BMJ Open

Test characteristics of history, examination and investigations in the evaluation for septic arthritis in the child presenting with acute non-traumatic limp. A systematic review

Jacky Tu,[1,2] Peter Gowdie,[3,4] Julian Cassar,[5] Simon Craig ⓘ [4,6,7]

ABSTRACT

Background   Septic arthritis is an uncommon but potentially significant diagnosis to be considered when a child presents to the emergency department (ED) with non-traumatic limp. Our objective was to determine the diagnostic accuracy of clinical findings (history and examination) and investigation results (pathology tests and imaging) for the diagnosis of septic arthritis among children presenting with acute non-traumatic limp to the ED.

Methods   Systematic review of the literature published between 1966 and June 2019 on MEDLINE and EMBASE databases. Studies were included if they evaluated children presenting with lower limb complaints and evaluated diagnostic performance of items from history, physical examination, laboratory testing or radiological examination. Data were independently extracted by two authors, and quality assessment was performed using the Quality Assessment Tool for Diagnostic Accuracy Studies 2 tool.

Results   18 studies were identified, and included 2672 children (560 with a final diagnosis of septic arthritis). There was substantial heterogeneity in inclusion criteria, study setting, definitions of specific variables and the gold standard used to confirm septic arthritis. Clinical and investigation findings were reported using varying definitions and cut-offs, and applied to differing study populations. Spectrum bias and poor-to-moderate study design quality limit their applicability to the ED setting. Single studies suggest that the presence of joint tenderness (n=189; positive likelihood ratio 11.4 (95% CI 5.9 to 22.0); negative likelihood ratio 0.2 (95% CI 0.0 to 1.2)) and joint effusion on ultrasound (n=127; positive likelihood ratio 8.4 (95% CI 4.1 to 17.1); negative likelihood ratio 0.2 (95% CI 0.1 to 0.3)) appear to be useful. Two promising clinical risk prediction tools were identified, however, their performance was notably lower when tested in external validation studies.

Discussion   Differentiating children with septic arthritis from non-emergent disorders of non-traumatic limp remains a key diagnostic challenge for emergency physicians. There is a need for prospectively derived and validated ED-based clinical risk prediction tools.

Strengths and limitations of this study

► This review has synthesised and analysed in depth 18 studies relating to history, examination, imaging and pathology testing relevant to the diagnosis of septic arthritis in children.

► The review has performed an analysis of the quality of the studies, as well as the performance of various investigations and clinical findings.

► We searched the two major electronic medical databases from 1966, and limited our analysis to English-language articles, so there is a possibility that some studies were missed.

► There was substantial heterogeneity between studies, so we were unable to combine results.

► Changes to epidemiology (increasing prevalence of methicillin-resistant *Staphylococcus aureus*) and vaccine coverage (*Streptococcus pneumoniae* and *Haemophilus influenzae*) over recent years may have influenced how relevant older studies are to current clinical practice.

INTRODUCTION

Non-traumatic limp is a common paediatric presentation to the emergency department (ED).[1] The differential diagnoses are broad, ranging from non-emergent disorders, such as transient synovitis and Legg-Calvé-Perthes disease, to urgent problems such as septic arthritis and slipped upper femoral epiphysis (SUFE).[2 3] After excluding radiographic abnormalities such as Legg-Calvé-Perthes disease and SUFE in high-prevalence age groups, an important clinical question is how to differentiate septic arthritis from non-emergent disorders.

While the majority of children presenting with acute non-traumatic limp have transient synovitis (a benign, self-limiting cause that can be managed conservatively), a small proportion have septic arthritis, which can cause significant morbidity if diagnosis is

delayed.[4] Early diagnosis and treatment (ie, ≤4 days from infection onset) is considered the most important prognostic factor in preventing acute complications such as sepsis and osteomyelitis, as well as long-term problems such as osteonecrosis, joint deformity and early onset osteoarthritis.[5] Thus, the key issue for emergency physicians when evaluating the child presenting with non-traumatic limp is the accurate identification of children with septic arthritis from non-emergent disorders, while also minimising the use of painful and costly overinvestigation.[6 7]

This article aims to describe the published performance characteristics of clinical findings, laboratory testing, imaging and clinical prediction tools for the identification of septic arthritis among children presenting with acute non-traumatic limp in the ED setting.

