Critical Review Form
Diagnostic Test

Preoperative Sonographic andClinical Characteristics as
Predictors of Ovarian Torsion, *J Ultrasound Med* 2008; 27:7–13.

Objectives: "To evaluate preoperative sonographic and clinical characteristics predictive of ovarian torsion in an effort to expedite diagnosis and therefore improve clinical outcomes in both pediatric and adult female patients". (p. 8)

Methods: Retrospective chart review without clear methods (**Gilbert 1996**, **Worster 2004**) of female patients presenting to either Vanderbilt University Medical Center or Vanderbilt Children's Hospital from July 2000 through December 2005 with both pathologically proven ovarian torsion (OT) and reviewable sonographic studies. Medical records were reviewed for pain duration, nausea/vomiting, fever, leukocytosis, pregnancy status, and age of menarche or menopause.

Sonograms were evaluated by an "independent" (does that mean blinded to the clinical data or blinded to the clinical sonographers interpretation?) radiologist for the imaging characteristics of ovarian vascularity, morphologic appearance, associated cysts or masses, intraperitoneal free fluid, and right/left lower quadrant location of the abnormality. The pathologists (no statement of **independence** or blinding) reviewed histopathologic findings via standard hematoxylin-eosin staining techniques.

Investigators do not report who performed the sonograms, how experienced they were, or whether they were blinded to additional clinical data. Pregnant, immediate postpartum, and virginal patients were examined transabdominally only. Two dimensional color Doppler sonography was conducted with HDI 4000, HDI 5000, and iU22 scanners, PowerVision 8000 and Aplio scanners, and Acuson Sequoia and Aspen scanners using 2- to 5-MHz curved array probes, 4- to 8- MHz vector probes, 6-MHz linear probes, and 4- to 8-MHz curved transvaginal probes.

	Guide	Comments
I.	**Are the results valid?**	
A.	**Did clinicians face diagnostic uncertainty?**	Yes, but readers are not certain which clinicians since the setting of patient presentation (ED? GYN clinic?) is not specified nor are any other patient characteristics or physical exam findings reported.
B.	**Was there a blind comparison with an independent gold standard applied similarly to the treatment group and to the control group?** **(Confirmation Bias)**	No. Thirty subjects without an ultrasound, three without Doppler, one with fallopian torsion, and three with surgical detorsion were excluded from analysis. In addition, multiple other abdominal pain patients where OT was not considered and/or where Doppler studies were unremarkable likely never underwent surgical or histopathological exam. Therefore, verification bias is most certainly present and sensitivity was likely artificially increased.
C.	**Did the results of the test being evaluated influence the decision to perform the gold standard?** **(Ascertainment Bias)**	Yes. As noted above clinicians' decisions to pursue laparoscopic evaluation of abdominal pain were most certainly influenced by the physical exam and sonographic findings being studied. Again, this probably biases sensitivity upward.
II.	**What are the results?**	

A.	**What likelihood ratios were associated with the range of possible test results?**	- 39 patients (9 premenarche, 28 reproductive, 2 post menopausal) with a mean age of 21 years (range 3-61 years) included in the analysis.

Sensitivity of History, Physical Exam, and Labs

Finding	# Pts	Sensitivity (%)
WBC > 11,000	22	56%
Nausea/vomiting	33	85%
Temp > 100.5°F	7	18%
Pain duration		
≤ 1 day	13	33%
2 day	12	31%
≥ 3 day	14	36%

Sensitivity of Ovarian Doppler Findings

Duration of Pain	Arterial Flow Absent	Venous Flow Absent
≤ 1 d	54%	85%
2 d	50%	75%
≥ 3 d	36%	43%

- 56% (22/39) only underwent transabdominal US.

Ultrasound Findings

- 56% of OT was associated with a cyst (25%) or benign mass (31%).

- 21% showed free fluid in the abdomen.

- 26% had ground-glass appearance of the ovary.

- Ovary or ovary/mass complex were uniformly enlarged ranging from 24 cm^3 in 3-yr old child (normal mean 1cm^3) and 957 cm^3 in an adult (normal mean 15cm^3 while reproductive or 8cm^3 post-menopausal).

- There were insufficient numbers of premenarchal or postmenopausal women to conduct meaningful comparisons between age groups, but among 28 reproductive age women arterial flow was only absent in 50% while venous flow was absent in 61%.

