Accuracy of imaging modalities for adnexal torsion: A systematic review and meta-analysis

Baihas Wattar[1]; Michael Rimmer[2]; Ewelina Rogozinska[3]; Mark Macmillian[4]; Khalid S. Khan[5]; Bassel H. Al Wattar[6]

[1]West Suffolk NHS Foundation Trust, Bury St Edmunds, Suffolk, UK

[2]MRC Centre for Reproductive Health, University of Edinburgh, Edinburgh, UK

[3]MRC Clinical Trials Unit, University College London, London, UK

[4]MRC Centre for Regenerative Medicine, The University of Edinburgh, Edinburgh UK

[5]Department of Preventive Medicine and Public Health, University of Granada, 18071 Granada, Spain

[6]Warwick Medical School, University of Warwick, Coventry, UK

Corresponding author: Bassel H. Al Wattar, Warwick Medical School, University of Warwick. Email: dr.basselwa@gmail.com

Short title: Imaging modalities for adnexal torsion

Abstract:

Background: Adnexal torsion (AT), a serious gynaecological emergency, often presents with non-specific symptoms leading to delayed diagnosis.

Objective: To compare the test accuracy of ultrasound (USS), computerized tomography (CT), and magnetic resonance (MRI) to diagnose AT.

Search Strategy: We searched EMBASE, MEDLINE, and Cochrane CENTRAL until December 2019.

Selection criteria: Studies reporting on the accuracy of any imaging modality (Index Test) in females (paediatric and adults) suspected of AT compared to surgical diagnosis and/or standard clinical/radiological follow-up period until resolution of symptoms (Reference Standard).

Data collection and Analysis: We assessed study quality using QUADAS-2. We conducted test accuracy meta-analysis using a univariate model or a hierarchical model.

Main Results: We screened 3836 citations, included 18 studies (1654 women, 665 cases), and 15 in the meta-analyses. USS pooled sensitivity (n=12, 1187 women) was 0.79 (95%CI 0.63–0.92) and specificity was 0.76 (95%CI 0.54–0.93), with a negative and positive likelihood ratio of 0.29 (95%CI 0.13-0.66) and 4.35 (95%CI 2.03-9.32) respectively. Using Doppler with USS (n=7, 845 women) yielded similar sensitivity (0.80, 95%CI 0.67-0.93) and specificity (0.88, 95%CI 0.72-1.00). For MRI (n=3, 99 women), the pooled sensitivity was 0.81 (95%CI 0.63-0.91) and specificity was 0.91 (95%CI 0.80-0.96). A meta-analysis for CT was not possible with two case-control and one cohort studies (n=3, 232 women). Its sensitivity range was 0.74-0.95, and specificity was 0.80-0.90.

Conclusions: Ultrasound has good performance as a first-line diagnostic test for suspected AT. Magnetic resonance could offer improved specificity to investigate complex ovarian morphology, but more evidence is needed.

Funding: None

Keywords: Ovary, adnexa, torsion, ultrasound, Doppler, magnetic resonance, computerized tomography, test accuracy, meta-analysis

Tweetable abstract: To investigate adnexal torsion, ultrasound is a good first-line diagnostic test with a pooled sensitivity of 0.79 and specificity of 0.76.

Introduction:

Adnexal torsion (AT) is a serious gynaecological emergency which involves a partial or complete twisting of the infundibulopelvic vascular pedicle. It acutely compromises the vascular supply of the ovary and the adjunct fallopian tube eliciting ischemia, tissue necrosis, reduced ovarian follicular reserve, subfertility and early menopause (1). While its prevalence is unclear, it is estimated to affect 2–7% of women undergoing surgery for acute pelvic pain (2). Most affected women present with non-specific symptoms such as abdominal pain, vomiting and fever leading to delayed diagnosis and increased risk of emergency oophorectomy (3). As such, establishing a prompt diagnosis is key to enable early surgical untwisting and restoration of the compromised vascular supply.