METHODS

Inclusion and exclusion criteria

Studies that described paediatric patients presenting with monoarticular complaints (ie, limp, altered gait, non-weight bearing, limb pain or swelling) and contained original data on the utility of clinical history, physical examination, biochemical or radiographic findings in the diagnosis of septic arthritis were included. We only included studies which described a reference standard for diagnosis of septic arthritis. This could be achieved using any combination of abnormal synovial fluid macroscopic appearance, elevated white blood cell (WBC) count, synovial fluid and/or blood culture result.

Articles were excluded if they described patients who did not initially present with a monoarticular complaint; examined an adult study population or described a mixed study population (containing both adult and paediatric patients), where the reviewers were unable to extract data to determine sensitivities and specificities for the paediatric population; or if an article did not describe a reference standard for the diagnosis of septic arthritis.

As referral practices and the extent and timing of orthopaedic involvement in the diagnostic process vary between health systems, we did not exclude studies conducted outside the ED setting.

Apart from the information included in this publication and its supplementary material, no additional data are available.

Patient and public involvement

Patients and the public were not involved in the design or conduct of this research.

Search strategy

A structured search of the medical literature (Ovid MEDLINE and EMBASE) from January 1966 to June 2019 was conducted to identify articles that reported the diagnostic value of clinical, biochemical or radiographic findings for differentiating septic arthritis in the limping child from other aetiologies. The following Medical

Subject Headings were used in the search strategy: infectious arthritis combined with medical history taking, physical examination, routine diagnostic tests, diagnostic imaging, differential diagnosis or sensitivity and specificity AND cohort studies, observational study, retrospective studies or prospective studies. The full search strategy is available in online supplemental material 1. Two authors (JT, JC) independently screened the titles and abstracts of the search results. The full manuscript of each article was reviewed if at least one author considered it as potentially relevant.

Quality of the evidence

The overall quality of the included studies was evaluated by two authors (JT, JC) using the Quality Assessment Tool for Diagnostic Accuracy Studies 2 (QUADAS-2).[8] Discrepant quality assessments were adjudicated by discussion with a third author (SC) and resolved by consensus. A kappa analysis using SPSS V.25.0 (SPSS, Chicago, Illinois, USA) was used to assess inter-rater agreement.[9]

The 'ideal' study population was defined as 'children presenting with a lower limb monoarticular complaint (limp, non-weight bearing status, altered gait, limb pain or swelling) to the ED, where 'septic arthritis' and 'non-septic arthritis' were the two evaluated outcomes. Spectrum bias, however, may limit the validity of such studies when applied more broadly to an ED population.[10] As such, if individual trials did not specifically recruit patients from the ED, the 'spectrum' portion of the QUADAS-2 tool was assessed as 'no'. Additionally, if the definition of the reference standard or blinding of index testers to the reference standard was not explicitly stated, these portions of the QUADAS-2 tool were assessed as 'no'. If the follow-up period of patients was not explicitly stated, 'Domain 4: Flow and Timing' was also assessed as 'no'.

Data analysis

A standardised data collection form was used by two authors (JT, JC) to independently extract data from the included studies. Data extracted included study characteristics (such as setting, sample size, inclusion and exclusion criteria) and study definitions for key diagnostic and outcome variables. The diagnostic test properties of key clinical, biochemical and radiographic findings were also extracted or calculated using the available information from the published paper.

Prior to our review, we standardised our definitions for the terms 'false negative', 'false positive', 'true positive' and 'true negative'. 'Disease' was defined as septic arthritis and 'no disease' was defined as an acute arthritis without bacterial aetiology or arthritis secondary to *Borrelia burgodferi* (ie, Lyme arthritis). The latter category included transient synovitis, Lyme arthritis and other non-emergent aetiologies of the limping child.

The diagnostic accuracy of the evaluated clinical and investigation parameters are presented as sensitivities and specificities and positive and negative likelihood ratios (LR), with 95% CIs where appropriate. Discrepancies in

assessment were discussed between three authors (JT, JC, SC) and resolved by consensus.

