



CONTINUING MEDICAL EDUCATION

Continuing Medical Education Activity in *Academic Emergency Medicine*

CME Editor: Hal Thomas, MD

Authors: Christopher R. Carpenter, MD, MS, Jeremiah D. Schuur, MD, MHS, Worth W. Everett, MD, and Jesse M. Pines, MD, MBA, MSCE

Article Title: Evidence-based Diagnostics: Adult Septic Arthritis

If you wish to receive free CME credit for this activity, please refer to the website: <http://www.wileyblackwellcme.com>.

Accreditation and Designation Statement:

Blackwell Futura Media Services designates this journal-based CME activity for a maximum of 1 *AMA PRA Category 1 Credit*[™]. Physicians should only claim credit commensurate with the extent of their participation in the activity.

Blackwell Futura Media Services is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians.

Educational Objectives

After completing this activity the participant will be able to restate criteria for the diagnosis of septic arthritis.

Activity Disclosures

No commercial support has been accepted related to the development or publication of this activity.

Faculty Disclosures:

CME editor – Hal Thomas, MD: No relevant financial relationships to disclose.

Authors – Christopher R. Carpenter, MD, MS, Jeremiah D. Schuur, MD, MHS, Worth W. Everett, MD, and Jesse M. Pines, MD, MBA, MSCE: No relevant financial relationships to disclose.

This manuscript underwent peer review in line with the standards of editorial integrity and publication ethics maintained by *Academic Emergency Medicine*. The peer reviewers have no relevant financial relationships. The peer review process for *Academic Emergency Medicine* is double-blinded. As such, the identities of the reviewers are not disclosed in line with the standard accepted practices of medical journal peer review.

Conflicts of interest have been identified and resolved in accordance with Blackwell Futura Media Services's Policy on Activity Disclosure and Conflict of Interest. No relevant financial relationships exist for any individual in control of the content and therefore there were no conflicts to resolve.

Instructions on Receiving Free CME Credit

For information on applicability and acceptance of CME credit for this activity, please consult your professional licensing board.

This activity is designed to be completed within an hour; physicians should claim only those credits that reflect the time actually spent in the activity. To successfully earn credit, participants must complete the activity during the valid credit period, which is up to two years from initial publication.

Follow these steps to earn credit:

- Log on to <http://www.wileyblackwellcme.com>
- Read the target audience, educational objectives, and activity disclosures.
- Read the article in print or online format.
- Reflect on the article.
- Access the CME Exam, and choose the best answer to each question.
- Complete the required evaluation component of the activity.

This activity will be available for CME credit for twelve months following its publication date. At that time, it will be reviewed and potentially updated and extended for an additional twelve months.



Evidence-based Diagnostics: Adult Septic Arthritis

Christopher R. Carpenter, MD, MS, Jeremiah D. Schuur, MD, MHS, Worth W. Everett, MD, and Jesse M. Pines, MD, MBA, MSCE

Abstract

Background: Acutely swollen or painful joints are common complaints in the emergency department (ED). Septic arthritis in adults is a challenging diagnosis, but prompt differentiation of a bacterial etiology is crucial to minimize morbidity and mortality.

Objectives: The objective was to perform a systematic review describing the diagnostic characteristics of history, physical examination, and bedside laboratory tests for nongonococcal septic arthritis. A secondary objective was to quantify test and treatment thresholds using derived estimates of sensitivity and specificity, as well as best-evidence diagnostic and treatment risks and anticipated benefits from appropriate therapy.

Methods: Two electronic search engines (PUBMED and EMBASE) were used in conjunction with a selected bibliography and scientific abstract hand search. Inclusion criteria included adult trials of patients presenting with monoarticular complaints if they reported sufficient detail to reconstruct partial or complete 2×2 contingency tables for experimental diagnostic test characteristics using an acceptable criterion standard. Evidence was rated by two investigators using the Quality Assessment Tool for Diagnostic Accuracy Studies (QUADAS). When more than one similarly designed trial existed for a diagnostic test, meta-analysis was conducted using a random effects model. Interval likelihood ratios (LRs) were computed when possible. To illustrate one method to quantify theoretical points in the probability of disease whereby clinicians might cease testing altogether and either withhold treatment (test threshold) or initiate definitive therapy in lieu of further diagnostics (treatment threshold), an interactive spreadsheet was designed and sample calculations were provided based on research estimates of diagnostic accuracy, diagnostic risk, and therapeutic risk/benefits.

Results: The prevalence of nongonococcal septic arthritis in ED patients with a single acutely painful joint is approximately 27% (95% confidence interval [CI] = 17% to 38%). With the exception of joint surgery (positive likelihood ratio [+LR] = 6.9) or skin infection overlying a prosthetic joint (+LR = 15.0), history, physical examination, and serum tests do not significantly alter posttest probability. Serum inflammatory markers such as white blood cell (WBC) counts, erythrocyte sedimentation rate (ESR), and C-reactive protein (CRP) are not useful acutely. The interval LR for synovial white blood cell (sWBC) counts of 0×10^9 – 25×10^9 /L was 0.33; for 25×10^9 – 50×10^9 /L, 1.06; for 50×10^9 – 100×10^9 /L, 3.59; and exceeding 100×10^9 /L, infinity. Synovial lactate may be useful to rule in or rule out the diagnosis of septic arthritis with a +LR ranging from 2.4 to infinity, and negative likelihood ratio (–LR) ranging from 0 to 0.46. Rapid polymerase chain reaction (PCR) of synovial fluid may identify the causative organism within 3 hours. Based on 56% sensitivity and 90% specificity for sWBC counts of $>50 \times 10^9$ /L in

From the Division of Emergency Medicine, Washington University in St. Louis School of Medicine (CRC), St. Louis, MO; the Department of Emergency Medicine Brigham and Women's Hospital (JDS), Salt Lake City, UT; the Department of Emergency Medicine, Skagit Valley Hospital (WWE), Mount Vernon, WA; and the Department of Emergency Medicine, George Washington University (JMP), Washington, DC.

Received March 8, 2011; revision received April 8, 2011; accepted April 11, 2011.

Dr. Carpenter was supported by an institutional KM1 Comparative Effectiveness Award (1KM1CA156708-01). Dr. Schuur is a member of the Primary Care and Emergency Medicine Scientific Advisory Board for United Health Care (UHC). He advises UHC on emergency medicine practice. He is supported by a Jahnigen Career development award from the American Geriatrics Society and the Atlantic Philanthropies. He receives grant funding from the Agency for Healthcare Research and Quality and for unrelated work. The authors have no conflicts of interest to declare.

The authors have no relevant financial information or potential conflicts of interest to disclose.

Supervising Editor: Jeffrey Kline, MD.

A related commentary appears on page 797.

Address for correspondence and reprints: Christopher R. Carpenter, MD, MS; e-mail: carpenterc@wusm.wustl.edu. Reprints will not be available.

conjunction with best-evidence estimates for diagnosis-related risk and treatment-related risk/benefit, the arthrocentesis test threshold is 5%, with a treatment threshold of 39%.

Conclusions: Recent joint surgery or cellulitis overlying a prosthetic hip or knee were the only findings on history or physical examination that significantly alter the probability of nongonococcal septic arthritis. Extreme values of sWBC ($>50 \times 10^9/L$) can increase, but not decrease, the probability of septic arthritis. Future ED-based diagnostic trials are needed to evaluate the role of clinical gestalt and the efficacy of nontraditional synovial markers such as lactate.

ACADEMIC EMERGENCY MEDICINE 2011; 18:782–796 © 2011 by the Society for Academic Emergency Medicine

Acute monoarticular arthritis in adults presenting to the emergency department (ED) has multiple potential etiologies including infection (bacterial, fungal, mycobacterial, viral), crystalloid arthropathies, rheumatoid arthritis, lupus, and trauma.^{1–3} Septic (i.e., bacterial) arthritis has an annual incidence of 10 per 100,000 individuals in the United States and is more common among those with rheumatoid arthritis or a prosthetic joint, with up to 70 cases per 100,000.⁴ Patients with human immunodeficiency virus (HIV) are also at increased risk for nongonococcal septic arthritis.⁵ Septic arthritis most commonly affects the knee, which accounts for approximately 50% of cases. In decreasing order of frequency, septic arthritis also affects the hip, shoulder, and elbow, although virtually any articular surface can become infected.⁶ Most cases result from hematogenous spread, since bacterial organisms can easily enter the synovial fluid because synovial tissue lacks a basement membrane. Prompt diagnosis to facilitate appropriate antibiotic management of septic arthritis is essential, since cartilage can be destroyed within days, and in-hospital mortality of treated infections can be as high as 15%.⁷ Permanent disability and increased mortality are associated with delayed presentations and diagnosis.^{7–9} Prior research suggests that using history, physical examination, and synovial tests, clinicians are able to deduce the etiology of acute nontraumatic monoarticular arthritis within 3 days in most cases.¹⁰ Since emergency physicians often lack the luxury of 3-day admissions for most monoarticular arthritis patients, identification of key diagnostic findings to accurately differentiate septic from nonseptic arthritis within minutes to hours is essential.