To aid its diagnosis, numerous imaging modalities have been used and evaluated in the literature(4). Ultrasound (USS) is commonly used to evaluate ovarian pathology due to its safety, availability and affordability. However, several factors could limit its accuracy to diagnose AT such as operators experience, machine quality, pregnancy and presence of complex ovarian morphology (5). Doppler is often used to highlight the compromised vascular supply to the adnexa, however, its added diagnostic value remains imprecise (4). Both Computerized Tomography (CT) and Magnetic Resonance Imaging (MRI) have been used to evaluate complex ovarian morphology, however, their use to diagnose AT could be hampered by the variations in diagnostic criteria and the experience of the assessor (4). Test accuracy for these modalities is not precisely known thus, increasing variations in practice and hindering effective policymaking (3).

We aimed to compare the test accuracy of the various imaging modalities used to diagnose AT by conducting a systematic review and meta-analysis.

Methods:

We conducted a systematic review using an established methodology for test accuracy research(6) and a prospectively registered protocol (CRD42018112048). We reported findings of our review as per established guidelines (7). Patients were not involved in the design and conduct of this review. We searched the COMET database and did not identify any relevant core outcome sets on the topic of interest.

Literature search

We searched the major electronic databases (EMBASE, MEDLINE, and Cochrane CENTRAL) for primary diagnostic accuracy studies for adnexal torsion from inception until December 2019. We performed complementary searches in ClinicalTrials.gov, Google Scholar and Scopus to capture any relevant additional citations. We did not employ any search filters or language restrictions. We used MeSH terms (ovarian, ovary, tube, fallopian, twisted, torsion, adnexa, adnexal, adnexa) and combined them using the Boolean operators AND/OR to produce a sensitive search (Appendix 1). We searched the bibliographies of potentially relevant articles to identify any additional citations not captured by our search.

Study selection and data extraction

We performed the study selection and inclusion process in two stages. First, two reviewers (BW and MPR) screened the titles and abstracts of potentially relevant articles. In the second stage we assessed relevant articles in full against our inclusion criteria before inclusion. We

included all primary studies reporting on the diagnostic accuracy of any imaging modality (Index Test) used in females (paediatric and adults) presenting with symptoms suggestive of AT (acute/sub-acute abdominal/pelvic pain, fever, nausea, vomiting, pelvic mass) compared to surgical diagnosis and/or standard clinical/radiological follow-up period until resolution of symptoms (Reference Standard) in no preferential order. We excluded studies reporting only on foetal/neonatal adnexal torsion or on isolated tubal torsion. We also excluded reviews, case reports and case series. Studies that identified their population by 'asymptomatic ovarian mass' were also excluded as this can overestimate the diagnostic accuracy. Any disagreements were resolved in consensus with a third reviewer (BHA). Studies that were of case-control design were included in our systematic review but not in the meta-analysis (8).

We extracted data in duplicate onto a piloted electronic data extraction sheet. We collected data on population characteristics, description of the index and reference tests, used diagnostic criteria, treatment algorithm in each study, and the duration of follow-up.

Quality assessment of included studies

Two reviewers (BW and MPR) independently assessed the risk of bias and applicability of the included studies using the QUADAS-2 (12) in four domains: patient selection, conduct of the index test, conduct of the reference standard, and patient flow. We considered a study to be of high quality if it used a patient spectrum matching the review question, enrolled a consecutive or random sample of patients, used the index test as first-line imaging with a pre-defined benchmark for a positive test, all participants had surgical confirmation within 48 hours as reference standard, and the majority of recruited participants were included in analyses. The following were considered to be inappropriate patient spectrums that

introduced bias: cohorts limited to only paediatric, pregnant or non-pregnant women, studies involving women with asymptomatic pelvic mass, and those with inappropriate exclusions. Lack of blinding to index test results upon the interpretation of the results of the reference standard was not considered to pose a high risk of bias.