Accuracy of signs and symptoms

A number of clinical, biochemical and radiographic variables did not share the same definition across studies. For instance, a 'history of fever' had variable definitions across studies (ie, temperature >37.0°C or ≥38.5°C) and whether it was assessed prior, at time of ED presentation or during their inpatient stay. In each case, we defined the absence or presence of each variable on the basis of each individual study's specific definition.

Statistical methods

Sensitivity and specificity (with calculation of 95% CIs), forest plots and summary receiver operating curves were generated using Reviewer Manager (RevMan) (computer program, V.5.3, Copenhagen: The Nordic Cochrane Collaboration, 2014). LR were calculated using MedCalc for Windows (MedCalc Software, Ostend, Belgium) and SPSS V.25.0.

Our literature review commenced as a narrative literature review during a university honours year. We later attempted to register the systematic review on the PROSPERO database, but were unable to do so as the original literature searches and data extraction conducted as part of the honours year project rendered the review ineligible for prospective registration.

RESULTS

A total of 500 articles were identified, including 357 from the MEDLINE and 143 EMBASE medical literature databases, respectively (figure 1). Following exclusion of duplicate articles, 408 unique articles were screened for study eligibility. Of these, 282 articles were excluded on review of their titles and abstracts, and another 108
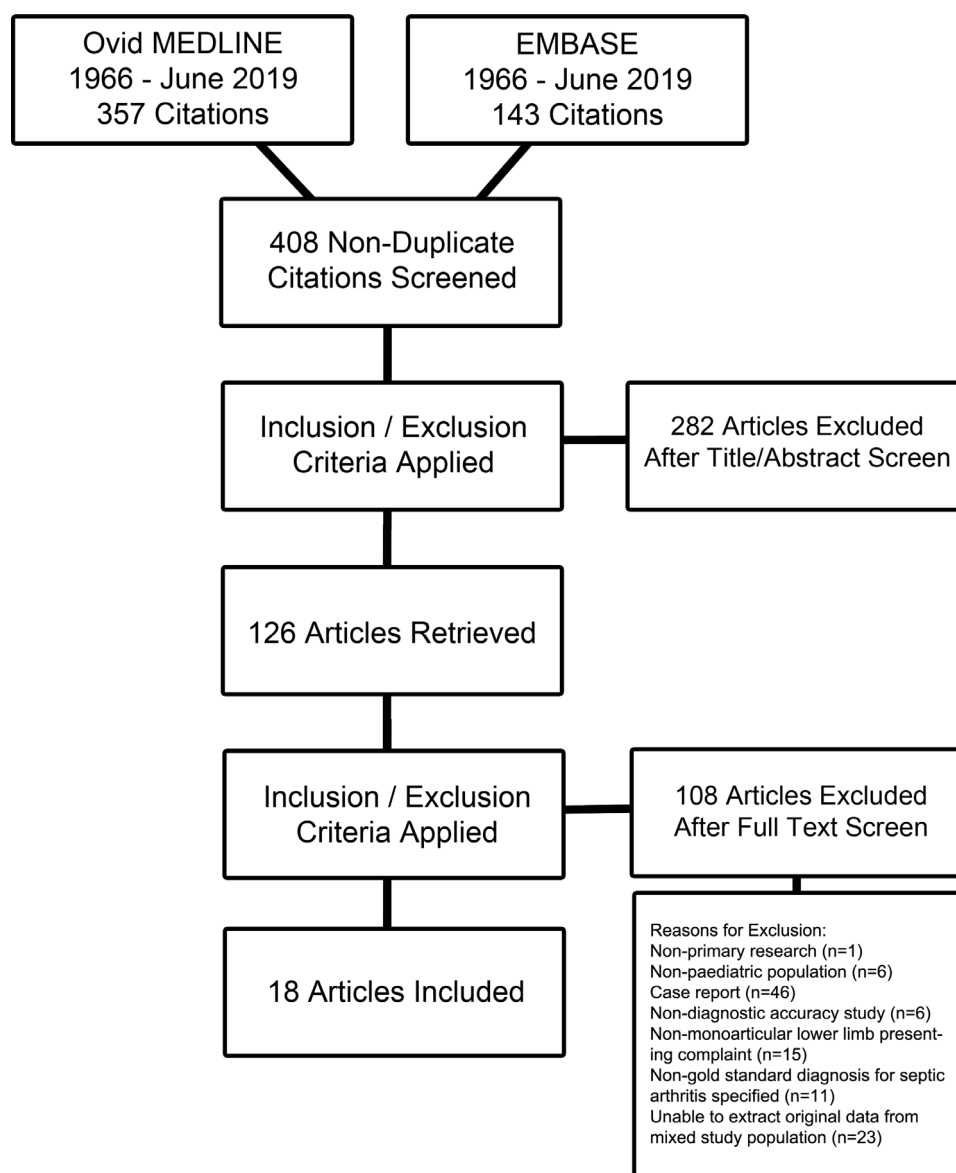
Figure 1 Study flow diagram.