When conceptualized quantitatively, clinical decision-making is a continuum of disease probabilities from 0% to 100%.¹¹ Health care providers continually revise disease probabilities throughout the clinical encounter based on multiple factors, including elements of the current and past medical examination, imaging and laboratory studies, and therapeutic responses.¹² In 1980, Drs. Pauker and Kassirer described one theoretical model to compute test and treatment thresholds.¹³ Basically, the Pauker-Kassirer algebraic equation provides estimates whereby patients can be divided into three groups: 1) disease probability below the test threshold with further diagnostic testing likely to be more harmful than helpful; 2) disease probability intermediate between the test and treatment thresholds for the diagnosis in question so further testing would be beneficial; 3) disease probability exceeds the treatment threshold with further confirmatory testing a risk to harm

patients, either via therapeutic delay or via unintended consequences of diagnostic test related adverse events.¹⁴

Multiple narrative reviews have been published summarizing the ED management of acute monoarticular joint complaints.^{1,3,15} All of these reviews provide general qualitative statements about the diagnostic tests available when contemplating septic arthritis, but none provide quantitative summaries, evidence quality assessments, or clinician relevant likelihood ratios (LRs) for bedside application.¹¹ Two recent systematic reviews have reported quantitative estimates of diagnostic accuracy.^{2,16} Mathews et al.¹⁶ focused on septic arthritis therapy with little detail provided for diagnostic trials. Margaretten et al.² provided a detailed analysis of diagnostic accuracy in heterogeneous settings, but they did not focus on the ED. They classified the quality of evidence using two previously described, nonvalidated scales.^{17,18} No prior reviews have classified diagnostic evidence using recently validated scales.^{19,20} Additionally, neither traditional textbooks nor prior diagnostic reviews provide ED-relevant estimates of test or treatment thresholds or interval LRs.^{21–23}

The primary objective of this meta-analysis was to assess the pretest probability 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.¹³

METHODS

Search Strategy

The design and manuscript structure of this systematic review conform to the recommendations from the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) statement.²⁴ In conjunction with a medical librarian, one investigator (CRC) searched the medical literature from 1966 to December 2010 using PUBMED and EMBASE for the search terms *septic arthritis* and *infectious arthritis*. The results from these two searches were combined with the MeSH terms *emergency medicine*, *physical examination*, *history*, *diagnostic tests*, *sensitivity*, and *specificity*. The same MeSH headings and combination strategy was used with the PUBMED clinical query setting “clinical

prediction guides/broad" to identify any validated clinical decision rules (CDRs) for septic arthritis. To identify the risk of arthrocentesis for the test threshold analysis, a PUBMED search was conducted using the search terms *arthrocentesis* AND *risk*. To identify the potential risks and benefit of antimicrobial or operative management of septic arthritis, a PUBMED clinical query with the setting "therapy/broad" for the search term *septic arthritis* was conducted with meta-analyses favored to define the least biased point estimate. In addition, the Cochrane Database of Systematic Reviews of Effect was searched using the same search term. All search results were limited to human studies and English language articles. Two authors (CRC, JS) reviewed the titles and abstracts to identify potentially relevant articles, which were then retrieved to review the full manuscript. These authors then independently reviewed these articles for inclusion criteria. The authors also reviewed the references from selected articles as well as the current editions of textbooks of emergency medicine²¹⁻²³ to identify other relevant published research. In addition, the authors conducted online bibliographic searches of abstract submissions to *Academic Emergency Medicine* and *Annals of Emergency Medicine* from 1990 through March 2011.

Studies were included if they recruited adult patients with acutely swollen or painful joints and reported sufficient detail on diagnostic test and criterion standard results to reconstruct two-by-two tables in whole or in part (i.e., if isolated sensitivity or specificity could be computed based on the available data, the study was included). Since synovial fluid culture is only 75% to 95% sensitive for the diagnosis of septic arthritis,²⁵ we also included studies that used positive Gram stain, positive blood culture, purulent drainage on arthrocentesis, or a surrogate outcome of antibiotic response in the setting of acute arthritis with suspected septic arthritis. Articles were excluded if gonococcal arthritis was the primary focus of the research. Narrative reviews, case reports, and studies focused on children or therapy were not included. In addition, trials that assessed diagnostic tests not readily available in the typical ED were excluded.

Individual Evidence Quality Appraisal

Two authors (CRC, JS) used the Quality Assessment Tool for Diagnostic Accuracy Studies (QUADAS) for systematic reviews to evaluate the overall quality of evidence for the trials identified.¹⁹ Discrepant quality assessments were adjudicated by discussion and coming to consensus. Statistical agreement between the two reviewers was assessed via a kappa analysis using SPSS v16.0 (SPSS Inc., Chicago, IL).²⁶ For the purposes of this diagnostic systematic review, several considerations were established a priori to assess the quality of individual trials. The ideal patient population would be those presenting to an ED with a swollen joint and clinical equipoise for the diagnosis of nongonococcal septic arthritis. Although other populations were incorporated into this review, *spectrum bias* may limit the validity of findings in ED settings,²⁷ so if individual trials did not include patients from the ED, then the "spectrum" portion of the QUADAS was assessed as "no." If the

criterion standard used in a trial was not explicitly stated or if the authors did not clearly state blinding of index testers to the criterion standard or vice versa, these portions of the QUADAS were marked as "no." If individual authors did not provide sufficient detail to compute sensitivity or specificity point estimates for varying cutoff values for continuous data, then the QUADAS question for intermediate test results was marked as "no."

Data Analysis

Two authors (CRC, JS) independently abstracted data from the included studies. Information abstracted included study setting, study inclusion criteria, the criterion standard employed, disease prevalence, and diagnostic test properties. Since the terms "false negative," "false positive," "true positive," and "true negative" are usually undefined in studies of septic arthritis that report diagnostic test properties, we standardized our vocabulary a priori. We defined "disease" as nongonococcal bacterial arthritis and "no disease" as the absence of a bacterial etiology for the acute arthritis. The latter category included rheumatoid arthritis, crystalloid arthritis, osteoarthritis, and trauma. A "true positive" was a diagnostic test that correctly identified septic arthritis at a given threshold, whereas a "false positive" indicated an abnormal test result suggesting septic arthritis when the criterion standard did not demonstrate septic arthritis. Similarly, a "true-negative" test indicated the absence of septic arthritis when the criterion standard confirmed no bacterial etiology, while a "false-negative" suggested no septic arthritis when in fact a bacterial etiology was identified by the criterion standard.

To compute meta-analysis summary estimates when more than one study assessed the same index test, we combined trials' patients when they reported septic arthritis and nonseptic arthritis patients using Meta-DiSc (Hospital Universitario Ramón y Cajal, Madrid, Spain) and a random-effects model.^{28,29} If a trial only reported sensitivity estimates for septic arthritis patients without any assessment of specificity, the results of that trial were not used in the meta-analysis. Interstudy heterogeneity was assessed for pooled estimates of sensitivity and specificity using the DerSimonian-Laird random effects model.³⁰ Publication bias was not assessed because of the questionable validity of this approach when assessing diagnostic test meta-analyses.³¹

Synovial white blood cell (sWBC) count and other synovial or serum tests represent continuous variables. Although imposing cutoff values via observation of the receiver operating characteristic curves can be useful clinically, the full value of nondichotomous data is lost in the process. Therefore, we also computed previously unreported interval LRs for individual trials when the original manuscript provided sufficient detail.³² Interval LRs were computed by dividing the proportion of septic arthritis patients within a range of sWBC values over the proportion of those without septic arthritis within the same range of sWBC.^{32,33} When interval LRs from multiple trials were available, the data were combined to identify a summary interval LR.

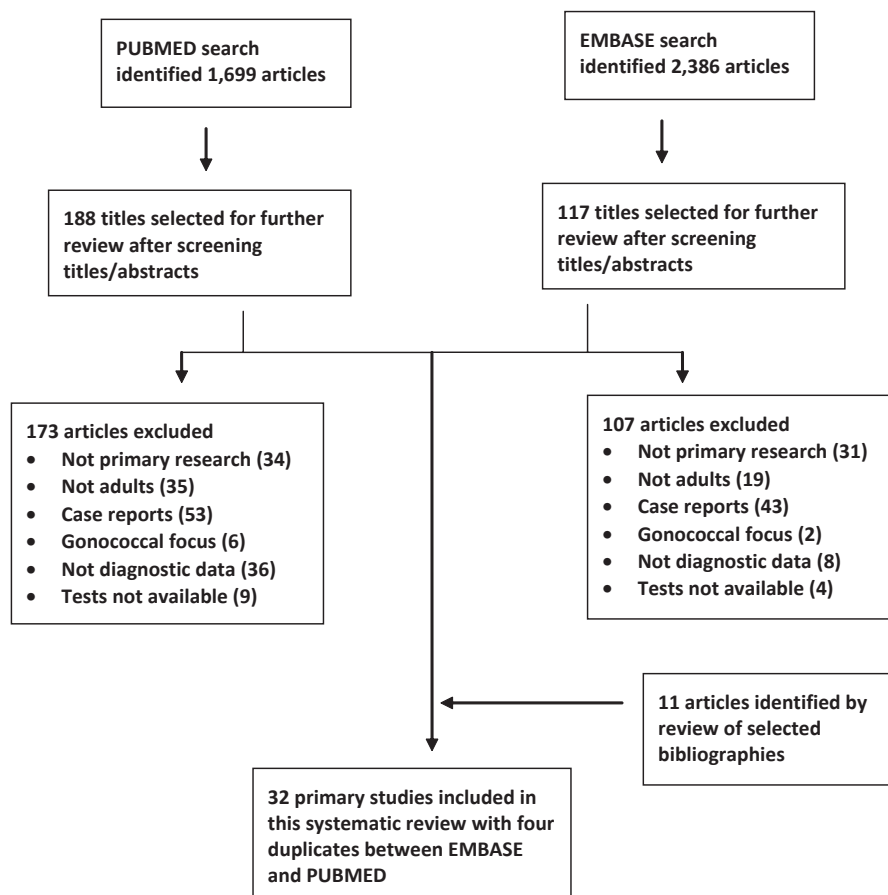


Figure 1. Study selection process.

Test–Treatment Threshold

The Pauker and Kassirer decision threshold model is based on seven variables: false-negative and false-positive proportions, sensitivity, specificity, risk of a diagnostic test, risk of treatment, and anticipated benefit of treatment.¹³ Evidence-based estimates for each of these variables were abstracted from our systematic review to derive theoretical test and treatment thresholds for emergency management of patients with potential septic arthritis. Recognizing that these estimates may be based on incomplete or biased literature, an interactive calculator (in development) is provided with this article to permit readers to alter assumptions and recompute test and treatment thresholds based on test performance or anticipated benefits more applicable to the end-user's patient populations and clinical environment.

RESULTS

The PUBMED search identified 1,699 citations while the EMBASE search identified 2,386 (Figure 1). No additional studies were identified by electronic searches of abstract submissions or by bibliography review of prominent emergency medicine textbooks. As detailed in Figure 1, a total of 32 diagnostic trials were included in the current analysis.^{4,5,7–9,25,34–59} A summary of the trials that were conducted from 1947 until 2007 is presented in Data Supplement S1 (available as supporting

information in the online version of this paper). Eighteen trials were retrospective case series, 12 were prospective case series, and two were case-control studies. No diagnostic randomized controlled trials or adult CDRs were identified. The trials included heterogeneous populations and various diagnostic tests ranging from elements of the history and physical examination, to serum and synovial fluid tests. The majority of trials assessed only patients with septic arthritis, limiting diagnostic conclusions to sensitivity estimates with no ability to compute specificities or LRs.