Data synthesis

We constructed 2×2 tables for each imaging modality and calculated sensitivity, specificity and likelihood ratios for positive and negative test results with 95% confidence intervals (CIs). We pooled the accuracy parameters using a hierarchical model (random effect) when a sufficient number of studies (>4) were available (9). When fewer than four studies were available, we used a univariate model (10). We investigated heterogeneity visually from forest plots of sensitivity and specificity estimates. We considered the use of Doppler to be a potential effect-modifier in studies evaluating the use of USS and investigated it using a meta-regression. We performed subgroup analyses to evaluate the effect of potential confounders (e.g population age, pubertal status etc..). We did not assess the publication bias due to the small number of studies included for each imaging modality. We conducted our analysis using RevMan version 5.3, Open Mata-analyst software version 12.11.14, and Stata version 14 (StataCorp, College Station, TX, 2015).

Funding

No funding received directly to support this work.

Results:

Characteristics of included studies

We identified 3836 potentially relevant citations, of these 124 were reviewed in full against our inclusion criteria and 18 were included reporting on 1654 women (Figure 1). Most studies (15/18, 83%) were cohorts (14 retrospectives and one prospective) while three were retrospective case-controls (3/18, 17%), one reporting on CT, one on USS and one on USS and CT. The median sample size was 71 (range 29-323) with 665 confirmed cases of AT (665/1654, 40%). There were four studies from each of the United States of America and Israel (4/18, 22%), three from Korea (3/18, 17%), two from France (2/18, 11%) and one from each of India, Iran, China, Canada, and Saudi Arabia (Table S1). Two-thirds of studies used surgical exploration as the Reference Standard (12/18, 67%), while six used a mixture of surgical exploration and clinical follow-up (6/18, 33%). Only three studies reported on each of CT(11–13) and MRI(14–16) (3/18, 17%). Fourteen studies reported on the accuracy of USS (14/18, 44%), of these nine included the use of Doppler (9/14, 64%) and five included adults only (5/14, 36%) while the remaining included a mixture of paediatric and adults or did not report on it. Ten USS studies only used surgical exploration as a Reference test (10/14, 71%) while the remaining four used a mixture of surgical and clinical follow-up.

Quality of included studies

The overall quality of included studies was moderate with two-thirds of included studies showing a high risk of bias for patient selection and applicability (Figure 2). The conduct and the applicability of the index and the reference tests were thought to be adequate in the majority of studies with only four showing a high risk of bias (4/18, 22%) for the index test. Seven studies showed a high risk of bias for patient flow and timing of testing in the study (7/18, 39%) while six studies showed no risk of bias for these items (6/18, 33%) (Figure 2).

Test accuracy meta-analysis

The pooled sensitivity and specificity for USS (12 studies, 1187 women) (16–27) were 0.79 (95%CI 0.63–0.92) and 0.76 (95%CI 0.54–0.93) with a negative and positive likelihood ratio of 0.29 (95%CI 0.13-0.66) and 4.35 (95%CI 2.03-9.32) respectively. Visual inspection of heterogeneity showed greater variability in the sensitivity than the specificity measures (Figure 3). We evaluated the additional use of Doppler with USS in a meta-regression (7 studies, 845 women)(18–20,22–24,26) which showed a slight improvement in sensitivity (0.80, 95%CI 0.67-0.93) and specificity (0.88, 95%CI 0.72-1.00), though not statistically significant (joint model, p-value=0.7). We also conducted subgroup analyses in studies using surgical exploration only as Reference test (n=9, sensitivity 0.81, 95%CI 0.61-0.94, specificity 0.73, 95%CI 0.42-0.94)(18–24,26,27) and in those reporting on adults only (n=3, sensitivity 0.84, 95%CI 0.34-0.98, specificity 0.78, 95%CI 0.42-0.94)(19,20,27). Both subgroups showed similar estimates to the whole population.

