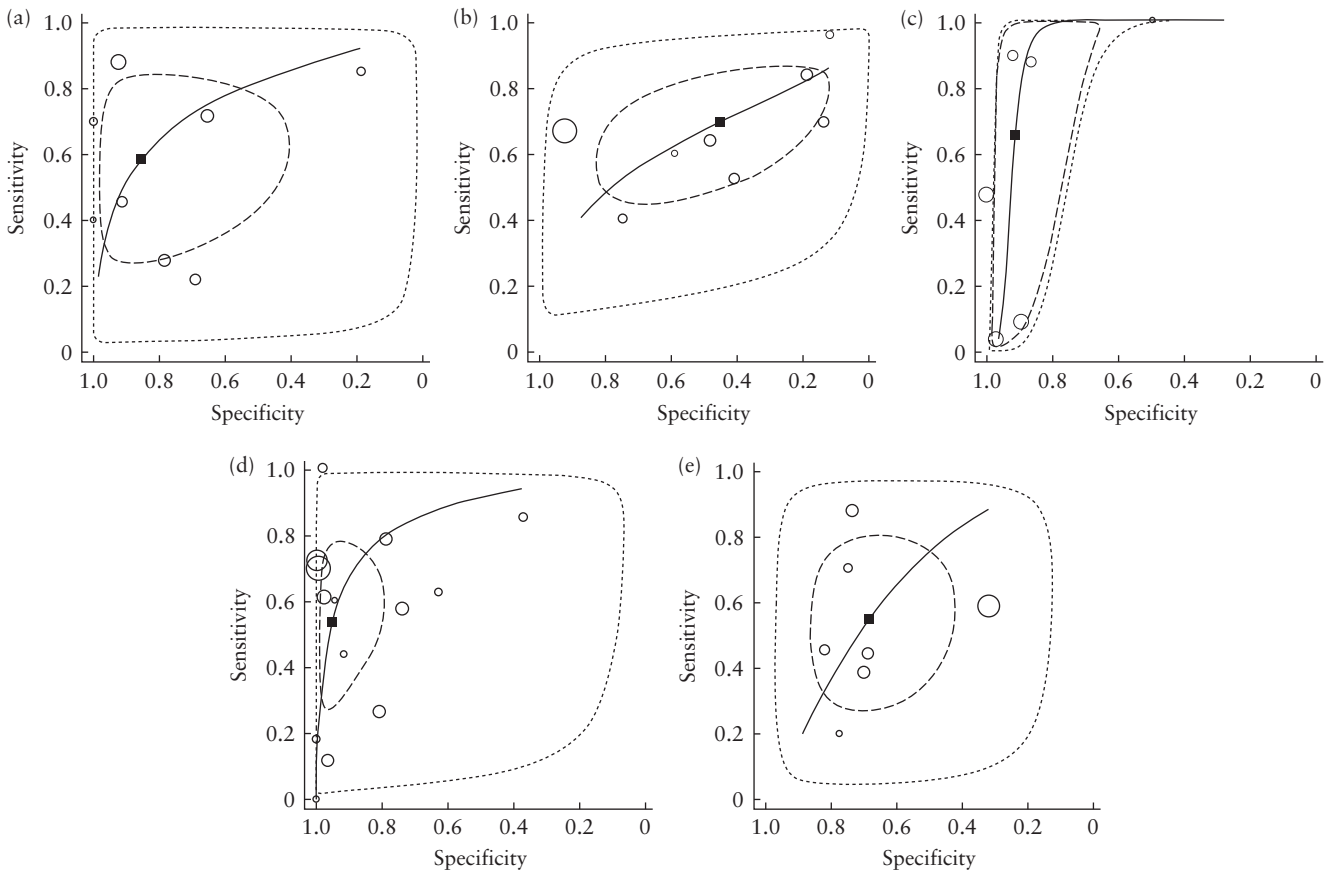


Figure 7 Hierarchical summary receiver-operating-characteristics curves (—) showing diagnostic performance for adnexal torsion of ovarian edema (a), adnexal mass (b), the whirlpool sign (c), ovarian Doppler flow (d) and pelvic fluid (e). ○, study estimate; ■, summary point; – – –, 95% confidence region; ·····, 95% prediction region.

had a mixed cohort of pediatric patients, pregnant and non-pregnant premenopausal women and postmenopausal women[21,29,30].