The authors' QUADAS assessment of quality had a kappa of range 0.612 to 1. The quality of the diagnostic trials for septic arthritis is highly variable (Table 1). Only four studies specifically note inclusion of ED populations.^{8,47,58,59} Several trials did not explicitly describe any inclusion criteria for their study populations^{9,34,37,38} or which criterion standard(s) were employed for the diagnosis of septic arthritis.^{35,36,44} Most studies do not report the interval between the index test and the criterion standard. In addition, few studies explicitly describe blinding the assessors of the index test from the criterion standard or vice versa. For continuous data such as sWBC count, the majority of studies defined dichotomous normal and abnormal cutoffs based on their hospital norms, rather than reporting all of the results that would allow readers to compute diagnostic accuracy at various thresholds in addition to interval LRs.^{32,33}

Table 1
Quality of the Evidence (QUADAS)

Study	Spectrum	Selection Criteria	Reference Standard	Interval Testing	Verification Uniform	No Incorporation Bias	Reproducible Index Test	Reproducible Gold Standard	Blinded Index Tester	Blinded Criterion Standard	Additional Clinical Data	Intermediate Test Results Reported	Withdrawals Explained
Ward 1960 ³⁴	N	N	Y	N	N	N	N	N	N	N	Y	N	N
Argen 1966 ³⁵	N	Y	N	N	N	N	N	N	N	N	Y	Y	Y
Goldenberg 1976 ³⁶	N	Y	N	N	N	N	N	N	N	N	Y	N	N
Brook 1978 ³⁷	N	N	Y	N	Y	N	Y	Y	N	N	N	Y	N
Krey 1979 ³⁸	N	N	Y	Y	Y	Y	Y	N	Y	N	Y	Y	N
Rosenthal 1980 ⁴⁰	N	Y	Y	N	N	N	N	N	N	N	Y	N	N
Mossman 1981 ³⁹	Y	N	N	N	N	Y	Y	N	Y	N	N	Y	Y
Riordan 1982 ⁴¹	N	N	Y	N	Y	N	Y	Y	N	N	Y	Y	Y
Cooper 1986 ⁸	Y	Y	Y	N	Y	N	N	N	N	N	Y	N	Y
Schmerling 1990 ²⁵	N	Y	Y	Y	Y	N	N	N	N	N	Y	Y	Y
Deesomchok 1990 ⁴²	N	Y	Y	N	Y	N	N	N	N	N	Y	N	N
McCutchan 1990 ⁴³	N	Y	Y	N	Y	N	N	N	N	N	Y	N	N
Schlapbach 1990 ⁴⁴	N	Y	N	N	Y	N	N	N	N	N	Y	Y	N
Kortekangas 1992 ⁴⁵	N	Y	N	N	N	N	N	N	N	N	Y	Y	N
Kaandorp 1995 ⁴	N	Y	Y	N	N	N	N	N	N	N	Y	N	Y
Gratacós 1995 ⁴⁶	N	N	N	Y	N	N	Y	N	N	N	Y	Y	Y
Saraux 1997 ⁵	N	Y	Y	N	N	Y	Y	Y	Y	N	Y	N	N
Jeng 1997 ⁴⁷	Y	Y	Y	Y	Y	N	Y	Y	N	N	Y	N	N
Söderquist 1998 ⁹	N	N	Y	Y	N	N	Y	N	N	N	Y	Y	Y
Gupta 2001 ⁷	N	Y	Y	Y	N	N	N	Y	N	N	Y	N	N
Faraj 2002 ⁴⁸	N	N	Y	N	Y	N	N	N	N	N	Y	N	Y
Trampuz 2004 ⁴⁹	N	Y	Y	Y	Y	N	Y	Y	N	N	Y	Y	Y
Li 2004 ⁵⁰	N	Y	Y	N	N	N	N	N	N	N	Y	Y	N
Martinot 2005 ⁵¹	N	Y	Y	N	Y	N	Y	N	N	N	Y	Y	Y
Abdullah 2007 ⁵²	Y	Y	N	Y	Y	N	Y	N	N	N	Y	N	Y
Li 2007 ⁵³	N	Y	Y	N	N	N	N	N	N	N	Y	N	N
McGillicuddy 2007 ⁵⁴	N	Y	Y	N	N	N	N	N	N	N	Y	N	N
Fottner 2008 ⁵⁵	Y	Y	Y	Y	Y	Y	Y	N	N	N	Y	Y	N
Hügler 2008 ⁵⁶	N	Y	Y	N	Y	Y	Y	N	N	N	Y	Y	N
Wiener 2008 ⁵⁷	N	Y	Y	N	Y	N	Y	Y	N	N	Y	Y	N
Yang 2008 ⁵⁸	Y	Y	Y	N	Y	Y	Y	Y	N	N	N	Y	Y
Ernst 2010 ⁵⁹	Y	Y	Y	N	N	N	N	Y	N	N	Y	N	N
Kappa	0.818	0.612	1	0.855	0.938	0.765	0.874	1	0.619	1	0.632	0.938	1

QUADAS = Quality Assessment Tool for Diagnostic Accuracy Studies.

Prevalence

Based on one ED-based prospective study published in 1997 and conducted in Taiwan, the best-estimate prevalence of nongonococcal septic arthritis in ED patients with monoarticular arthritis is 27% (95% confidence interval [CI] = 17% to 38%).^{4,7} Other estimates of prevalence in research cohorts that are either undefined or a mix of medicine, surgery, and rheumatology clinic patients ranged from 0.4% to 45%.^{4,25,41,45,53,58,59} Prospective and retrospective designs appeared to offer similar ranges of septic arthritis prevalence. One study prospectively assessed the prevalence of septic arthritis in Rwandan HIV patients, noting a 0.5% prevalence.⁵ Estimates of treated septic arthritis-related short-term mortality have ranged from 3% to 11%.^{4,7,8,42}

History

Risk factors from the history and physical examination are listed in Table 2, along with estimates of their diagnostic accuracy.^{2,4} Except for HIV as a risk factor, these

were derived from a prospective analysis of almost 5,000 Dutch rheumatology patients.⁴ The interval between joint surgery and septic arthritis exceeded 3 months in one-third of cases. None of these findings significantly reduce the probability of septic arthritis in their absence. HIV was assessed as a risk factor in Rwanda, but was not a significant predictor of septic arthritis.⁵ Several other trials assessed different components of the history in synovial culture-positive patients alone, so that specificities and LRs could not be evaluated. Joint pain (sensitivity range = 85% to 100%) and tenderness (sensitivity 100% in a single study) may be sufficiently sensitive, but their overall diagnostic accuracy (ability to alter pretest probability) remains uncertain without data on specificities.⁶⁰

Physical Examination

With the exception of fever, no studies have evaluated physical examination findings for all patients with suspected septic arthritis, so specificities and LRs

Table 2
Risk Factors for Septic Arthritis from History and Physical Examination

Risk factor	Sensitivity, %	Specificity, %	+LR (95% CI)	-LR (95% CI)	OR
Age > 80 yr*	18.9	94.6	3.5 (1.7–6.4)	0.86 (0.70–0.96)	3.5
Diabetes mellitus*	10.8	96.0	2.7 (1.1–6.2)	0.93 (0.79–1.0)	3.3
Rheumatoid arthritis*	67.6	72.5	2.5 (1.9–2.9)	0.45 (0.27–0.67)	4.0
Joint surgery (<3 months ago)*	24.0	96.5	6.9 (3.7–11.6)	0.78 (0.63–0.90)	5.1
Hip or knee prosthesis*	35.1	88.6	3.1 (1.9–4.5)	0.73 (0.55–0.88)	15.0
Skin infection (no prosthesis)*	32.4	88.4	2.8 (1.7–4.2)	0.76 (0.58–0.91)	27.2
Prosthesis and skin infection*	24.3	98.4	15.0 (8.0–26.0)	0.77 (0.62–0.88)	72.7
HIV infection	75.0	38.8	1.2 (0.76–1.5)	0.64 (0.23–1.37)	N/A
Rigors	16.0–21.0	N/A	N/A	N/A	N/A
Subjective fever	44.0–97.0	N/A	N/A	N/A	N/A
Sweats	31.0	N/A	N/A	N/A	N/A
Pain affected joint	85.0	N/A	N/A	N/A	N/A
New joint swelling	77.0	N/A	N/A	N/A	N/A
Physical examination					
Pain with motion	100.0	N/A	N/A	N/A	N/A
Limited motion	92.0	N/A	N/A	N/A	N/A
Tender	68.0–100	N/A	N/A	N/A	N/A
Swelling	45.0–92.0	N/A	N/A	N/A	N/A
Joint effusion	92.0	N/A	N/A	N/A	N/A
Increased heat	18.0–92.0	N/A	N/A	N/A	N/A
Redness	13.0–64.0	N/A	N/A	N/A	N/A
Fever > 37.5°C	34.0–54.0	N/A	N/A	N/A	N/A
Axial load pain	36.0	N/A	N/A	N/A	N/A

*These history risk factors are derived from Kaandorp 1995.⁴
HIV = human immunodeficiency virus; +LR = positive likelihood ratio; -LR = negative likelihood ratio; N/A = not available because not reported.

cannot be reported. Several studies have described objective examination findings in patients with confirmed septic arthritis (Table 2). The definition of fever has varied from study to study, ranging from >37.5 to >39°C. Goldenberg and Cohen³⁶ noted 90% sensitivity for fever within the first 48 hours, but did not specify whether subjects were febrile at the initial presentation. Kortekangas et al.⁴⁵ assessed fever for bacterial arthritis, probable bacterial arthritis, and reactive arthritis patients. Combining the first two categories into one category labeled bacterial arthritis, compared with reactive arthritis, yields a positive likelihood ratio (+LR) of 0.89 and a negative likelihood ratio (-LR) of 1.2. No studies assessed the parameters, diagnostic accuracy, or reliability of clinical gestalt.