Test accuracy meta-analysis for MRI (3 studies, 99 women)(14–16) showed a pooled sensitivity of 0.81 (95%CI 0.63-0.91) and specificity of 0.91 (95%CI 0.80-0.96) (Figure 3). With two case-control and one cohort studies (n=3, 232 women), a meta-analysis for CT was not possible. It had a reported sensitivity ranging from 0.74 to 0.95, and specificity from 0.80 to 0.90. Figure 4 illustrates the scatter of the accuracy parameters for all reported imaging modalities across included studies.

Discussion:

Main findings

Our findings support an overall good performance for USS as a first line diagnostic tool for AT. Evaluating the ovarian vascular blood flow using Doppler slightly improved the diagnostic accuracy of USS, though this was not statistically significant with overlapping confidence intervals. We assessment of CT and MRI was limited by the number of available studies on those two modalities. Overall, MRI seemed to offer higher specificity which could be of value when investigating ambiguous adnexal masses with high suspicion of torsion, but more studies are needed to define the role of MRI in the diagnostic pathway of AT. Data pooling was not possible for CT, though its reported range was consistent with that of USS.

Strengths and limitations

We conducted our review using a standard methodology for diagnostic accuracy reviews, registered our protocol prospectively, and reported according to established guidelines. We adopted a pragmatic search strategy and inclusion criteria including all suspected cases of AT to offer the most comprehensive patient spectrum for evidence synthesis. We considered the potential effect of Doppler on the accuracy of USS using a meta-regression and performed sub-group analyses where possible.

Our findings are not without limitations. Overall, our pooled estimates suffered from heterogeneity likely due to variations in the characteristics of included women (such as age and reproductive status) in our meta-analysis, thus we interpret the findings with caution. Our inclusion criteria are pragmatic and comprehensive to capture the whole literature on the diagnosis of AT. However, we acknowledge the increased heterogeneity and the potential effect of several confounders such as variations in age, reproductive status, operator experience and sequential testing. Majority of studies included a mixed population of

paediatric and adult females which limited our ability to adjust for important factors such as USS route (trans-abdominal vs trans-vaginal) and the underlying ovarian pathology (e.g dermoid cysts). Adjustment for such factors would only be possible using an IPD meta-analysis which was not feasible in our review. Still, we believe our review to offer the most comprehensive evidence synthesis at present to advise current clinical practice.

Interpretation

Establishing an accurate diagnosis in women with suspected AT remains a clinical challenge due to the non-specific presentation and the varied deferential diagnosis. Several ovarian pathologies could produce similar radiological signs (ovarian oedema, unilateral enlargement, midline shift, etc..) as well as overlap with an acute AT (e.g teratoma, endometrioma, haemorrhagic cyst) complicating the radiological diagnosis. As a gynaecological emergency, rapid diagnosis of AT is crucial to optimize the outcomes of affected women and advise any planned surgical intervention (e.g laparoscopy for smaller masses vs laparotomy for large complex torsion). Our estimates support the role of USS as a reliable first-line diagnostic tool for AT. Certainly, several emergency departments now offer rapid-access USS to aid the diagnosis in women with non-specific abdominal pain which seems to optimize the diagnosis and management process (28). Our findings depict relatively wide confidence intervals for the accuracy of USS to diagnose AT. Therefore, clinicians should consider the diagnostic limitations of USS, especially when faced with complex ovarian morphology such as very large cysts, complex masses or paediatric cases (4) which might increase the rate of false-negative findings. Given the established limitations of USS, clinicians should correlate the clinical, biochemical and radiological findings before deciding to operate on symptomatic women. Such practice is key specifically when planning the management of particular patient

groups (e.g prepubertal girls and pregnant women) to aid the decision making for the surgical route of choice (e.g laparotomy for large complex masses) and the surgical approach (oophorectomy vs conservative surgery) (2).