Adnexal mass

Eight studies assessed the presence of an adnexal mass as an ultrasound sign associated with adnexal torsion[17,18,22,23,27–30]. All studies were retrospective, and one of them had a case–control design. The studies included for this analysis comprised 1044 patients. The mean prevalence of adnexal torsion at surgery was 61% (range, 23–82%).

The pooled sensitivity, specificity, and LR+ and LR− of adnexal mass in the detection of adnexal torsion were 69% (95% CI, 55–81%), 46% (95% CI, 22–71%), 1.3 (95% CI, 0.8–1.9) and 0.67 (95% CI, 0.41–1.10). The DOR was 2 (95% CI, 1–5). Heterogeneity was high for both sensitivity (Cochran's $Q = 53.88$, $P = 0.001$; $I^2 = 87.0\%$) and specificity (Cochran's $Q = 247.72$, $P = 0.001$; $I^2 = 97.2\%$) (Figure 3). Based on metaregression, differences in prevalence among the assessed studies could explain heterogeneity of specificity.

The sROC curve is shown in Figure 7b. The area under the curve was 0.65 (95% CI, 0.52–0.75). The Fagan nomogram showed that a positive result for adnexal mass on ultrasound increased the post-test probability of adnexal torsion from 61% to 67%, while a negative test decreased the post-test probability from 61% to 51%. No publication bias was observed ($P = 0.06$).

We attempted to perform a subgroup analysis according to the population assessed in the studies. However, this was not possible owing to the small number of studies assessing a specific population. Two studies focused only on pediatric patients[17,28]. Three studies focused only on non-pregnant pre- and/or postmenopausal women[18,22,23]. Two studies included a mixed cohort of pregnant and non-pregnant women[19,27]. In addition, two studies included a mixed cohort of pediatric patients, pregnant and non-pregnant premenopausal women and postmenopausal women[29,30].

Whirlpool sign

Six studies assessed the whirlpool sign[18,24,25,29,30,34]. The studies included in this analysis comprised 545 patients. The mean prevalence of adnexal torsion at surgery was 61% (range, 23–82%).

The pooled sensitivity, specificity, and LR+ and LR− of the whirlpool sign in the detection of adnexal torsion were 65% (95% CI, 12–96%), 91% (95% CI, 81–96%), 7.6 (95% CI, 3.8–15.6) and 0.38 (95% CI, 0.07–1.94). The DOR was 20 (95% CI, 2–164). Heterogeneity was high for both sensitivity (Cochran's $Q = 191.79$, $P = 0.001$; $I^2 = 97.4\%$) and specificity (Cochran's $Q = 91.36$, $P = 0.001$; $I^2 = 94.53\%$) (Figure 4). Based on metaregression, heterogeneity could not be explained by any of the variables analyzed.

The sROC curve is shown in Figure 7c. The area under the curve was 0.92 (95% CI, 0.81–0.97). The Fagan nomogram showed that a positive result for the whirlpool sign on ultrasound increased the post-test probability of adnexal torsion from 61% to 92% and a negative result decreased it from 61% to 37%. No publication bias was observed ($P = 0.20$).

We attempted to perform a subgroup analysis according to the population assessed in the studies. However, this was not possible owing to the small number of studies assessing a specific population. Two studies focused only on non-pregnant pre- and/or postmenopausal women[18,24]. Two studies had a mixed cohort of pregnant and non-pregnant women[25,34], and two studies had a mixed cohort of pediatric patients, pregnant and non-pregnant premenopausal women and postmenopausal women[29,30].

Ovarian Doppler flow

Fourteen studies assessed ovarian Doppler findings as an ultrasound sign for diagnosing adnexal torsion[17–23,26,27,29–33]. The studies in this analysis included 1742 patients. The mean prevalence of adnexal torsion at surgery was 47% (range, 12–82%).