Serum Tests

Two studies provided estimates of sensitivity and specificity of leukocytosis for the diagnosis of septic arthritis.^{47,53} Using varying thresholds to define leukocytosis, six additional studies assessed the sensitivity of an elevated peripheral white blood cell (WBC) count (Table 3). Although Deesomchok and Tumrasvin⁴² and McCutchan and Fisher⁴³ did not define leukocytosis, they reported sensitivities of 80 and 57%, respectively. Regardless of the threshold selected, no study demonstrated an acceptable sensitivity or overall diagnostic accuracy of peripheral WBC count for septic arthritis.

Multiple studies demonstrated acceptable sensitivity for erythrocyte sedimentation rate (ESR) of >30 mm/hour, but the results were not consistent across trials and, when reported, the specificities were uniformly poor (Table 3). No cutoff for ESR or C-reactive protein

(CRP) significantly increases or decreases the posttest probability of septic arthritis. Li et al.⁵⁰ demonstrated a sensitivity of 96% for ESR of >30 mm/hour. In another retrospective review, Li et al.⁵³ demonstrated an ESR sensitivity of 75% and a specificity of 11% using ESR of >20 mm/hour as the definition of abnormal. Using ESR of >15 mm/hour, and CRP of >0.8 mg/L, Ernst et al.⁵⁹ demonstrated sensitivities of 66 and 90%, respectively. Söderquist et al.⁹ demonstrated that both ESR (mean = 81 mm/hour vs. 54 mm/hour) and CRP (mean = 182 mg/L vs. 101 mg/L) were significantly higher on admission in patients with septic arthritis than in those with crystalloid arthropathy.

Procalcitonin, tumor necrosis factor, and various cytokines including interleukin (IL)-6 and IL- β , were generally specific with very poor sensitivity. Procalcitonin levels are generally elevated when the etiology of septic arthritis is systemic rather than local.⁵¹ Four studies assessed the sensitivity of blood cultures, which ranged from 23% to 36%, but no research evaluated the specificity.^{5,8,9,44}

Synovial Tests

Gross inspection of synovial fluid by a rheumatologist is 94% sensitive and 58% specific for differentiating inflammatory and noninflammatory causes of acute arthritis.⁵² As summarized in Table 4 and Figure 2, the summary +LR for a sWBC count of $>50 \times 10^9/L$ is 4.7 (95% CI = 2.5 to 85), and the -LR is 0.52 (95% CI = 0.38 to 0.72), while for a sWBC count of $>100 \times 10^9/L$ the +LR is 13.2 (95% CI = 3.6 to 51) and -LR is 0.83 (95% CI = 0.80 to 0.89). Significant heterogeneity ($I^2 > 25%$) was noted between trials for sensitivity ($I^2 = 53%$, 70%)

Table 3
Serum Inflammatory Markers

Serum Marker	Sensitivity, %	Specificity, %	+LR	-LR
WBC count				
>10 × 10 ⁹ /L				
Argen 1966 ³⁵	74	N/A	N/A	N/A
Goldenberg 1976 ³⁶	60	N/A	N/A	N/A
Schlapbach 1990 ⁴⁴	42	N/A	N/A	N/A
Jeng 1997 ⁴⁷	90	36	1.4	0.28
Li 2004 ⁵⁰	62	N/A	N/A	N/A
>11 × 10 ⁹ /L				
Li 2007 ⁵³	75	55	1.7	0.84
>14 × 10 ⁹ /L				
Cooper 1986 ⁸	23	N/A	N/A	N/A
Söderquist 1998 ⁹	30	N/A	N/A	N/A
ESR				
>15 mm/hour				
Ernst 2010 ⁵⁹	66	48	1.3	0.71
>20 mm/hour				
Li 2007 ⁵³	75	11	0.84	2.4
>30 mm/hour				
Schlapbach 1990 ⁴⁴	76	N/A	N/A	N/A
Jeng 1997 ⁴⁷	95	29	1.3	0.17
Li 2004 ⁵⁰	97	N/A	N/A	N/A
>50 mm/hour				
Cooper 1986 ⁸	42	N/A	N/A	N/A
Li 2004 ⁵⁰	92	N/A	N/A	N/A
Martinot 2005 ⁵¹	82	42	1.4	0.4
>100 mm/hour				
Cooper 1986 ⁸	18	N/A	N/A	N/A
Li 2004 ⁵⁰	59	N/A	N/A	N/A
Martinot 2005 ⁵¹	46	94	7.0	0.6
CRP				
>10 mg/L				
Fottner 2008 ⁵⁵	87	39	1.4	0.3
Ernst 2010 ⁵⁹	91	15	1.1	0.6
>100 mg/L				
Söderquist 1998 ⁹	83	27	1.1	0.6
Martinot 2005 ⁵¹	82	70	2.8	0.3
>150 mg/L				
Martinot 2005 ⁵¹	73	83	4.5	0.3
> 200 mg/L				
Söderquist 1998 ⁹	44	85	2.9	0.7
Procalcitonin				
>0.3 ng/mL				
Martinot 2005 ⁵¹	73	94	11	0.3
Fottner 2008 ⁵⁵	73	94	13	0.3
>0.5 ng/mL				
Martinot 2005 ⁵¹	54	94	8	0.5
Fottner 2008 ⁵⁵	53	100	Infinity	0.5
>0.7 ng/mL				
Martinot 2005 ⁵¹	54	94	8	0.5
Fottner 2008 ⁵⁵	47	100	Infinity	0.5
>1 ng/mL				
Söderquist 1998 ⁹	34	93	5.1	0.7
Fottner 2008 ⁵⁵	20	100	Infinity	0.8
TNF- α (pg/mL)				
Söderquist 1998 ⁹	30	100	Infinity	0.7
IL-6 (pg/mL)				
Söderquist 1998 ⁹	26	83	1.5	0.9
IL- β (pg/mL)				
Söderquist 1998 ⁹	21	93	3.2	0.8

CRP = C-reactive protein; ESR = erythrocyte sedimentation rate; IL = interleukin; +LR = positive likelihood ratio; -LR = negative likelihood ratio; TNF- α = tumor necrosis factor alpha; WBC = white blood cell.

and specificity ($I^2 = 70\%$, 68%) for a sWBC count of $>50 \times 10^9$ and $>100 \times 10^9$, respectively. Sensitivity analysis was performed for a sWBC count of $>50 \times 10^9$ by sequentially excluding each trial and recomputing summary sensitivity and specificity. Exclusion of the

Kortenagas et al.⁴⁵ 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.,⁴⁵ Soderquist et al.,⁹ and Schmerling et al.²⁵ trials

Table 4
sWBC

Synovial Marker (sWBC)	Sensitivity, %	Specificity, %	+LR	-LR
>1.7 × 10 ⁹ /L				
Trampuz 2004 ⁴⁹	94	88	7.8	0.07
>10 × 10 ⁹ /L				
Jeng 1997 ⁴⁷	90	36	1.3	0.28
Trampuz 2004 ⁴⁹	71	98	34.9	0.30
>17.5 × 10 ⁹ /L				
Li 2007 ⁵³	83	67	2.5	0.25
>20 × 10 ⁹ /L				
Schlapbach 1990 ⁴⁴	78	N/A	N/A	N/A
Kortekangas 1992 ⁴⁵	63	60	1.6	0.63
Martinot 2005 ⁵¹	100	71	3.4	<0.001
>25 × 10 ⁹ /L				
McCutchan 1990 ⁴³	66	N/A	N/A	N/A
Schmerling 1990 ²⁵	70	84	4.3	0.35
Söderquist 1998 ⁹	76	58	1.8	0.42
Li 2004 ⁵⁰	79	N/A	N/A	N/A
<i>Summary estimate</i>	73 (64–81)	77 (73–81)	3.2 (2.3–4.4)	0.35 (0.23–0.50)
>50 × 10 ⁹ /L				
Krey 1979 ³⁸	70	87	5.5	0.34
McCutchan 1990 ⁴³	32	N/A	N/A	N/A
Schlapbach 1990 ⁴⁴	52	N/A	N/A	N/A
Schmerling 1990 ²⁵	63	97	19.3	0.38
Kortekangas 1992 ⁴⁵	31	75	1.3	0.92
Gratacós 1995 ⁴⁶	55	94	9.0	0.48
Söderquist 1998 ⁹	58	74	2.2	0.57
Li 2004 ⁵⁰	62	N/A	N/A	N/A
Martinot 2005 ⁵¹	50	86	3.4	0.54
McGillicuddy 2007 ⁵⁴	61	N/A	N/A	N/A
Li 2007 ⁵³	50	88	4.0	0.57
<i>Summary estimate</i>	56 (49–63)	90 (88–92)	4.7 (2.5–8.5)	0.52 (0.38–0.72)
>100 × 10 ⁹ /L				
Schmerling 1990 ²⁵	18	100	Infinity	0.88
McCutchan 1990 ⁴³	10	N/A	N/A	N/A
Kortekangas 1992 ⁴⁵	6	100	Infinity	0.94
Söderquist 1998 ⁹	30	94	4.7	0.75
Li 2004 ⁵⁰	31	N/A	N/A	N/A
<i>Summary estimate</i>	19 (14–20)	99 (96–100)	13.2 (3.6–51.1)	0.83 (0.80–0.89)

+LR = positive likelihood ratio; -LR = negative likelihood ratio; sWBC = synovial white blood cells.