The role of MRI in investigating larger and more complex ovarian morphology is well established (29–32). However, considering its higher cost and limited availability, reserving its use as a second-line diagnostic tool seems reasonable within the context of our findings. We were unable to identify unified diagnostic criteria to establish an ultrasonographic diagnosis of AT due to the varied reporting across included studies. This was also the case for reported diagnostic radiological features on CT and MRI. Certain features seem to be more suggestive of AT (e.g. ovarian oedema> 5cm, twisted pedicles on color Doppler, free fluid in the pelvis, and the whirlpool sign) (17,19,33), however, future consensus work is needed to evaluate the accuracy of unified diagnostic criteria that correlates with the clinical presentation.

Establishing a well-defined care pathway for women presenting with acute abdominal/pelvic pain shared across multiple disciplines is key for efficient diagnosis and management of AT(34). Currently, care for affected women is heterogeneous, often tailored by the attending clinician and their speciality of interest (emergency medicine, general surgery, urology, gynaecology…) increasing the chance of delayed diagnosis and treatment. Developing and evaluating standardized care pathways with rapid access to imaging services is needed to improve the longterm outcomes of women with AT.

Conclusion

Ultrasound has good performance as a first-line diagnostic test for women with suspected adnexal torsion. Magnetic resonance could offer improved specificity to investigate complex ovarian morphology, but more evidence is needed.

Acknowledgement: None

Conflict of interest: All authors have nothing to disclose. BHA holds a personal Lecturership from the UK National Health Institute of Research. KSK is Distinguished Investigator at the University of Granada with a grant awarded by the Beatriz Galindo Program (senior modality) of the Spanish Ministry of Science, Innovation and Universities.

Contribution to Authorship: BW and MR conducted the search, data extraction and 1st drafting of the manuscript. BHA and ER conducted the statistical analysis. MM and KSK contributed to data interpretation and final editing of the manuscript.

References:

1. Galinier P, Carfagna L, Delsol M, Ballouhey Q, Lemasson F, Le Mandat A, et al. Ovarian torsion. Management and ovarian prognosis: a report of 45 cases. J Pediatr Surg. 2009;44(9):1759–65.

2. Huchon C, Fauconnier A. Adnexal torsion: a literature review. Eur J Obstet Gynecol Reprod Biol. 2010;150(1):8–12.

3. Shadinger LL, Andreotti RF, Kurian RL. Preoperative sonographic and clinical characteristics as predictors of ovarian torsion. J Ultrasound Med. 2008;27(1):7–13.

4. Chang HC, Bhatt S, Dogra VS. Pearls and pitfalls in diagnosis of ovarian torsion. Radiographics. 2008;28(5):1355–68.

5. 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(9):1205–10.

6. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. Cochrane handbook for systematic reviews of diagnostic test accuracy. Version 09 0 London Cochrane Collab. 2010;

7. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. Jama. 2018;319(4):388–96.

8. Burch J, St John J, Duffy S, Smith S, Soares-Weiser K, Kleijnen J, et al. Should data from diagnostic case-control studies be included in systematic reviews alongside diagnostic cohort studies? J Med Screen. 2014;14(3):132–7.

9. Harbord RM, Whiting P. Metandi: meta-analysis of diagnostic accuracy using

hierarchical logistic regression. Stata J. 2009;9(2):211–29.

10. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. Bmj. 2004;329(7458):168–9.

11. 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(4):733–8.

12. Lee MS, Moon MH, Woo H, Sung CK, Oh S, Jeon HW, et al. CT findings of adnexal torsion: A matched case-control study. PLoS One. 2018;13(7):e0200190.

13. Jung S Il, Park HS, Yim Y, Jeon HJ, Yu MH, Kim YJ, et al. Added value of using a CT coronal reformation to diagnose adnexal torsion. Korean J Radiol. 2015;16(4):835–45.

14. Birchard KR, Brown MA, Hyslop WB, Firat Z, Semelka RC. MRI of acute abdominal and pelvic pain in pregnant patients. Am J Roentgenol. 2005;184(2):452–8.