articles were excluded on review of their full manuscripts. Eighteen articles met all inclusion criteria.

Study characteristics and definitions for key diagnostic and outcome variables for the included 18 studies are summarised in table 1. Overall, 2672 patients presenting with lower limb complaints were included in our review, of whom 560 (21.0%) were diagnosed with septic arthritis. These studies took place between 1979 and 2013 and their patient cohorts ranged from 18 to 474 patients. Of the 18 included studies, 12 were retrospective observational cohort studies, 4 were prospective observational cohort studies and 2 were case-control studies.

Only one study included children presenting to the ED with reduced range of motion (ROM) of a skeletal segment.[11] Of the remaining studies of children presenting with lower limb complaints, nine studies examined children who also underwent joint aspiration,[12–20] five examined children admitted as hospital inpatients for exclusion of septic arthritis,[21–25] two examined children who underwent hip MRI for acute hip pain or limp[26 27] and one examined children who had a hip effusion identified on ultrasonography.[28] Only three studies explicitly included ED populations.[11 16 17]

Two authors (JT, JC) independently performed a quality assessment of the 18 included studies using the QUADAS-2 assessment tool for diagnostic accuracy studies (see online supplemental material 2). The authors' QUADAS-2 assessment of quality had a kappa of range 0.82–1.

The quality of the diagnostic accuracy studies for septic arthritis among children presenting with non-traumatic limp is highly variable. Eight studies included patients with a 'recent history of antibiotic use' without an explicit definition for this variable. Four studies also excluded patients on the basis of 'later development of rheumatological disease, Legg-Calvé-Perthes disease or associated proximal femoral osteomyelitis'.

The definition of 'septic arthritis' also varied between studies. Thirteen studies provided broad and detailed definitions for 'septic arthritis' on all key parameters including macroscopic appearance, synovial WBC count and presence/absence of a positive synovial fluid or blood culture result. Five provided limited definitions for septic arthritis: three studies defined it as 'gross pus on joint aspiration or drainage' only, but provided no comment on blood culture, synovial fluid or WBC count[22 24 25]; one defined it on the basis of a positive synovial fluid culture, but provided no comment on blood culture, synovial WBC count or macroscopic appearance[23] and one defined it as a positive synovial fluid or blood culture, with associated 'numerous polymorphs seen on high-screen microscopy', but provided no explicit definition for 'numerous polymorphs'.[21]

No study reported the interval between the index test and the reference standard. In addition, no study explicitly described blinding assessors of the index test from the reference standard or vice versa. Six retrospective and two prospective cohort studies also included children in a control group who did not receive the same reference standard as those subsequently classified as 'septic arthritis' (ie, children classified as 'transient synovitis' was on the basis of 'clinical improvement from bed rest and analgesics alone without undergoing a joint aspiration', while children classified as 'septic arthritis' was on the basis of those who 'underwent a joint aspiration and diagnosed on the basis of synovial fluid or blood culture findings'). Of the six retrospective studies, none of these children was followed up to ensure they had not presented elsewhere and been treated for septic arthritis.