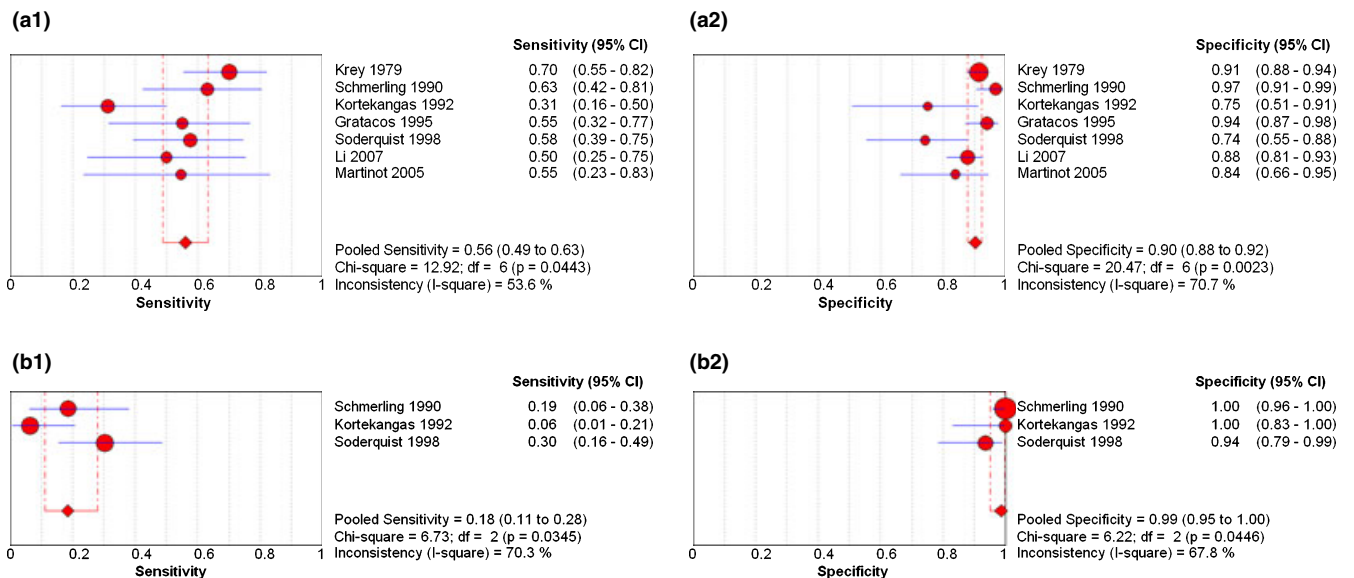


Figure 2. Meta-analysis. (A) sWBC count ≥ 50 × 10⁹/L. (B) sWBC count ≥ 100 × 10⁹/L. sWBC = synovial white blood cell.

Table 5
Other Synovial Tests

Synovial Marker	Sensitivity, %	Specificity, %	+LR	-LR
Synovial PMNs > 90%				
Schmerling 1990 ²⁵	59	83	3.4	0.49
Kortekangas 1992 ⁴⁵	54	60	1.4	0.76
Gratacós 1995 ⁴⁶	70	73	2.6	0.41
Summary estimate	60 (51–68)	78 (75–80)	2.7 (2.1–3.5)	0.51 (0.39–0.65)
Synovial glucose				
Schmerling 1990 ²⁵	56	85	3.7	0.52
Söderquist 1998 ⁹	64	85	4.2	0.43
Synovial protein > 30 g/L				
Schmerling 1990 ²⁵	50	47	0.94	1.1
Synovial LDH > 250 U/L				
Schmerling 1990 ²⁵	100	51	2.0	0
Synovial LDH > 600 U/L				
Schmerling 1990 ²⁵	60	68	1.9	0.59
Synovial lactate				
Brook 1978 ³⁷				
>5.6 mmol/L	67	72	2.4	0.46
>11 mmol/L	55	100	Infinity	0.45
Mossman 1981 ³⁹				
>10 mmol/L	86	100	Infinity	0.14 (0.14–0.31)
Riordan 1982 ⁴¹				
>12 mmol/L	100	95	19	0 (0–0.16)
Gratacós 1995 ⁴⁶				
>0.05 mmol/L	85	96	21	0.16

LDH = lactate dehydrogenase; +LR = positive likelihood ratio; -LR = negative likelihood ratio; PMN = polymorphonuclear leukocytes.

($I^2 = 27\%$) with a summary estimate of 91% for specificity. Trampuz et al.⁴⁹ demonstrated that prosthetic knee septic arthritis produce lower sWBC counts than native joint infections.

There were four trials reporting sWBC counts with sufficient detail and homogeneous patient populations to compute interval LR_s.^{9,25,38,45} The interval LR was 0.33 for the range 0–25 × 10⁹/L, 1.06 for the range 25 × 10⁹–50 × 10⁹/L, and 3.59 for the range 50 × 10⁹–100 × 10⁹/L.

Synovial polymorphonuclear cells greater than 90% does not significantly increase or decrease the probability of septic arthritis (Table 5). Seven trials assessed the sensitivity of synovial Gram stains to diagnose septic arthritis, with point estimates ranging from 29% to 65%.^{35,36,41,43,48,54,61} No trials have assessed the specificity of synovial Gram stains.

Synovial glucose and protein levels do not significantly increase or decrease the posttest probability of septic arthritis (Table 5). Söderquist et al.⁹ defined an abnormal synovial glucose as less than 1.5 mmol/L, or greater than 2.5 mmol/L discrepancy with a concurrent serum glucose specimen, whereas Schmerling et al.²⁵ used the definition of synovial glucose less than 4 mmol/L. Based on a single trial, a synovial lactate dehydrogenase (LDH) of <250 U/L may be sufficient to exclude the diagnosis of septic arthritis.²⁵

Four trials using varying techniques and cutoff values have demonstrated excellent diagnostic accuracy for synovial lactate. Brook et al.,³⁷ Mossman et al.,³⁹ and Riordan et al.⁴¹ used liquid chromatography to assess lactate levels. Each demonstrated high +LRs for synovial lactate levels above 10 mmol/L with reasonable -LRs (Table 5). Gratacós et al.⁴⁶ used spectropho-

tometry to assess D-lactate isomers that are only produced by certain bacterial species and demonstrated +LR and -LR of 21 and 0.16, respectively, although their population did not include any gonorrhoea or Gram-negative cases. Wiener et al.⁵⁷ used single-voxel ¹H magnetic resonance spectroscopy to measure synovial lactate concentrations in vitro with moderate correlation.

Yang et al.⁵⁸ used polymerase chain reaction (PCR) pathogen-specific probes and Gram stains for 36 relevant organisms to assess synovial fluid from a cohort of patients including ED patients. This uniprobe demonstrated a +LR 31.7 (95% CI = 14.3 to 45.3) and a -LR 0.05 (95% CI = 0.009 to 0.189) and was able to provide organism-specific results within 3 hours.

Test-Treatment Threshold Estimates

The management options for septic arthritis are needle, arthroscopic, or open drainage of the affected joint, in addition to 8 to 10 weeks of antibiotics.²¹ One recent systematic review of therapy for septic arthritis failed to identify any high-quality data to favor conservative or surgical management.¹⁶ One prospective multicenter report conducted over 2 years identified an 11% mortality in treated septic arthritis patients at a median of 39 days after admission.⁷ Lacking any randomized controlled trials of 10 weeks of antibiotics in septic arthritis patients, the literature for chronic osteomyelitis was reviewed because it incorporates similar antimicrobials for an equal length of time and a closely related pathology. Antibiotic therapy for osteomyelitis is associated with moderate or severe adverse events in 4.8% of oral antibiotic patients and 15.5% of parenteral antibiotic patients.⁶²

$$T_{\text{testing threshold}} = [(P_{\text{pos/nd}}) \times (R_{\text{rx}}) + R_t] \div [(P_{\text{pos/nd}} \times R_{\text{rx}}) + (P_{\text{pos/d}} \times B_{\text{rx}})] = 0.052$$

$$T_{\text{treatment threshold}} = [(P_{\text{neg/nd}}) \times (R_{\text{rx}}) - R_t] \div [(P_{\text{neg/nd}} \times R_{\text{rx}}) + (P_{\text{neg/d}} \times B_{\text{rx}})] = 0.387$$

Where assumptions are based upon the summary estimates for sWBC count > 50 × 10⁹ cells/L from Table 4.

$P_{\text{pos/nd}}$ = probability of a positive result in patients without disease = 1-specificity = 1-0.9 = 0.10*

$P_{\text{neg/nd}}$ = probability of a negative result in patients without disease = specificity = 0.90*

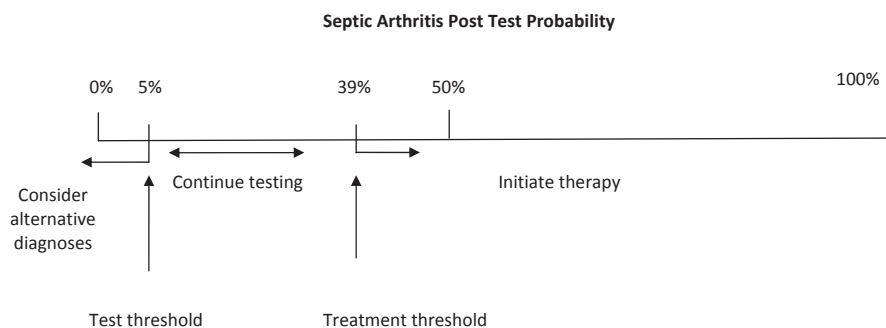
R_{rx} = risk of treatment in patients without disease = 0.155

R_t = risk of diagnostic test = 0.00037

$P_{\text{pos/d}}$ = probability of a positive result in patients with disease = sensitivity = 0.56*

$P_{\text{neg/d}}$ = probability of a negative result in patients with disease = 1 - sensitivity = 1-0.54= 0.44*

B_{rx} = benefit of treatment in patients with disease = 0.50†



* Using sWBC > 50 × 10⁹/L.

† Assumes NNT = 2 for joint debridement plus antibiotics to reduce short-term mortality. This is an estimate since no randomized controlled trials were identified. One prospective trial provided an estimated 11% mortality of treated septic arthritis at a median of 39 days. If the mortality of untreated septic arthritis was 61%, then the NNT would be 2.

Figure 3. Test and treatment threshold formulas. NNT = number needed to treat; sWBC = synovial white blood cell.