III.	**How can I apply the results to patient care?**	
A.	**Will the reproducibility of the test result and its interpretation be satisfactory in my clinical setting?**	Uncertain since the patient and clinical setting details were not provided nor did the investigators provide any estimates of interrater reliability for physical exam or Doppler findings. Also, 95% CIs and post-hoc power calculations are not provided.
B.	**Are the results applicable to the patients in my practice?**	Probably since we default to GYN sonography for the definitive diagnosis and management of OT, but more details about patients, the clinical setting and sonographer skill sets would enhance readers' ability to confidently apply this data in various settings like the ED
C.	**Will the results change my management strategy?**	Yes, by recognizing that the absence of venous flow with pain duration < 1 day is much more sensitive (85% sensitivity) than the absence of arterial flow (54% sensitivity) for the diagnosis of OT.
D.	**Will patients be better off as a result of the test?**	Yes, although future trials are needed of consecutive abdominal pain patients with heterogeneous sonographers and interrater reliability assessments in order to fully understand the sensitivitiy, specificity, and likelihood ratios. However, for now "Suspicion of ovarian torsion should be high in the setting of clinical symptoms and ovarian enlargement regardless of the presence or absence of ovarian Doppler signals, and surgical exploration should not be delayed. Further research is necessary to develop methods of determining the viability of the ovary because the presence or absence of spectral Doppler signals does not appear to be a reliable tool for viability assessment". (p 13)

Limitations

1) Failure to report sufficient detail for chart review methods (Worster 2004, Gilbert 1996). Specifically,
    a) How cases were identified (ICD-9 codes, surgical records, etc)?
    b) What proportions of data were missing?
    c) How were conflicting data coded?
    d) Were chart abstractors trained, monitored, and/or blinded to the primary hypothesis?
    e) Were standardized data abstraction instruments/forms used?
    f) Were outcome assessors (Pathology) blinded to clinical and sonographic data?

2) Failure to identify results as sensitivity or to provide 95% CI around point estimates.

3) Failure to stratify data by patient site of entry (ED, non-ED) or by physician obtaining the data (EM vs. non-EM).

4) Failure to describe the prevalence and diagnostic test characteristics for additional signs and symptoms that are dogma for the evaluation of abdominal pain: pain location, diarrhea/constipation, urinary symptoms, vaginal bleeding or discharge, time since last menstrual period, palpable mass, cervical motion tenderness, pain descriptors, or quantitative pain severity.

5) Failure to assess the sensitivity for constellations of symptoms.

6) Failure to compute/report post-hoc power calculation.

7) Failure to explicitly describe what surgical criteria were used to diagnose ovarian torsion.

8) Failure to describe the sonographer experience or to perform an interrater reliability assessment of US interpretations.

9) Retrospective case series without the ability to compute specificity or likelihood ratios. In fact, the authors grossly misinterpret and overstate the relevance of their findings:

> "Therefore, whereas the presence of lower quadrant pain plus 1 or more of those symptoms should raise clinical suspicion for torsion, the absence of these additional symptoms does not preclude the diagnosis". (p.10)

For this statement to be true the specificity would need to be acceptably high and the positive LR > 10 (**Brown 1999**). We have <u>no idea</u> what the specificity or LR's are because no disease-negative subjects were included.

Since $LR^+ = \dfrac{sen}{(1 - spec)}$ and $LR^- = \dfrac{(1 - sen)}{spec}$

we can manufacture a range of specificities to test what the LR+ and LR- would become.

Sample LR's (not actual data) for venous flow symptoms < 1 day

Reported Sensitivity	Invented Specificity	LR^+	LR^-
85%	5%	1	3
85%	50%	2	0.3
85%	95%	17	0.16

Bottom Line

In adult patients with acute (less than 3-days since symptom onset) lower quadrant abdominal pain the absence of ovarian venous flow is more sensitive for the diagnosis of ovarian torsion than is the absence of arterial flow (67% vs. 46%, respectively). Venous flow absence is more sensitive at less than 1-day since the symptoms began (85%) than later on (2-days, 75%, ≥ 3-days, 43%). All patients have enlarged ovaries and 56% are associated with a cyst or benign mass. Future prospective studies of consecutive patients with and without ovarian torsion are essential to fully understand the actual diagnostic accuracy (sensitivity/specificity/likelihood ratios) and reproducibility of history, physical exam and sonographic findings for the diagnosis of ovarian torsion.