Prevalence of septic arthritis

Deanehan et al[16] found that 19 (3%) of the 673 children presenting to one of two urban paediatric centres in a Lyme disease-endemic area with acute knee monoarthritis between 1992 and 2012 had septic arthritis. Other estimates of septic arthritis in different study populations, including children who were admitted as a hospital inpatient for exclusion of septic arthritis and children who underwent joint aspiration are notably higher, with a prevalence of septic arthritis ranging from 5.2% to 75.6%. No studies described ED prevalence of septic arthritis in a non-Lyme disease-endemic area.

History and examination findings

Eight studies provided sensitivity and specificity data enabling the calculation of positive and negative LR for the risk factors of septic arthritis among children presenting with non-traumatic limp (table 2).[12–16 21–23] None of male gender, history of tick bites, previous antibiotic use, history of chills, joint pain, Lyme season, a previous healthcare visit and recent illness had consistently useful LR to either rule in or rule out a diagnosis of septic arthritis.

Twelve studies examined the role of objective examination findings (table 3).[11–16 18 19 21–24] The definition of fever varied from study to study, ranging from a temperature ≥37.0°C to ≥ 38.5°C. The presence of any documented fever, irrespective of threshold, increased the risk of septic arthritis, although positive LR ranged from 2.0 to 25.2. The absence of fever had negative LR ranging from 0.2 to 0.8.

The ability to weight-bear was assessed by eight studies.[12–16 18 19] This finding had variable performance in different studies, with positive LR ranging from 1.2 to 1.7, and negative LR ranging from 0.1 to 0.9.

Joint tenderness, limited ROM and joint warmth were less commonly assessed clinical variables. A single study suggested that joint tenderness may be a useful finding, with a positive LR of 11.4 (95% CI 5.9 to 22.0), and a negative LR of 0.3 (95% CI 0.2 to 0.5).

Biochemical variables

Six studies provided primary diagnostic data to evaluate the performance of an elevated WBC count in diagnosing septic arthritis among children presenting with lower limb complaints (table 4).[12 18 21–24] The definition of an

Table 1 Studies of the performance of clinical, laboratory and clinical findings in diagnosing septic arthritis

Study	Location	SA cases/total patients	Mean age, years	Study design	Inclusion criteria	Exclusion criteria	Case selection	Control selection	Prevalence of SA, %	Variables examined	Risk of bias?	Applicability concerns to ED setting?
Faesch et al[11]	France	48/339*	4	Prospective observational study	Presented with non-traumatic decreased active motion of a skeletal segment	Neonates, recent antibiotic use	Either confirmed (ie, positive bacteriological culture (blood, synovial fluid aspiration OR bone aspiration)) or presumed infection (purulent aspiration or positive bone scan AND elevated WBC count or CRP, but negative cultures)†	Children presenting with a limp and reduced ROM with 'normal' laboratory and radiographic findings and symptom resolution within 1 week without antibiotic treatment as assessed by orthopaedic surgeons	2.7%–14.2%	Fever Procalcitonin	Yes, due to study flow and timing	Yes, due to patient selection
Kocher et al[13]	USA	82/168	5.6	Retrospective study	Underwent hip joint aspiration	Immunocompromised, renal failure, neonatal sepsis, postoperative infection of the hip, lateral development of rheumatological disease or Perthes disease or associated proximal femoral osteomyelitis	Culture-positive (synovial WBC count >50 000×10^9 cells/L with positive blood culture) or culture-negative septic arthritis (synovial WBC count count >50 000 ×10^9 cells/L with negative blood culture)	Synovial WBC count count <50 000×10^9 cells/L with negative blood culture, symptom resolution without antibiotic therapy and no further development of disease documented	48.8%	Male gender Recent antibiotic use Fever Non-weight bearing History of chills Radiographic effusion	Yes, due to patient selection	Yes, due to patient selection