The pooled sensitivity, specificity, and LR+ and LR− of ovarian Doppler flow were 53% (95% CI, 34–72%), 95% (95% CI, 86–98%), 11.0 (95% CI, 3.8–31.8) and 0.49 (95% CI, 0.32–0.74). The DOR was 22 (95% CI, 7–76). Heterogeneity was high for both sensitivity (Cochran's $Q = 172.35$, $P = 0.001$; $I^2 = 92.5\%$) and specificity (Cochran's $Q = 199.31$, $P = 0.001$; $I^2 = 93.5\%$) (Figure 5). Based on metaregression, differences in prevalence among the assessed studies could explain heterogeneity of Doppler specificity.

The sROC curve is shown in Figure 7d. The area under the curve was 0.86 (95% CI, 0.76–0.92). The Fagan nomogram showed that a positive result regarding ovarian Doppler flow on ultrasound increased significantly the post-test probability of adnexal torsion from 47% to 91%, while a negative test decreased the post-test probability moderately, from 47% to 30%. Publication bias was observed ($P = 0.02$).

We attempted to perform a subgroup analysis according to the population assessed in the studies. This was possible only for studies focusing on non-pregnant pre- and/or postmenopausal women[18,20,22,23,32,33]. In this population, the diagnostic performance was similar to that in the whole aggregate analysis of ovarian Doppler flow, with pooled sensitivity, specificity, and LR+ and LR− of 51% (95% CI, 14–88%), 99% (95% CI, 89–100%), 35.6 (95% CI, 4.4–289.7) and 0.49 (95% CI, 0.19–1.30). The DOR was 72 (95% CI, 6–896). Heterogeneity was high for both sensitivity (Cochran's $Q = 55.4$, $P = 0.001$; $I^2 = 90.1\%$) and specificity (Cochran's $Q = 52.7$, $P = 0.001$; $I^2 = 90.5\%$).

Pelvic fluid

Seven studies assessed the presence of pelvic fluid[17,18,23,27–30]. The studies included for this analysis comprised 981 patients. The mean prevalence of adnexal torsion at surgery was 59% (range, 23–82%).

The pooled sensitivity, specificity, and LR+ and LR− of the whirlpool sign in the detection of adnexal torsion were 55% (95% CI, 38–71%), 69% (95% CI, 54–80%), 1.7 (95% CI, 1.1–2.9) and 0.66 (95% CI, 0.44–0.99). The DOR was 3 (95% CI, 1–6). Heterogeneity was high for both sensitivity (Cochran's $Q = 72.09$, $P = 0.001$; $I^2 = 91.7\%$) and specificity (Cochran's $Q = 129.02$, $P = 0.001$; $I^2 = 95.4\%$) (Figure 6). Based on metaregression, heterogeneity could not be explained by any of the variables analyzed.

The sROC curve is shown in Figure 7e. The area under the curve was 0.67 (95% CI, 0.54–0.77). The Fagan nomogram showed that a positive result for the whirlpool sign on ultrasound increased the post-test probability of adnexal torsion from 59% to 72% and a negative result decreased it from 59% to 49%. No publication bias was observed ($P = 0.14$).

We attempted to perform a subgroup analysis according to the population assessed in the studies. However, this was not possible owing to the small number of studies assessing a specific population. Two studies focused only on non-pregnant pre- and/or postmenopausal women[18,23]. Two studies focused on a pediatric population[17,28]. One study included a mixed cohort of pregnant and non-pregnant women[27]. In addition, two studies included a mixed cohort of pediatric patients, pregnant and non-pregnant premenopausal women and postmenopausal women[29,30].

DISCUSSION

Summary evidence

According to our results, ovarian edema, the whirlpool sign and ovarian Doppler flow are ultrasound signs with a high specificity but moderate sensitivity for the diagnosis of adnexal torsion. The presence of an adnexal mass or pelvic fluid have poor diagnostic performance. However, it should be borne in mind that objective diagnostic criteria for ovarian edema, adnexal mass, pelvic fluid and ovarian Doppler flow were not stated clearly in many studies.

