Critical Review Form

Meta-analysis

Evidence based diagnostics: Adult septic arthritis, *Acad Emerg Med 2011;18: 782-796*

Objective: **"To assess the pretest probability of and diagnostic test characteristics (sensitivity, specificity, LRs) for nongonococcal septic arthritis from elements of the history, physical examination, and laboratory tests available at the bedside. A secondary objective was to define arthrocentesis test and treatment thresholds using the Pauker-Kassirer method based on best estimates of sensitivity, specificity, diagnostic risks, and treatment benefits and risks derived from this systematic literature review." (p. 783)**

Methods: **Adhering to MOOSE criteria, one investigator conducted an electronic search of PUBMED and EMBASE from 1966-2010, in conjunction with a medical librarian. To identify original diagnostic research the search terms septic arthritis and infectious arthritis were combined with MeSH terms emergency medicine, physical examination, history, diagnostic tests, sensitivity, and specificity. In order to find CDR's the same search terms and MeSH headings were used with the PUBMED clinical query setting "clinical prediction guides/broad". For the test-treatment threshold assessment, PUBMED was searched using the terms arthrocentesis and risk, while also searching for interventional effectiveness using PUBMED clinical query "therapy/broad" and the search term septic arthritis. Search results were limited to human studies and English language. Manuscript and textbook bibliographies were hand searched, as were scientific assembly abstracts from ACEP and SAEM (1990-2011).**

Two authors reviewed titles and abstracts for inclusion. Studies were included if they enrolled adult patients with acutely swollen or painful joints. Synovial culture was the preferred gold standard, but imperfect gold standards (positive gram stain or blood culture, purulence on operative drainage, clinical improvement on antibiotics) were also accepted. Studies were excluded if they evaluated only pediatric patients, gonococcal arthritis, or diagnostic tests not available in the contemporary ED.

Two authors independently assessed individual manuscript quality using QUADAS. If the individual studies were not ED-based patient populations, then the spectrum question of QUADAS was "No". Similarly, if the criterion standard was not explicitly defined or if not clearly stated that the index interpreter and gold standard outcome assessor were blinded to the other test, the relevant QUADAS questions were answered "No". Reliability between these two authors was assessed using Kappa (κ).

Two authors independently abstracted data from the original studies: setting, inclusion criteria, criterion standard, disease prevalence, and diagnostic test characteristics. "Disease" was defined as non-gonococcal bacterial arthritis using the original study criterion standard, whereas "no disease was defined as the absence of a bacterial etiology for the acute arthritis. The following definitions were used to construct 2x2 tables:

> **True positive – diagnostic test correctly identified bacterial arthritis at given threshold.**
> **False positive – abnormal test result suggesting bacterial arthritis when the criterion standard did not demonstrate septic arthritis.**
> **True negative – test correctly noted no bacterial arthritis and the criterion standard confirmed no bacterial etiology.**
> **False negative – test suggested no bacterial arthritis when a bacterial etiology was identified by the criterion standard.**

When appropriate meta-analysis was conducted with a random-effects model using Meta-Disc software. Interstudy heterogeneity was assessed for sensitivity and specificity pooled estimates using I^2 (inconsistency index). When sufficient detail was available interval LR's were computed. The Pauker-Kassirer decision model was used to derive test-treatment estimates.

Guide	Question	Comments
I	***Are the results valid?***	
1.	Did the review explicitly address a sensible question?	Yes – the diagnostic accuracy of history, physical exam, and labs to distinguish septic arthritis from other etiologies of acute monoarticular joint pain/swelling in the ED.