These summary estimates (or point estimates for single trials) were used to compute the treatment and test thresholds as illustrated in Figure 3. Using the summary estimate of sWBC counts of >50 × 10⁹/L (see Table 4; sensitivity 56%, specificity 90%), for example, the evidence-based assumptions yield a test threshold of 5.2% and a treatment threshold of 38.7%. In other words, if the clinical probability of septic arthritis is below 5%, then further testing may harm more patients than it helps. Similarly, if the probability of septic arthritis is above 39%, treatment should be started instead of additional diagnostic testing, or patient harm may outweigh the yield of additional diagnostic tests. The test threshold approximates previously described analyses amongst rheumatologists where the threshold for potential septic arthritis was lower than for gout (20% vs. 65%).⁶³ Using the online calculator associated with this article (in development), the various literature-based assumptions can be altered to determine how the test and treatment thresholds would vary. For example, because no randomized controlled trials exist by which to estimate the benefit of surgical and antibiotic management as a whole, we estimated a 50% absolute risk reduction for our computations. If the absolute risk

reduction were instead 10% ($B_{\text{rx}} = 0.1$), then the test threshold would increase to 22%, and the treatment threshold to 76%. Similarly, the other estimates (risk of arthrocentesis, risk of antibiotics, sensitivity, and specificity for various diagnostic findings) can be adjusted to define thresholds in accordance with one's perception of the true values for these parameters.

DISCUSSION

The overall quality of evidence for the diagnosis of non-gonococcal septic arthritis is relatively low. The majority of the studies are retrospective hospital-based case series derived from administrative data with no control group. Verification bias occurs when patients with positive index test results are more likely to have criterion standard testing performed. Studies that only enroll patients with confirmed diagnoses may suffer from this bias, which falsely increases sensitivity and decreases specificity.³³

Another often unstated bias of septic arthritis diagnostic accuracy studies is spectrum bias. When the population studied is skewed toward higher-severity illness, sensitivity can be falsely elevated. On the

other hand, healthier populations will falsely elevate specificity.³³ Therefore, recognizing disease prevalence and the specific population evaluated in each study is imperative to appropriately translate diagnostic test performance in other clinical settings.⁶⁴ For example, estimates of diagnostic accuracy for elements of the history can only be derived from a single Dutch trial of rheumatology patients, limiting external validity for ED populations. In these patients, no finding from the history significantly *decreases* the probability of septic arthritis. Only recent joint surgery or a joint prosthesis with overlying skin infection significantly *increases* the risk of septic arthritis.

Overall, serum inflammatory markers are not useful acutely (Table 3). Because history, physical examination, and serum markers are not helpful to significantly adjust the posttest probability of septic arthritis, synovial fluid analysis is essential for the diagnosis. Although generally considered safe, arthrocentesis is not a risk-free procedure. One textbook of emergency procedures estimates that the risk of iatrogenic infection from arthrocentesis is 0.01% in the general population and as high as 0.05% in immunocompromised patients.²³ One retrospective review in Iceland suggested that the incidence of postarthrocentesis iatrogenic septic arthritis was 0.037%.⁶⁵ Another risk is bleeding. However, arthrocentesis can be safely performed even in patients with therapeutic warfarin anticoagulation with a risk of clinically significant hemorrhage below 10%.⁶⁶

Textbooks suggest the following guidelines to interpret synovial WBC counts: 2×10^3 to 2×10^9 /L noninflammatory, 2×10^9 to 50×10^9 /L inflammatory, and 2×10^9 to $>50 \times 10^9$ /L infectious.^{23,67} Although 13 trials assessed sensitivity, only seven trials have assessed both the sensitivity and the specificity of sWBC count, and significant heterogeneity was noted (Figure 2). Sources of heterogeneity could include the populations studied, interval between disease onset and synovial testing, personnel or equipment used to quantify cell counts, or criterion standard used. For example, manually counting sWBC may be less accurate than automated approaches.^{68,69}

Our meta-analysis suggests summary +LRs of 4.7 (95% CI = 2.5 to 8.5) and 13.2 (95% CI = 3.6 to 51.1) for a sWBC count of $>50 \times 10^9$ or $>100 \times 10^9$ /L, respectively (Table 4). In other words, if one starts with a pretest probability for septic arthritis of 27%, and the sWBC count is $>50 \times 10^9$ but $< 100 \times 10^9$ /L, then the posttest probability would increase to 64%. If the sWBC count was $>100 \times 10^9$ /L, then the posttest probability would increase to 83%.¹¹ However, if the sWBC count is 75×10^9 /L, should one use the +LR for a sWBC count of $>50 \times 10^9$ /L or the -LR for sWBC count of $<100 \times 10^9$ /L? This conundrum can be eliminated by using the interval LR of 3.59 for a sWBC count between 50×10^9 and 100×10^9 /L. Interval LR's minimize the waste of valuable diagnostic detail that occurs when continuous data are artificially dichotomized for clinical convenience.⁷⁰ Interval LR's also tend to reduce the biased estimates of accuracy yielded by lower quality studies.⁷¹ Using the interval LR for 0 to 25×10^9 /L (0.33), 25×10^9 – 50×10^9 /L (1.06), 50×10^9 to 100×10^9 /L (3.59), or $>100 \times 10^9$ /L (infinity) would change a pretest

probability for septic arthritis of 27 to 11, 28, 57, and near 100%, respectively. The sWBC count should not be used in isolation to rule in or rule out the diagnosis of septic arthritis, but should augment the entire clinical evaluation. For example, the presence of organisms on Gram stain in a patient with a sWBC count of 25×10^9 – 50×10^9 /L should increase the probability of septic arthritis. Similarly, the presence of synovial crystals in a patient with a sWBC count of 25×10^9 – 50×10^9 /L should reduce the probability of septic arthritis.

However, clinicians need to be wary of the limitations of routine synovial tests such as Gram stains and crystal analysis. A negative Gram stain does not by itself rule out septic arthritis since the sensitivity of synovial fluid Gram stain is poor, with 45% to 71% false-negative rates, and the specificity remains undefined.^{35,36,41,43,48,54} Although synovial fluid should be evaluated for uric acid or calcium pyrophosphate crystals, since the differential diagnosis includes crystalloid arthropathies, septic arthritis can rarely coexist with gout or pseudogout, in about 1.5% of cases.⁷² Therefore, synovial fluid culture is indicated even when crystals are identified, sWBC count is lower, and Gram stain reveals no organisms.⁷³

Several recent trials have assessed newer synovial biomarkers that may be useful to diagnose septic arthritis in the ED. Two trials have evaluated a low synovial glucose (defined as a serum-to-synovial fluid glucose ratio < 0.5 , or synovial fluid glucose < 27 g/dL), and one trial each has evaluated LDH, protein, and tumor necrosis factor alpha (TNF- α).^{9,25,38,45,47} Besides LDH, none of these synovial tests effectively discriminate septic arthritis from other etiologies of acute arthritis. On the other hand, synovial lactate has consistently demonstrated desirable diagnostic properties to rule in septic arthritis using a threshold of >10 mmol/L. No studies to date have evaluated the diagnostic accuracy or effect on ED management of septic arthritis using point-of-care lactate assays. Future investigators might keep in mind that sepsis investigators using point-of-care lactate testing have demonstrated accurate and reliable results, while saving a mean of 151 minutes in ED settings.^{74,75}

No CDRs have been derived to assist with the diagnosis of adult septic arthritis. Single-center validation of a pediatric CDR identified four variables associated with increased risk of hip septic arthritis with a receiver operating characteristic curve area under the curve of 0.86: history of a fever, non-weight-bearing, ESR of >40 mm/hour, and a serum WBC count of $>12 \times 10^6$ cells/L.⁷⁶ Similar methodology could be used to derive a decision aid for adult septic arthritis.

Recently, opinion leaders have demanded full incorporation of modern Web-based technology to transform static literature into dynamic research.⁷⁷ By assessing one model to quantitatively estimate theoretical test-treatment thresholds using best-evidence approaches to locate and summarize risks and benefits for one clinical challenge, we have taken one step toward a more dynamic style of reporting diagnostic science. Specifically, we have provided an interactive calculator (available as supporting information in the online version of

this paper) that permits critical readers to adjust the best-evidence assumptions that our analysis has provided as future research findings become available.

One common criticism of diagnostic evidence-based medicine is that clinicians are not cognitively wired to think as Bayesian decision-makers.^{78,79} In fact, Campbell et al.⁸⁰ hypothesized seven cognitive biases related to prevalence misperceptions, including base-rate neglect and “gambler’s fallacy.” By providing the range of disease prevalence within the context of heterogeneous patient populations and potential ascertainment bias, we have empowered astute bedside clinicians to use a reasonable estimate of septic arthritis baseline prevalence by which to adjust diagnostic test-driven probabilities using the summary LRs.^{33,81,82} In addition to recognizing which tests facilitate the diagnosis of septic arthritis (sWBC count $> 50 \times 10^9/L$, synovial lactate, possibly synovial probe-based PCR), we are hopeful that our review will also quantitatively highlight those tests which do *not* alter posttest disease probability (WBC, ESR, CRP, synovial glucose, synovial protein) and help clinicians recognize optimally efficient ordering strategies.⁸³

Implications for Future Research

Diagnostic research can be fraught with bias.^{84–87} Our review of the septic arthritis diagnostic literature revealed overall poor to moderate quality trial designs. Basic research considerations such as defining the patient population and reporting the criterion standard employed were often lacking. More complex research considerations, such as reporting the interval between the index test and the criterion standard, blinding outcome assessors to ancillary clinical data (including the index tests being evaluated), and providing sufficient detail to reconstruct two-by-two tables, were virtually never reported. Future investigators need to be cognizant of these biases and should incorporate the QUADAS instrument and the Standards for Reporting of Diagnostic Accuracy Studies guidelines when designing new research protocols.^{19,85,88} Based on our findings, we suggest the following priorities for future septic arthritis diagnostic research.

First, more ED-based septic arthritis diagnostic and therapeutic trials are needed. The ideal study would prospectively recruit patients suspected of having septic arthritis before the diagnosis is established to assess clinicians’ pretest probability estimates along with the elements and overall accuracy of emergency provider gestalt. These investigators should blind clinicians who obtain the index tests (history, physical examination, labs) from the criterion standard. The individuals who categorize patients into “septic arthritis” or “not septic arthritis” subsets should be unaware of the constellation of clinical findings used by bedside clinicians. All of these design-related details should be explicitly described in the article.¹⁹

Second, in analyzing and reporting data, investigators should at a minimum report LRs, including interval LRs for continuous data,^{32,33} since sensitivity/specificity have multiple limitations to bedside application.⁶⁰ Histograms and scatter plots allow readers to reconstruct LRs and interval LR estimates around alternative thresholds.