Continued

Table 1 Continued

Study	Location	SA cases/total patients	Mean age, years	Study design	Inclusion criteria	Exclusion criteria	Case selection	Control selection	Prevalence of SA, %	Variables examined	Risk of bias?	Applicability concerns to ED setting?
Luhmann et al[15]	USA	47/165	5.3	Retrospective study	Underwent hip joint aspiration	Immunocompromised, renal failure, neonatal sepsis, postoperative infection of the hip, lateral development of rheumatological disease or Legg-Calvé-Perthes disease or associated proximal femoral osteomyelitis	Culture-positive (synovial WBC count >50 000×10⁹ cells/L with positive blood culture) or culture-negative septic arthritis (synovial WBC count count >50 000 ×10⁹ cells/L with negative blood culture)	Synovial WBC count count <50 000×10⁹ cells/L with negative blood culture, symptom resolution without antibiotic therapy and no further development of disease documented	28.5%	Male gender Previous healthcare visit Recent antibiotic use Fever Non-weight bearing	Yes, due to patient selection	Yes, due to patient selection
Deanehan et al[16]	USA	13/474*	7.4	Retrospective study	Presented with knee monoarthritis	Recent history of significant knee trauma, knee surgery within the past 30 days, previous arthritis of any joint, history of rheumatological disease, immunocompromised, multiple joint involvement, knee cellulitis or other overlying infection or critical illness (defined as hypotension requiring vasoactive medications or respiratory distress requiring assisted ventilation)	Positive synovial fluid culture or synovial fluid pleocytosis (synovial WBC count >40 000 cells/mL) with positive blood culture result	Documented history of erythema migrans rash, clinical manifestations of Lyme disease with a positive Lyme serology (ie, western blot analysis) and patients with a negative synovial fluid and blood culture, normal synovial WBC count and negative Lyme serology	1.7%–3.1%	Male gender Lyme season Recent illness History of tick bite Recent antibiotic use Fever Joint pain Joint warmth Limited ROM Non-weight bearing	Yes, due to study flow and timing	Yes, due to patient selection

Continued

Table 1 Continued

Study	Location	SA cases/total patients	Mean age, years	Study design	Inclusion criteria	Exclusion criteria	Case selection	Control selection	Prevalence of SA, %	Variables examined	Risk of bias?	Applicability concerns to ED setting?
Caird et al[18]	USA	34/48*	5.5	Retrospective study	Underwent hip aspiration	Immunosuppressed, history of inflammatory arthritis, active Lyme arthritis	Culture-positive (synovial WBC count >50 000×10^9 cells/L and positive synovial fluid, blood culture or synovial gram stain) OR culture-negative (synovial WBC count >50 000×10^9 cells/L ONLY)	Synovial WBC count <50 000×10^9 cells/L and no evidence of bacterial infection (negative synovial fluid, blood culture or synovial gram stain) and did not represent to ED with hip pain or refusal to weight bear	70.8%–75.6%	Recent antibiotic use Fever Non-weight bearing Elevated WBC count Elevated ESR Elevated CRP	Yes, due to study flow and timing	Yes, due to patient selection
Heyworth et al[12]	USA	15/46*	7.6	Prospective observational study	Underwent hip aspiration	Previous or newly established oncological or immunological disorder, history of rheumatological disease, synovial fluid WBC count <25 000 or >75 000×10^9 cells/L	Culture-positive (≥1 positive bacterial cultures not deemed contaminants by treating clinicians) OR culture-negative (based on author consensus)	Lyme arthritis (positive western blot analysis or synovial fluid Lyme PCR result), transient synovitis (negative Lyme arthritis, bacterial cultures or positive cultures deemed contaminants by treating clinicians) or other non-septic arthritis diagnoses	31%–32.6%	Male Fever Non-weight bearing Elevated WBC count Elevated ESR Elevated CRP Synovial WBC count 25 000–50 000×10^9 cells/L Synovial WBC count 50 000–75 000×10^9 cells/L	Unclear, due to study flow and timing	Yes, due to patient selection