15. Singh AK, Desai H, Novelline RA. Emergency MRI of acute pelvic pain: MR protocol with no oral contrast. Emerg Radiol. 2009;16(2):133.

16. Béranger-Gibert S, Sakly H, Ballester M, Rockall A, Bornes M, Bazot M, et al. Diagnostic value of MR imaging in the diagnosis of adnexal torsion. Radiology. 2016;279(2):461–70.

17. Jourjon R, Morel B, Irtan S, Audureau E, Coulomb-L'Herminé A, Larroquet M, et al. Analysis of clinical and ultrasound determinants of adnexal torsion in children and adolescents. J Pediatr Adolesc Gynecol. 2017;30(5):582–90.

18. 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(1):64–6.

19. Bar-On S, Mashiach R, Stockheim D, Soriano D, Goldenberg M, Schiff E, et al. Emergency laparoscopy for suspected ovarian torsion: are we too hasty to operate? Fertil Steril. 2010;93(6):2012–5.

20. 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(2):83–9.

21. Resapu P, Rao Gundabattula S, Bharathi Bayyarapu V, Pochiraju M, Surampudi K, Dasari S. Adnexal torsion in symptomatic women: a single-centre retrospective study of diagnosis and management. J Obstet Gynaecol (Lahore). 2019;39(3):349–54.

22. Niafar F, Mirfendereski S, Rostamzadeh A. Diagnostic efficacy of sonography for diagnosis of ovarian torsion. Zahedan J Res Med Sci. 2014;16(11):37–9.

23. 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(6):1283–7.

24. Mawaldi L, Gupta C, Bakhsh H, Saadeh M, Abolfotouh MA. Validity of ultrasound in patients with acute pelvic pain related to suspected ovarian torsion. Surg Sci. 2011;2(6):326.

25. Graif M, Itzchak Y. Sonographic evaluation of ovarian torsion in childhood and adolescence. Am J Roentgenol. 1988;150(3):647–9.

26. Feng J, Lei T, Xie H, Li L, Du L. Spectrums and outcomes of adnexal torsion at different ages. J Ultrasound Med. 2017;36(9):1859–66.

27. Cohen SB, Weisz B, Seidman DS, Mashiach S, Lidor AL, Goldenberg M. Accuracy of the preoperative diagnosis in 100 emergency laparoscopies performed due to acute abdomen in nonpregnant women. J Am Assoc Gynecol Laparosc. 2001;8(1):92–4.

28. Villalba M, Lambert M. Gynaecological Ultrasound in Emergency Medicine. Emerg Point-of-Care Ultrasound. 2017;152.

29. Duan N, Rao M, Chen X, Yin Y, Wang Z, Chen R. Predicting necrosis in adnexal torsion in women of reproductive age using magnetic resonance imaging. Eur Radiol. 2020;30(2):1054–61.

30. Moribata Y, Kido A, Yamaoka T, Mikami Y, Himoto Y, Kataoka M, et al. MR imaging findings of ovarian torsion correlate with pathological hemorrhagic infarction. J Obstet Gynaecol Res. 2015;41(9):1433–9.

31. Sofic A, Husic-Selimovic A, Katica V, Jahic E, Delic U, Sehic A, et al. Magnetic resonance imaging (MRI) and transvaginal ultrasonography (TVU) at ovarian pain caused by benign ovarian lesions. Acta Inform Medica. 2018;26(1):15.

32. Emil S, Youssef F, Arbash G, Baird R, Laberge J-M, Puligandla P, et al. The utility of magnetic resonance imaging in the diagnosis and management of pediatric benign ovarian lesions. J Pediatr Surg. 2018;53(10):2013–8.

33. Jung S Il, Park HS, Jeon HJ, Yu MH, Kim YJ, Chung J, et al. Whirlpool sign of adnexal torsion on CT: where can we find it? Acta radiol. 2020;61(5):714–20.

34. Macaluso CR, McNamara RM. Evaluation and management of acute abdominal pain in the emergency department. Int J Gen Med. 2012;5:789.