Third, patients, clinicians, and policy-makers ultimately expect medical testing to improve patient-centered outcomes. Traditional randomized controlled trials are expensive and time-consuming to conduct, particularly for rare diagnoses. Different trial designs should be considered for minimally biased yet still feasible diagnostic test impact evaluations.⁸⁹ Accordingly, criteria to prioritize diagnostic test randomized trials are now being defined.⁹⁰ These evolving criteria should be applied to the ED diagnosis of septic arthritis to determine the variables and outcomes of importance to optimize patient outcomes.

Finally, pending identification of new serum or synovial markers to distinguish septic arthritis from other sources of nontraumatic acute joint pain, the current review suggests three priorities for ED investigators: 1) confirm the accuracy and external validity of rapidly available synovial PCR techniques to identify the bacterial pathogen as was described in a single trial;⁵⁸ 2) assess the accuracy of point-of-care synovial fluid lactate assays; and 3) evaluate the potential for combinations of findings from history, physical examination, and bedside tests (i.e., a CDR) to enhance the clinician’s ability to rule in or rule out septic arthritis.⁹¹

LIMITATIONS

First, we limited our analysis to English-language articles and searched only two electronic sources. We sought to enhance our strategy by analyzing the bibliographies of selected articles and current textbooks within our field, but we may have missed older manuscripts or non-English research reports. Second, the overall quality of the evidence as judged by the QUADAS instrument was poor to moderate, so our summary estimates may be biased by multiple unmeasured confounders. Few trials reported sufficient detail to reconstruct two-by-two tables, and specificity was generally not assessed. Even among those trials that did report sensitivity and specificity, significant between-trial heterogeneity was noted, limiting the confidence in the meta-analysis results. Third, we excluded gonococcal arthritis patients from this review so the results cannot be extrapolated to this condition. Fourth, we were unable to identify any controlled trials to estimate treatment benefit for septic arthritis, so our test and treatment thresholds are derived from such data available from osteomyelitis trials. Although osteomyelitis is a closely related pathologic process with similar organisms and prognostic risk factors, we recognize that these treatment risk-benefit estimates are suboptimal. However, pending septic arthritis interventional trials to derive more accurate, disease-specific estimates, we believe that some information is better than no information. Furthermore, we have provided an interactive instrument by which readers can revise diagnostic and therapeutic risk-benefit estimates to recalculate test and treatment thresholds based on their preferred assumptions. Finally, our test and treatment thresholds do not incorporate patient preferences into the equation. Future modifications would optimally include patient-centric risk-benefit communication instruments and

algorithms to efficiently and effectively bring them into the decision-making process.

CONCLUSIONS

We estimate the prevalence of nongonococcal septic arthritis in adults presenting to the ED with acute monoarticular joint complaints to be approximately 27%. With the exception of recent joint surgery or cellulitis overlying a prosthetic knee or hip, the history, physical examination, and routine blood tests do not distinguish acute septic arthritis from other forms of arthritis. In other words, neither the presence nor the absence of these findings significantly changes the probability of septic arthritis. On the other hand, a synovial white blood cell count of $>50 \times 10^9/L$ can increase the probability of septic arthritis, while a synovial white blood cell count from 0 to $25 \times 10^9/L$ can reduce the probability of septic arthritis, and values of 25×10^9 to $50 \times 10^9/L$ require additional testing and perhaps empiric antibiotics pending definitive culture results. Future prospective trials are needed to understand the sensitivity, specificity, and positive and negative likelihood ratios for elements of the history and physical examination and point-of-care inflammatory synovial markers such as lactate, which may be useful to rule in or rule out septic arthritis. In the meantime, clinicians should be aware of the risk factors for nongonococcal septic arthritis in expeditiously selecting appropriate diagnostic and therapeutic options while consulting orthopedic surgery for early management when clinical evaluation remains less than definitive.

The authors acknowledge Dr. Andrew Worster, Dr. Eddy Lang, and Dr. David Cone for their substantial assistance in devising the flow and format for this manuscript. In addition, the authors acknowledge Susan Fowler, MLIS, for her assistance devising and employing the search strategy.

References

1. Goldenberg DL. Septic arthritis. *Lancet*. 1998; 351:197–202.
2. Margaretten ME, Kohlwes J, Moore D, Bent S. Does this adult patient have septic arthritis? *JAMA*. 2007; 297:1478–88.
3. Till SH, Snaith ML. Assessment, investigation, and management of acute monoarthritis. *J Accid Emerg Med*. 1999; 16:355–61.
4. Kaandorp CJ, van Schaardenburg D, Krijnen P, Habbema JD, van De Laar MA. Risk factors for septic arthritis in patients with joint disease. *Arthritis Rheum*. 1995; 38:1819–25.
5. Saraux A, Taelman H, Blanche P, et al. HIV infection as a risk factor for septic arthritis. *Br J Rheumatol*. 1997; 36:333–7.
6. Kaandorp CJ, Dinant HJ, van De Laar MA, Moens HJ, Prins AP, Dijkmans BA. Incidence and sources of native and prosthetic joint infection: a community based prospective survey. *Ann Rheum Dis*. 1997; 56:470–5.
7. Gupta MN, Sturrock RD, Field M. A prospective 2-year study of 75 patients with adult-onset septic arthritis. *Rheumatology*. 2001; 40:24–30.
8. Cooper C, Cawley MI. Bacterial arthritis in an English health district: a 10 year review. *Ann Rheum Dis*. 1986; 45:458–63.
9. Söderquist B, Jones I, Fredlund H, Vikerfors T. Bacterial or crystal-associated arthritis? Discriminating ability of serum inflammatory markers. *Scan J Infect Dis*. 1998; 30:591–6.
10. Freed JF, Nies KM, Boyer RS, Louie JS. Acute monoarticular arthritis: a diagnostic approach. *JAMA*. 1980; 243:2314–6.
11. Hayden SR, Brown MD. Likelihood ratio: a powerful tool for incorporating the results of a diagnostic test into clinical decision making. *Ann Emerg Med*. 1999; 33:575–80.
12. Zehtabchi S, Kline JA. The art and science of probabilistic decision-making in emergency medicine. *Acad Emerg Med*. 2010; 17:521–3.
13. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *N Engl J Med*. 1980; 302:1109–17.
14. Kline JA. Further illumination of the test threshold approach in the care of emergency department patients with symptoms of pulmonary embolism. *Ann Emerg Med*. 2010; 55:327–30.
15. Marsh RH, Brown DF, Nadel ES. Septic arthritis. *J Emerg Med*. 2006; 31:417–20.
16. Mathews CJ, Kingsley G, Field M, et al. Management of septic arthritis: a systematic review. *Ann Rheum Dis*. 2007; 66:440–5.
17. Holleman DR, Simel DL. Does the clinical examination predict airflow limitation? *JAMA*. 1995; 273:313–9.
18. Smetana GW, Schmerling RH. Does this patient have temporal arteritis? *JAMA*. 2002; 287:92–101.
19. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*. 2003; 3:e25.
20. Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Med Res Methodol*. 2006; 6:e9.
21. Adams BD, Lowery DW. Arthritis. In: Marx JA, Hockberger RS, Walls RM, et al. (eds). *Rosen's Emergency Medicine Concepts and Clinical Practice*. 7th ed. Philadelphia, PA: Mosby, 2010, pp 1477–80.
22. Burton JH. Acute disorders of the joints and bursae. In: Tintinalli JE, Stapczynski JS, Ma OJ, Cline DM, Cydulka RK, Meckler GD (eds). *Tintinalli's Emergency Medicine: A Comprehensive Study Guide*. 7th ed. New York, NY: McGraw-Hill, 2011, pp 1926–33.
23. Parrillo SJ, Morrison DS, Panacek EA. Arthrocentesis. In: Roberts JR, Hedges JR, Custalow CB, Chanmugam AS, Chudnofsky CR, McManus J (eds). *Clinical Procedures in Emergency Medicine*. 5th ed. Philadelphia, PA: Elsevier, 2010, pp 971–85.
24. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *Meta-analysis Of Observational*