<table>
<tr><td>2.</td><td>Was the search for relevant studies details and exhaustive?</td><td>Yes – although the authors neglected non-English studies. Could have also hand searched research abstracts in Ortho, ID, Rheumatology, and Medical Decision Making societies/scientific meetings.</td></tr>
<tr><td>3.</td><td>Were the primary studies of high methodological quality?</td><td>No. “The quality of the diagnostic trials for septic arthritis is highly variable (Table 1). Only four studies specifically note inclusion of ED populations. Several trials did not explicitly describe any inclusion criteria for their study populations or which criteria standard(s) were employed for the diagnosis of septic arthritis. Most studies do not report the interval between the index test and the criterion standard. In addition, few studies explicitly describe blinding the assessors for the index test from the criterion standard or vice versa.” (p. 785)</td></tr>
<tr><td>4.</td><td>Were the assessments of the included studies reproducible?</td><td>Yes. As noted in Table 1 (p. 786) the inter-rater reliability for the 13 domains of QUADAS ranged between κ 0.619 and 1.0. The lowest κ values were for the domains of selection criteria, blinded index tester, and presence of additional clinical data.</td></tr>
<tr><td>II.</td><td>What are the results?</td><td></td></tr>
<tr><td>1.</td><td>What are the overall results of the study?

History Risk Factors

<u>Risk</u> <u>LR+</u> <u>LR-</u>
Age > 80 3.5 0.86
DM 2.7 0.93
RA 2.5 0.45
Jt Surg* 6.9 0.78
Art hip/knee 3.1 0.73
Skin infect 2.8 0.76
Pros + infect 15.0 0.70
HIV 1.2 0.64

* Joint surgery < 3 months ago.</td><td>• PUBMED search yielded 1699 citations, EMBASE 2386 citations, and 11 additional references were identified by bibliometric hand search. A total of 32 original diagnostic trials were included in this systematic review, including 18 retrospective, 12 prospective, and 2 case-control designs.
• The majority of trials only assessed disease-positive patients so only sensitivity (not specificity or LR’s) is reportable.
• Only 4 studies specifically note inclusion of ED populations.
• Prevalence estimates ranged from 0.4% to 45%, but the only ED-based prospective study reporting prevalence estimated that 27% of acute monoarticular arthritis patients with suspected septic arthritis will have septic arthritis (95% CI 17%-38%).
• Only one study (5000 Dutch Rheumatology Clinic patients) evaluated the diagnostic accuracy of history (Table 2, page 787 – see at left). None of these risk factors significantly reduces the probability of septic arthritis when absent.
• No studies evaluated specificity of physical exam findings.
• No studies evaluated the sensitivity or specificity of clinical gestalt.</td></tr>
</table>

Serum Markers

	Range	LR+	LR-
WBC*		1.4-1.7	0.28-0.84
ESR*		1.3-7.0	0.17-2.4
CRP*		1.1-4.5	0.3-0.7
Procalcitonin		5-∞	0.3-0.7
TNF		∞	0.7
IL-6		1.5	0.9
IL-β		3.2	0.8

* Various thresholds.

- No clinical decision rules were identified for adult septic arthritis.
- With the exception of cytokines which are generally not available in most ED's (TNFα and IL-6) and procalcitonin, no serum inflammatory marker or threshold accurately distinguishes septic arthritis from non-SA (see table at left).
- Blood culture sensitivity ranged from 23%-36%, but no studies assessed specificity.
- Based upon 7 trials for sWBC > 50,000 (I^2 = 54% for sensitivity, I^2 = 71% for specificity) and 3 trials for sWBC > 100,000 (I^2 = 70% for sensitivity, I^2 = 68% for specificity), the following results were obtained via meta-analysis

sWBC	LR+ (95% CI)	LR- (95% CI)
>50,000	4.7 (2.5-8.5)	0.52 (0.38-0.72)
>100,000	13.2 (3.6-51)	0.83 (0.80-0.89)

- One study noted that prosthetic knee infections produce lower mean sWBC than do native knee joint infections.
- Based upon four trials, the following interval LR's were computed for sWBC.

sWBC interval	interval LR
0-25000	0.33
25000-50000	1.06
50000-100000	3.59
100000	∞