Interpretation of results

An accurate diagnosis is essential for the optimal management of women with a clinical suspicion of adnexal torsion. A delayed or false-negative diagnosis may lead to ovarian necrosis, while a false-positive diagnosis may lead to unnecessary surgical intervention with potential complications[35].

We have observed that ovarian edema, the whirlpool sign and ovarian Doppler findings show good specificity for diagnosing adnexal torsion. However, the sensitivity of these signs is moderate. Therefore, these signs should be assessed in every patient with a clinical suspicion of adnexal torsion. The presence of an adnexal mass and pelvic fluid may be a potential source of false-positive cases, and these signs should be interpreted while taking other ultrasound signs into consideration. Some studies have shown that combining more than one sign may improve the diagnostic performance[18,24].

The qualitative synthesis raises some concerns regarding the quality of the included studies because many included a heterogeneous population, did not provide a clear definition of the index test and mixed data obtained by transabdominal sonography and transvaginal sonography.

Comparison with previous literature

To date, three meta-analyses about the ultrasound diagnosis of adnexal torsion have been published[7–9]. Bronstein *et al.*[7] performed a meta-analysis assessing the role of B-mode ultrasound, Doppler ultrasound and computed tomography (CT) scan for diagnosing adnexal torsion in a pediatric population, including 18 studies using B-mode findings and 15 studies using Doppler. However, they did not perform a qualitative analysis of the studies, and not all studies included used surgical findings as a reference test. In fact, only three studies assessing B-mode findings and four studies assessing Doppler ultrasound findings were used to estimate pooled specificity. Furthermore, for B-mode ultrasound, no specific ultrasound sign was evaluated.

Adu-Bredu *et al.*[8] performed a meta-analysis including eight studies assessing only the whirlpool sign. Although the authors stated that a qualitative assessment was performed, data from the analysis were not reported. Furthermore, in the quantitative synthesis, the authors included six studies from which specificity could not be estimated because all patients included had adnexal torsion or because the reference standard was not surgery.

Wattar *et al.*[9] performed a meta-analysis assessing the diagnostic performance of ultrasound, CT and magnetic resonance imaging (MRI) for diagnosing adnexal torsion, including 12 studies assessing ultrasound findings. In this meta-analysis, the reference standard used in the included studies was both surgical findings and clinical follow-up. The inclusion of studies using clinical follow-up as the reference standard may pose a risk of bias because spontaneous detorsion can occur and cases with adnexal torsion may be classified as true-negative cases. Additionally, no specific analysis of different ultrasound signs was done. Qualitative synthesis was performed only for case–control studies and the data were not reported separately for studies using ultrasound, CT and MRI.

Three meta-analyses reported data on the diagnostic performance of other imaging techniques, including CT and MRI[7,9,36]. Two meta-analyses reported the overall diagnostic performance of these techniques but did not analyze specific signs[7,9], and one meta-analysis reported on the pooled proportion of different signs present in

cases of adnexal torsion using CT but did not assess the diagnostic performance of these signs[36]. Therefore, we cannot compare our data with those reported in these meta-analyses.

Strengths and limitations

We consider that the main strength of our meta-analysis is that it is the first one to perform separate quantitative and qualitative analyses on the diagnostic performance of multiple ultrasound signs.

A common problem with studies assessing imaging for adnexal torsion is that not all suspicious cases undergo surgery, which can lead to ascertainment bias. This may affect specificity of the test, which may be overestimated. However, all women in our meta-analysis underwent surgery, which can be considered a strength of this study.

The main limitation of this meta-analysis is the small number of studies and patients included. In addition, we could not assess the diagnostic performance of the different ultrasound signs in different populations. We observed that there were few objective, quantifiable and reproducible criteria available to reach an ultrasound diagnosis of adnexal torsion with high certainty. This is why we believe that it seems difficult to propose the development of a clinical guide for action in the face of this condition. From the methodological point of view, we assumed that a sign was negative in cases in which the authors did not specifically mention that it had been visualized. This assumption could be erroneous and potentially lead to our underestimating sensitivity.