- Studies in Epidemiology (MOOSE) group. *JAMA*. 2000; 283:2008–12.
25. Schmerling RH, Delbanco TL, Tosteson AN, Trentham DE. Synovial fluid tests: what should be ordered? *JAMA*. 1990; 264:1009–14.
 26. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull*. 1968; 70:213–20.
 27. Mulherin SA, Miller WC. Spectrum bias or spectrum effect? Subgroup variation in diagnostic test evaluation. *Ann Intern Med*. 2002; 137:598–602.
 28. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.9.0*. The Cochrane Collaboration, 2010.
 29. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol*. 2006; 6:e31.
 30. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986; 7:177–88.
 31. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol*. 2005; 58:882–93.
 32. Brown MD, Reeves MJ. Evidence-based emergency medicine skills for evidence-based emergency care. Interval likelihood ratios: another advantage for the evidence-based diagnostician. *Ann Emerg Med*. 2003; 42:292–7.
 33. Newman TB, Kohn MA. *Evidence-Based Diagnosis (Practical Guide to Biostatistics and Epidemiology)*. Cambridge, UK: Cambridge University Press, 2009.
 34. Ward J, Cohen AS, Bauer W. The diagnosis and therapy of acute suppurative arthritis. *Arthritis Rheum*. 1960; 3:522–35.
 35. Argen RJ, Wilson CH, Wood P. Suppurative arthritis. *Arch Intern Med*. 1966; 117:661–6.
 36. Goldenberg DL, Cohen AS. Acute infectious arthritis. A review of patients with nongonococcal joint infections (with emphasis on therapy and prognosis). *Am J Med*. 1976; 60:369–77.
 37. Brook I, Reza MJ, Bricknell KS, Bluestone R, Finegold SM. Synovial fluid lactic acid. A diagnostic aid in septic arthritis. *Arthritis Rheum*. 1978; 21:774–9.
 38. Krey PR, Bailen DA. Synovial fluid leukocytosis: a study of extremes. *Am J Med*. 1979; 67:436–42.
 39. Mossman SS, Coleman JM, Gow PJ. Synovial fluid lactic acid in septic arthritis. *N Z Med J*. 1981; 93:115–7.
 40. Rosenthal J, Bole GG, Robinson WD. Acute nongonococcal infectious arthritis. Evaluation of risk factors, therapy, and outcome. *Arthritis Rheum*. 1980; 23:889–97.
 41. Riordan T, Doyle D, Tabaqchali S. Synovial fluid lactic acid measurement in the diagnosis and management of septic arthritis. *J Clin Pathol*. 1982; 35:390–4.
 42. Deesomchok U, Tumrasvin T. Clinical study of culture-proven cases of non-gonococcal arthritis. *J Med Assoc Thai*. 1990; 73:615–23.
 43. McCutchan HJ, Fisher RC. Synovial leukocytosis in infectious arthritis. *Clin Orthop Relat Res*. 1990; 257:226–30.
 44. Schlapbach P, Ambord C, Blochlinger AM, Gerber NJ. Bacterial arthritis: are fever, rigors, leucocytosis and blood cultures of diagnostic value? *Clin Rheumatol*. 1990; 9:69–72.
 45. Kortekangas P, Aro HT, Tuominen J, Toivanen A. Synovial fluid leukocytosis in bacterial arthritis vs. reactive arthritis and rheumatoid arthritis in the adult knee. *Scan J Rheumatol*. 1992; 21:283–8.
 46. Gratacós J, Vila J, Moyá F, et al. D-lactic acid in synovial fluid. A rapid diagnostic test for bacterial synovitis. *J Rheumatol*. 1995; 22:1504–8.
 47. Jeng GW, Wang CR, Liu ST, et al. Measurement of synovial tumor necrosis factor-alpha in diagnosing emergency patients with bacterial arthritis. *Am J Emerg Med*. 1997; 15:626–9.
 48. Faraj AA, Omonbude OD, Godwin P. Gram staining in the diagnosis of acute septic arthritis. *Acta Orthop Belgica*. 2002; 68:388–91.
 49. Trampuz A, Hanssen AD, Osmon DR, Mandrekar J, Steckelberg JM, Patel R. Synovial fluid leukocyte count and differential for the diagnosis of prosthetic knee infection. *Am J Med*. 2004; 117:556–62.
 50. Li SF, Henderson J, Dickman E, Darzynkiewicz R. Laboratory tests in adults with monoarticular arthritis: can they rule out a septic joint. *Acad Emerg Med*. 2004; 11:276–80.
 51. Martinot M, Sordet C, Soubrier M, et al. Diagnostic value of serum and synovial procalcitonin in acute arthritis: a prospective study of 42 patients. *Clin Exp Rheumatol*. 2005; 23:303–10.
 52. Abdullah S, Young-Min SA, Hudson SJ, Kelly CA, Heycock CR, Hamilton JD. Gross synovial fluid analysis in the differential diagnosis of joint effusion. *J Clin Pathol*. 2007; 60:1144–7.
 53. Li SF, Cassidy C, Chang C, Gharib S, Torres J. Diagnostic utility of laboratory tests in septic arthritis. *Emerg Med J*. 2007; 24:75–7.
 54. McGillicuddy DC, Shah KH, Friedberg RP, Nathanson LA, Edlow JA. How sensitive is the synovial fluid white blood cell count in diagnosing septic arthritis? *Am J Emerg Med*. 2007; 25:749–52.
 55. Fottner A, Birkenmaier C, von Schulze-Pellengahr C, Wegener B, Jansson V. Can serum procalcitonin help to differentiate between septic and nonseptic arthritis? *Arthroscopy*. 2008; 24:229–33.
 56. Hügle T, Schuetz P, Mueller B, et al. Serum procalcitonin for discrimination between septic and non-septic arthritis. *Clin Exp Rheumatol*. 2008; 26:453–6.
 57. Wiener E, Zanetti M, Hodler J, Pfirrmann CW. Lactate and T₂ measurements of synovial aspirates at 1.5 T: differentiation of septic from non-septic arthritis. *Skeletal Radiol*. 2008; 37:743–8.
 58. Yang S, Ramachandran P, Hardick A, et al. Rapid PCR-based diagnosis of septic arthritis by early gram-type classification and pathogen identification. *J Clin Microbiol*. 2008; 46:1386–90.
 59. Ernst AA, Weiss SJ, Tracy LA, Weiss NR. Usefulness of CRP and ESR in predicting septic joints. *South Med J*. 2010; 103:522–6.
 60. Gallagher EJ. Evidence-based emergency medicine editorial. The problem with sensitivity and specificity. *Ann Emerg Med*. 2003; 42:298–303.

61. Belhaj A, Badawi A, Lee CH, Tweedle L. Utility of gram stain and cell counts for diagnosing septic arthritis. *Int J Antimicrob Agents*. 2009; 34(Suppl 2):S2.
62. Conterno LO, da Silva Filho CR. Antibiotics for treating chronic osteomyelitis in adults. *Cochrane Database Syst Rev*. 2009; 3:CD004439.
63. Eisenberg JM, Schumacher HR, Davidson PK, Kaufmann L. Usefulness of synovial fluid analysis in the evaluation of joint effusions. Use of threshold analysis and likelihood ratios to assess a diagnostic test. *Arch Intern Med*. 1984; 144:715–9.
64. Leeflang MM, Bossuyt PM, Irwig L. Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis. *J Clin Epidemiol*. 2009; 62:5–12.
65. Geirsson AJ, Statkevicius S, Vikingsson A. Septic arthritis in Iceland 1990–2002: increasing incidence due to iatrogenic infections. *Ann Rheum Dis*. 2008; 67:638–43.
66. Thumboo J, O'Duffy JD. A prospective study of the safety of joint and soft tissue aspirations and injections in patients taking warfarin sodium. *Arthritis Rheum*. 1998; 41:736–9.
67. El-Gabalawy HS. Diagnostic tests and procedures in rheumatic diseases: synovial fluid analysis, synovial biopsy, and synovial pathology. In: Firestein GS, Budd RC, Harris ED, McInnes IB, Ruddy S, Sargent JS (eds). *Kelly's Textbook of Rheumatology*. 8th ed. Philadelphia, PA: WB Saunders, 2008.
68. Schumacher HR, Sieck MS, Rothfuss S, et al. Reproducibility of synovial fluid analyses. A study among four laboratories. *Arthritis Rheum*. 1986; 29:770–4.
69. de Jonge R, Brouwer R, Smit M, et al. Automated counting of white blood cells in synovial fluid. *Rheumatology*. 2004; 43:170–3.
70. Berger VW. Improving the information content of categorical clinical trial endpoints. *Control Clin Trials*. 2002; 23:502–14.
71. Berger VW, Alperson SY. A general framework for the evaluation of clinical trial quality. *Rev Recent Clin Trials*. 2009; 4:79–88.
72. Shah K, Spear J, Nathanson LA, McCauley J, Edlow JA. Does the presence of crystal arthritis rule out septic arthritis? *J Emerg Med*. 2007; 32:23–6.
73. Yu KH, Luo SF, Liou LB, et al. Concomitant septic and gouty arthritis—an analysis of 30 cases. *Rheumatology*. 2003; 42:1062–6.
74. Goyal M, Pines JM, Drumheller BC, Gaijeski DF. Point-of-care testing at triage decreases time to lactate level in septic patients. *J Emerg Med*. 2010; 38:578–81.
75. Shapiro NI, Fisher C, Donnino M, et al. The feasibility and accuracy of point-of-care lactate measurement in emergency department patients with suspected infection. *J Emerg Med*. 2010; 39:89–94.
76. Kocher MS, Mandiga R, Zurakowski D, Barnewell C, Kasser JR. Validation of a clinical prediction rule for the differentiation between septic arthritis and transient synovitis of the hip in children. *J Bone Joint Surg Am*. 2004; 86-A:1629–35.
77. Schriger DL. Dynamic research, static literature. *Ann Emerg Med*. 2010; 56:339–40.
78. Cahan A, Gilon D, Manor O, Paltiel O. Probabilistic reasoning and clinical decision-making: do doctors overestimate diagnostic probabilities? *QJM*. 2003; 96:763–9.
79. Phelps MA, Levitt MA. Pretest probability estimates: a pitfall to the clinical utility of evidence-based medicine? *Acad Emerg Med*. 2004; 11:692–4.
80. Campbell SG, Croskerry P, Bond WF. Profiles in patient safety: a “perfect storm” in the emergency department. *Acad Emerg Med*. 2007; 14:743–9.
81. Richardson WS, Polashenski WA, Robbins BW. Could our pretest probabilities become evidence based? A prospective survey of hospital practice. *J Gen Intern Med*. 2003; 18:203–8.
82. Pines JM, Everett WW. *Evidence-Based Emergency Care: Diagnostic Testing and Clinical Decision Rules*. Oxford, UK: Blackwell Publishing, 2008.
83. Brownlee S. *Overtreated: Why Too Much Medicine Is Making Us Sicker and Poorer*. New York, NY: Bloomsbury USA, 2007.
84. Mower WR. Evaluating bias and variability in diagnostic test reports. *Ann Emerg Med*. 1999; 33:85–91.
85. Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA*. 1999; 282:1061–6.
86. Rutjes AW, Reitsma JB, Di Nisio M, Smidt N, van Rijn JC, Bossuyt PM. Evidence of bias and variation in diagnostic accuracy studies. *CMAJ*. 2006; 174:469–76.
87. Schünemann AH, Oxman AD, Brozek J, et al. GRADE: grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ*. 2008; 336:1106–10.
88. Bossuyt PM, Reitsma JB, Bruns DE, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med*. 2003; 138:W1–12.
89. Lijmer JG, Bossuyt PM. Various randomized designs can be used to evaluate medical tests. *J Clin Epidemiol*. 2009; 62:364–73.
90. Lord SJ, Irwig L, Simes RJ. When is measuring sensitivity and specificity sufficient to evaluate a diagnostic test, and when do we need randomized trials? *Ann Intern Med*. 2006; 144:850–5.
91. Stiell IG, Wells GA. Methodologic standards for the development of clinical decision rules in emergency medicine. *Ann Emerg Med*. 1999; 33:437–47.

Supporting Information

The following supporting information is available in the online version of this paper:

Data Supplement S1. Summary of included studies.

Data Supplement S2. Test treatment calculator.

The documents are in DOC and XLS formats.

Please note: Wiley Periodicals Inc. is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.