- Gram stain sensitivity ranges from 29% to 65%.
- Except for synovial lactate, other traditional synovial lactate tests are generally not helpful (see left).
- One study assessed magnetic resonance spectroscopy to measure synovial lactate, demonstrating moderate correlation (but no measures of diagnostic accuracy).
- One study assessed the diagnostic accuracy of PCR pathogen-specific probes yielding organism specific results within 3 hours with LR+ 31.7 and LR- 0.05.
- Based upon the meta-analysis estimates of sensitivity (56%) and specificity (90%) for sWBC > 50, as well as risk of treatment with patients without disease (15.5%), risk of diagnostic arthrocentesis (0.037%), and benefit of treatment in septic arthritis patients of

		50%, the Pauker-Kassirer test threshold was 5.2% and treatment threshold 38.7% (see attached Excel calculator to recomputed test- and treatment-thresholds based upon different estimates).
2.	How precise are the results?	See 95% CI provided above
3.	Were the results similar from study to study?	No, there was significant heterogeneity in the design of the diagnostic studies and statistically in the I^2 for the meta-analysis. However, the meta-analysis authors investigated the stability of their summary estimates of sen/spec for SWBC via a sensitivity analysis: "Sensitivity analysis was performed for a SWBC > 50,000 by sequentially excluding each trial and recomputing summary sensitivity and specificity. Exclusion of the Kortenagas et al trial eliminated heterogeneity for sensitivity (I^2 = 0%) with a summary estimate of 62% sensitivity. For specificity, heterogeneity could only be reduced by excluding the Kortenagas et al, Soderquist et al, and Schmerling et al trials (I^2 = 27%) with a summary estimate of 91% for specificity."
III.	***Will the results help me in caring for my patients?***	
1.	How can I best interpret the results to apply them to the care of my patients?	• Based upon one study, history and physical exam are generally inaccurate with the exception of prosthetic joint patients with overlying cellulitis. • Serum tests are inaccurate. • Synovial WBC > 100,000 (iLR = ∞) is very helpful to rule-in the diagnosis of SA, but sWBC 0-25,000 cannot definitively exclude the diagnosis (iLR 0.33). Therefore, cultures should always be sent and followed. • Involve patients in decision and awareness of the test- and treatment thresholds to facilitate informed shared decision making. • Synovial LDH and synovial lactate (D-lactate or L-lactate?) is probably worth considering in equivocal cases.
2.	Were all patient important outcomes considered?	No, those studies were all focused on diagnostic accuracy (Stage II of diagnostic study hierarchy) Whether any of these tests reduce suffering, mortality, or costs is unknown.
3.	Are the benefits worth the costs and potential risks?	Uncertain since no cost-effectiveness studies were reported or contemplated.

Limitations

1) **English language only.**

2) **Low to moderate quality evidence by QUADAS criteria.**

3) **Only one study assessed the diagnostic accuracy of history and none evaluated physical exam.**

4) **Exclusion of gonococcal-arthritis (by design).**

5) **Lack of definitive septic arthritis treatment randomized controlled trials or methodologically pristine observational trials yield suboptimal estimates of treatment risk/benefit for test-treatment estimates.**

6) **No patient-centric outcomes reported or incorporation of patient perspectives into the test-treatment equation.**

Bottom Line

Diagnostic studies of history and physical exam to evaluate septic arthritis in any setting are virtually non-existent. Few septic arthritis diagnostic studies adhere to STARD criteria. Serum tests (WBC, ESR, CRP) for septic arthritis are inaccurate and probably worthless acutely. Synovial gram stain has sensitivity 29% - 65% with an undefined specificity. A swWBC > 100,000 has an iLR of ∞, whereas a sWBC 0-25,000 has iLR 0.33. Synovial lactate and sLDH, as well as PCR, are promising tests for the future ED evaluation of suspected acute SA. The best estimate pre-test prob for septic arthritis in the ED is 27% and the test-threshold 5% with a treatment threshold of 39%.