Clinical significance of findings

The diagnostic performance in this systematic review of ultrasound signs for adnexal torsion are rather disappointing. The poor-to-moderate diagnostic performance may be explained by the significant heterogeneity found among the 18 studies ultimately selected. In fact, we could not assess the impact of factors such as quality of ultrasound machine used (owing to a significant range of the year of publication of the assessed studies), experience of the examiner, route of the ultrasound examination (transvaginal *vs* transabdominal) and the population studied (girls, adolescents, non-pregnant premenopausal women, pregnant women and postmenopausal women). All these factors may affect the performance of ultrasound in the diagnosis of adnexal torsion.

Many clinicians rely on ultrasound as an imaging technique when evaluating women with suspected adnexal torsion. However, we found that the sensitivity of all signs assessed in our meta-analysis was moderate at best. This fact means that false-negative cases are frequent, and this is relevant when we consider the consequences of adnexal torsion (loss of the adnexa). We believe that our findings should prompt the development of a scoring system combining several clinical features and ultrasound findings in an attempt to improve the diagnostic performance for adnexal torsion.

Conclusions

Ovarian edema, the whirlpool sign and absent intraovarian blood flow as assessed by Doppler ultrasound are highly specific sonographic signs for diagnosing adnexal torsion. The presence of an adnexal mass or pelvic fluid have moderate diagnostic performance for detecting adnexal torsion. However, the quality of the evidence is limited. Future research is needed to improve the performance of ultrasound in diagnosing adnexal torsion.