Continued

Table 1 Continued

Study	Location	SA cases/ total patients	Mean age, years	Study design	Inclusion criteria	Exclusion criteria	Case selection	Control selection	Prevalence of SA, %	Variables examined	Risk of bias?	Applicability concerns to ED setting?
Strouse et al[28]	USA	11/30*	4.3	Retrospective study	Hip effusion identified on ultrasound	<12 months of age	Culture-positive (positive synovial fluid culture) or culture-negative (macroscopic pus on synovial fluid aspiration, 'high' synovial fluid WBC count and 'abnormal' clinical parameters)	Negative synovial fluid cultures OR symptom resolution without antibiotic medications at time of follow-up	36.4%–40.7%	Joint effusion on X-ray Asymmetrical flow on Power-Doppler ultrasound Debris within effusion on ultrasound	Yes, due to study flow and timing	Yes, due to patient selection and index test
Baldwin et al[19]	USA	49/189	6.9	Prospective observational study	Underwent knee aspiration	Non-knee joint aspirations, diagnoses other than septic or Lyme arthritis, ESR, CRP or Lyme titre not obtained, incomplete medical records, contaminated culture, previous history of rheumatological or Lyme disease, positive synovial fluid culture AND Lyme titre	Culture-positive (positive synovial fluid culture) or culture-negative (synovial WBC count >60 000×10^9 cells/L AND negative Lyme titre)	Positive Western blot analysis	25.9%	Recent antibiotic use Recent illness Fever Joint pain Non-weight bearing Joint warmth Joint tenderness Joint effusion on X-ray	Yes, due to patient selection	Yes, due to patient selection
Sultan et al[21]	UK	5/96	5.7	Retrospective study	Admitted as an inpatient for irritable hip	Incomplete medical records, diagnosis other than septic arthritis or transient synovitis	Positive synovial fluid culture OR positive blood culture with 'numerous WBC' on high-power microscopy of hip aspirate	Negative synovial fluid and blood culture AND total symptom resolution at time of follow-up	5.2%	Male Fever Non-weight bearing Elevated WBC count Elevated ESR Elevated CRP	Unclear, due to patient selection, reference standard and study flow and timing	Yes, due to patient selection

Continued

Table 1 Continued

Study	Location	SA cases/ total patients	Mean age, years	Study design	Inclusion criteria	Exclusion criteria	Case selection	Control selection	Prevalence of SA, %	Variables examined	Risk of bias?	Applicability concerns to ED setting?
Del Beccaro et al[22]	USA	38/132*	4.1	Retrospective study	Admitted as an inpatient OR seen as an outpatient	Non-lower limb monoarthritis presenting complaint, history of trauma, fracture or bony abnormality on X-ray, evidence of invasive bacterial infection at another site, history of rheumatological disease, immunosuppressive disorders (including malignancy, renal failure, IBD or being treated with immunosuppressive or anti-inflammatory medications), diagnoses other than septic arthritis or transient synovitis	Culture-positive (macroscopic pus at time of joint aspiration or hip drainage AND positive synovial fluid, blood culture OR urine latex antigen test for *Haemophilus influenzae* type B) OR Culture-negative (macroscopic pus at time of joint aspiration or hip drainage only)	Negative synovial fluid, blood culture or urine latex antigen test AND symptom resolution without antibiotic therapy at time of follow-up	28.8%–31.2%	Male Fever Elevated WBC count Elevated bands Elevated ESR Joint effusion on X-ray	Yes, due to study flow and timing	Yes, due to patient selection and index test
Jung et al[23]	South Korea	27/124	6.5	Retrospective study	Admitted as an inpatient	Non-lower limb monoarthritis presenting complaint, history of trauma, fracture or bony abnormality on X-ray, evidence of invasive bacterial infection at another site, history of rheumatological disease, immunosuppressive disorders (including malignancy, renal failure, IBD or being treated with immunosuppressive or anti-inflammatory medications), diagnoses other than septic arthritis or transient synovitis	'Based on clinical and laboratory findings' AND confirmed with positive synovial fluid culture	'Based on clinical and laboratory findings' AND 'improved clinical manifestations' with analgesia and bed rest alone	21.8%	Male Fever Elevated WBC count Elevated ESR Elevated CRP 'Abnormal' X-ray findings	Unclear, due to study flow and timing	Yes, due to patient selection and reference standard