REFERENCES

1. Ssi-Yan-Kai G, Rivain AL, Trichot C, Morcelet MC, Prevot S, Deffieux X, De Laveaucoupet J. What every radiologist should know about adnexal torsion. *Emerg Radiol* 2018; **25**: 51–59.
2. Sasaki KJ, Miller CE. Adnexal torsion: review of the literature. *J Minim Invasive Gynecol* 2014; **21**: 196–202.
3. Huchon C, Fauconnier A. Adnexal torsion: a literature review. *Eur J Obstet Gynecol Reprod Biol* 2010; **15**: 8–12.
4. Fleischer AC, Stein SM, Cullinan JA, Warner MA. Color Doppler sonography of adnexal torsion. *J Ultrasound Med* 1995; **14**: 523–528.
5. Chang HC, Bhatt S, Dogra VS. Pearls and pitfalls in diagnosis of ovarian torsion. *Radiographics* 2008; **28**: 1355–1368.
6. Adnexal Torsion in Adolescents: ACOG Committee Opinion No, 783. *Obstet Gynecol* 2019; **134**: e56–e63.
7. Bronstein ME, Pandya S, Snyder CW, Shi Q, Muensterer OJ. A meta-analysis of B-mode ultrasound, Doppler ultrasound, and computed tomography to diagnose pediatric ovarian torsion. *Eur J Pediatr Surg* 2015; **25**: 82–86.
8. Adu-Bredu TK, Arkorful J, Appiah-Denkyira K, Wiafe YA. *J Clin Ultrasound* 2021; **49**: 746–753. Diagnostic value of the sonographic whirlpool sign in the diagnosis of ovarian torsion: A systematic review and meta-analysis.
9. Wattar B, Rimmer M, Rogozinska E, Macmillian M, Khan KS, Al Wattar BH. Accuracy of imaging modalities for adnexal torsion: a systematic review and meta-analysis. *BJOG* 2021; **128**: 37–44.
10. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analysis: the PRISMA statement. *BMJ* 2009; **339**: b2535.
11. Sotiriadis A, Papatheodorou SI, Martins WP. Synthesizing Evidence from Diagnostic Accuracy Tests: the SEDATE guideline. *Ultrasound Obstet Gynecol* 2016; **47**: 386–395.
12. Moro F, Bolomini G, Sibal M, Vijayaraghavan SB, Venkatesh P, Nardelli F, Pasciuto T, Mascilini F, Pozzati F, Leone FPG, Josefsson H, Epstein E, Guerriero S, Scambia G, Valentin L, Testa AC. Imaging in gynecological disease (20): clinical and ultrasound characteristics of adnexal torsion. *Ultrasound Obstet Gynecol* 2020; **56**: 934–943.
13. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; **155**: 529–536.
14. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557–560.
15. Kiechl-Kohlendorfer U, Maurer K, Unsinn KM, Gassner I. Fluid-debris level in follicular cysts: a pathognomonic sign of ovarian torsion. *Pediatr Radiol* 2006; **36**: 421–425.
16. Sibal M. Follicular ring sign: a simple sonographic sign for early diagnosis of ovarian torsion. *J Ultrasound Med* 2012; **31**: 1803–1809.
17. Otjen JP, Stanescu AL, Alessio AM, Parisi MT. Ovarian torsion: developing a machine-learned algorithm for diagnosis. *Pediatr Radiol* 2020; **50**: 706–714.
18. Yatsenko O, Vlachou PA, Glanc P. Predictive Value of Single or Combined Ultrasound Signs in the Diagnosis of Ovarian Torsion. *J Ultrasound Med* 2021; **40**: 1163–1172.
19. Bar-On S, Mashiach R, Stockheim D, Soriano D, Goldenberg M, Schiff E, Seidman DS. Emergency laparoscopy for suspected ovarian torsion: are we too hasty to operate? *Fertil Steril* 2010; **93**: 2012–2015.
20. Ben-Ami M, Perlitz Y, Haddad S. The effectiveness of spectral and color Doppler in predicting ovarian torsion. A prospective study. *Eur J Obstet Gynecol Reprod Biol* 2002; **104**: 64–66.
21. Budhram G, Elia T, Dan J, Schroeder M, Safain G, Schlech W, Friderici J, Knee A, Anthouard M, Schoenfeld E. A Case–control Study of Sonographic Maximum Ovarian Diameter as a Predictor of Ovarian Torsion in Emergency Department Females with Pelvic Pain. *Acad Emerg Med* 2019; **26**: 152–159.
22. Carugno J, Naem A, Ibrahim C, Ehinger N, Moore J, Garzon S, Laganà AS. Is color Doppler ultrasonography reliable in diagnosing adnexal torsion? A large cohort analysis. *Minim Invasive Ther Allied Technol* 2022; **31**: 620–627.
23. Ghulmiyyah L, Nassar A, Sassine D, Khoury S, Nassif J, Ramadan H, Najem E, Berjawi G. Accuracy of Pelvic Ultrasound in Diagnosing Adnexal Torsion. *Radiol Res Pract* 2019; **2019**: 1406291.
24. Gu X, Yang M, Liu Y, Liu F, Liu D, Shi F. The ultrasonic whirlpool sign combined with plasma d-dimer level in adnexal torsion. *Eur J Radiol* 2018; **109**: 196–202.

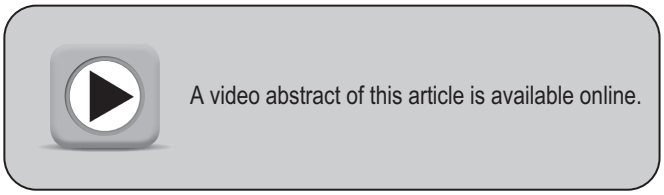
25. Lee EJ, Kwon HC, Joo HJ, Suh JH, Fleischer AC. Diagnosis of ovarian torsion with color Doppler sonography: depiction of twisted vascular pedicle. *J Ultrasound Med* 1998; **17**: 83–89.
26. Linam LE, Darolia R, Naffaa LN, Breech LL, O'hara SM, Hillard PJ, Huppert JS. US findings of adnexal torsion in children and adolescents: size really does matter. *Pediatr Radiol* 2007; **37**: 1013–1019.
27. Mashiach R, Melamed N, Gilad N, Ben-Shitrit G, Meizner I. Sonographic diagnosis of ovarian torsion: accuracy and predictive factors. *J Ultrasound Med* 2011; **30**: 1205–1210.
28. Melcer Y, Maymon R, Pekar-Zlotin M, Pansky M, Smorgick N. Clinical and sonographic predictors of adnexal torsion in pediatric and adolescent patients. *J Pediatr Surg* 2018; **53**: 1396–1398.
29. Meyer R, Meller N, Mohr-Sasson A, Abu-Bandora E, Cohen A, Tamir M, Mashiach R, Levin G. Prediction score for recurrent adnexal torsion in women with a previous adnexal torsion. *Int J Gynaecol Obstet* 2021; **155**: 411–416.
30. Meyer R, Meller N, Mohr-Sasson A, Toussia-Cohen S, Komem DA, Mashiach R, Levin G. A clinical prediction model for adnexal torsion in pediatric and adolescent population. *J Pediatr Surg* 2022; **57**: 497–501.
31. Naiditch JA, Barsness KA. The positive and negative predictive value of transabdominal color Doppler ultrasound for diagnosing ovarian torsion in pediatric patients. *J Pediatr Surg* 2013; **48**: 1283–1287.
32. Rostamzadeh A, Mirfendereski S, Rezaie MJ, Rezaei S. Diagnostic efficacy of sonography for diagnosis of ovarian torsion. *Pak J Med Sci* 2014; **30**: 413–416.
33. Swenson DW, Lourenco AP, Beaudoin FL, Grand DJ, Killelea AG, McGregor AJ. Ovarian torsion: Case–control study comparing the sensitivity and specificity of ultrasonography and computed tomography for diagnosis in the emergency department. *Eur J Radiol* 2014; **83**: 733–738.
34. Valsky DV, Esh-Broder E, Cohen SM, Lipschuetz M, Yagel S. Added value of the gray-scale whirlpool sign in the diagnosis of adnexal torsion. *Ultrasound Obstet Gynecol* 2010; **36**: 630–634.
35. Kives S, Gascon S, Dubuc É, Van Eyk N. No. 341-Diagnosis and Management of Adnexal Torsion in Children, Adolescents, and Adults. *J Obstet Gynaecol Can* 2017; **39**: 82–90.
36. Ling-Shan C, Jing L, Zheng-Qiu Z, Pin W, Zhi-Tao W, Fu-Ting T, Xu-Yu H, Zhong-Qiu W. Computed Tomography Features of Adnexal Torsion: A Meta-Analysis. *Acad Radiol* 2022; **29**: 317–325.

SUPPORTING INFORMATION ON THE INTERNET

The following supporting information may be found in the online version of this article:

Appendix S1 Protocol template: Systematic Review

Appendix S2 Papers excluded after reading full text



Ultrasound Obstet Gynecol 2023; **61**: 310–324
Published online 11 February 2023 in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.24976.

Precisión diagnóstica de indicios ecográficos para detectar la torsión anexial: revisión sistemática y metaanálisis

RESUMEN

Objetivo. Evaluar la precisión diagnóstica de diferentes indicios ecográficos para el diagnóstico de la torsión anexial, utilizando la cirugía como patrón de referencia.

Métodos. Se trata de una revisión sistemática y un metaanálisis de los estudios publicados entre enero de 1990 y noviembre de 2021 que evaluaron el edema ovárico, la masa anexial, los resultados del flujo Doppler ovárico, el "signo del remolino" (torsión ovárica) y el líquido pélvico como indicios ecográficos (pruebas de referencia) para detectar la torsión anexial, utilizando los resultados quirúrgicos como patrón de referencia. La búsqueda de estudios se realizó en las bases de datos PubMed/MEDLINE, CINAHL, Scopus, The Cochrane Library, ClinicalTrials.gov y Web of Science. Para evaluar la calidad de los estudios se utilizó la herramienta de Evaluación de Calidad de los Estudios de Precisión Diagnóstica-2 (QUADAS-2, por sus siglas en inglés). Se calcularon por separado los valores combinados de sensibilidad, especificidad y cocientes de verosimilitud positivos y negativos, y se determinó también la probabilidad posterior a la prueba de la torsión anexial tras una prueba positiva o negativa.

Resultados. La búsqueda identificó 1267 citas, una vez excluidas las duplicadas. Se incluyeron dieciocho estudios en las síntesis cualitativa y cuantitativa. Ocho estudios (809 pacientes) analizaron la presencia de edema ovárico, ocho estudios (1044 pacientes) analizaron la presencia de una masa anexial, 14 estudios (1742 pacientes) analizaron el flujo Doppler ovárico, seis estudios (545 pacientes) analizaron el "signo del remolino" y siete estudios (981 pacientes) analizaron la presencia de líquido pélvico como indicios ecográficos de torsión anexial. En general, la calidad de la mayoría de los estudios se consideró moderada o buena. Sin embargo, hubo un alto riesgo de sesgo en una proporción significativa de estudios en las áreas de selección de pacientes y texto de referencia (con la excepción del "signo del remolino"). Los valores combinados de la sensibilidad, la especificidad y los cocientes de probabilidad positivos y negativos de cada indicio ecográfico fueron del 58%, 86%, 4,0 y 0,49 para el edema ovárico, del 69%, 46%, 1,3 y 0,67 para la masa anexial, del 65%, 91%, 7,6 y 0,38 para el "signo del remolino", del 53%, 95%, 11,0 y 0,49 para los resultados del Doppler ovárico y del 55%, 69%, 1,7 y 0,66 para el líquido pélvico. La heterogeneidad fue elevada en todos los análisis.

Conclusión. La presencia de una masa anexial o de líquido pélvico apenas tienen precisión diagnóstica como indicios ecográficos de torsión anexial, mientras que la presencia de edema ovárico, el "signo del remolino" y la disminución o ausencia de flujo Doppler ovárico tienen una buena especificidad, pero una sensibilidad moderada para detectar la torsión anexial.

检测附件扭转的超声征象的诊断准确性：系统回顾和荟萃分析

摘要

目的 以手术为参考标准，评价不同超声征象诊断附件扭转的准确性。

方法 这是对1990年1月至2021年11月期间发表的研究的系统回顾和荟萃分析，评价卵巢水肿、附件肿块、卵巢多普勒血流结果、漩涡征和盆腔积液作为检测附件扭转的超声征象（指标测试），以手术结果作为参考标准。在PubMed/MEDLINE、CINAHL、Scopus、The Cochrane Library、ClinicalTrials.gov和Web of Science数据库中进行研究搜索。使用诊断准确性研究的质量评价工具-2（QUADAS-2）来评价这些研究的质量。分别计算了汇总灵敏度、特异性以及阳性和阴性的似然比，还确定了测试阳性或阴性后附件扭转的概率。

结果 在排除重复文献后，搜索确定了1267篇引文。最终18项研究被纳入定性和定量综合分析。其中8项研究（809名患者）分析了卵巢水肿，8项研究（1044名患者）分析了附件肿块，14项研究（1742名患者）分析了卵巢多普勒血流，6项研究（545名患者）分析了漩涡征，7项研究（981名患者）分析了盆腔液体，作为附件扭转的超声征象。总的来说，大多数研究的质量被认为属于中等或良好。然而，在相当一部分研究中，患者选择和索引文本方面（漩涡征除外）存在较高的偏倚风险。每个超声征象的汇合敏感性、特异性以及阳性和阴性似然比分别如下：卵巢水肿58%、86%、4.0和0.49，附件肿块69%、46%、1.3和0.67，漩涡征65%、91%、7.6和0.38，卵巢多普勒血流53%、95%、11.0和0.49，盆腔积液55%、69%、1.7和0.66。所有分析的异质性都很高。

结论 附件肿块或盆腔积液作为附件扭转的超声征象诊断准确性较差，而卵巢水肿、漩涡征和卵巢多普勒血流减少或消失对检测附件扭转有良好的特异性，但敏感性一般。