Continued

Table 1 Continued

Study	Location	SA cases/ total patients	Mean age, years	Study design	Inclusion criteria	Exclusion criteria	Case selection	Control selection	Prevalence of SA, %	Variables examined	Risk of bias?	Applicability concerns to ED setting?
Kuda et al[24]	Jordan	9/33	NA	Retrospective study	Admitted as an inpatient with irritable hip	Non-lower limb monoarthritis presenting complaint, history of trauma, fracture or bony abnormality on X-ray, evidence of invasive bacterial infection at another site, history of rheumatological disease, immunosuppressive disorders (including malignancy, renal failure, IBD or being treated with immunosuppressive or anti-inflammatory medications)	Culture-positive (macroscopic pus at time of joint aspiration or drainage AND positive synovial fluid or blood culture) OR culture-negative (macroscopic pus at time of joint aspiration or drainage only)	Negative synovial fluid culture and gram stain AND symptom resolution at time of follow-up	27.3%	Fever Elevated WBC count Elevated ESR	Unclear due to study flow and timing	Yes, due to patient selection
Zamzam et al[25]	USA	59/127	4.3	Retrospective study	Admitted as an inpatient with suspected hip septic arthritis and underwent ultrasound	'Abnormal' X-ray findings identified prior to ultrasound, had 'hip or extra-hip problems that mimic hip septic arthritis or transient synovitis clinically', patients with 'general diseases that predispose to or affect the course of septic arthritis'	Macroscopic pus or turbid fluid identified at time of joint aspiration	No macroscopic pus or turbid fluid identified at time of joint aspiration	46.4%	Hip effusion on ultrasound	Yes, due to study flow and timing	Yes, due to patient selection, index test and reference standard

Continued

Table 1 Continued

Study	Location	SA cases/ total patients	Mean age, years	Study design	Inclusion criteria	Exclusion criteria	Case selection	Control selection	Prevalence of SA, %	Variables examined	Risk of bias?	Applicability concerns to ED setting?
Kocher et al[14]	USA	51/154	5.8	Prospective observational study	Underwent joint aspiration	Immunocompromised, renal failure, neonatal sepsis, postoperative infection of the hip, lateral development of rheumatological disease or Legg-Calvé-Perthes disease or associated proximal femoral osteomyelitis	Culture-positive (synovial WBC count >50 000×10^9 cells/L with positive blood culture) or culture-negative septic arthritis (synovial WBC count count >50 000 ×10^9 cells/L with negative blood culture)	Synovial WBC count count <50 000 ×10^9 cells/L with negative blood culture, symptom resolution without antibiotic therapy and no further development of disease documented	33.1%	Male Recent antibiotic use Fever History of chills Non-weight bearing Joint effusion on X-ray	Yes, due to patient selection	Yes, due to patient selection
Levine et al[20]	USA	39/133*	6.6	Retrospective study	Underwent joint aspiration	Did not have a CRP reported within 24 hours of presentation	Culture-positive (positive synovial fluid culture) OR culture-negative (synovial fluid WBC count >50 000 ×10^9 cells/L with polymorphic neutrophils and negative western blot analysis or Lyme ELISA test)	Negative synovial fluid culture OR synovial fluid WBC count <50 000×10^9 cells/L OR positive western blot analysis or Lyme ELISA test	29.3%–31.5%	Elevated ESR Elevated CRP	Unclear, due to study flow and timing	Yes, due to patient selection

Continued

Table 1 Continued

Study	Location	SA cases/ total patients	Mean age, years	Study design	Inclusion criteria	Exclusion criteria	Case selection	Control selection	Prevalence of SA, %	Variables examined	Risk of bias?	Applicability concerns to ED setting?
Lee et al[26]	South Korea	9/23	11.8	Case-control study	Underwent hip MRI for acute hip pain or limp	Diagnoses other than septic arthritis or transient synovitis	Culture-positive (positive synovial fluid, blood culture or synovial histological findings) or culture-negative (macroscopic pus on joint aspiration and >20 000×10⁹ cells/L with polymorphic neutrophils)	Negative synovial fluid, blood culture or histological findings, clinical improvement with analgesia and bed rest alone	39.1%	Joint effusion on MRI Soft tissue signal intensity alterations on MRI	Yes, due to patient selection and study flow and timing	Yes, due to patient selection and index test
Kwack et al[27]	South Korea	7/18	5.5	Case-control study	Underwent hip MRI for acute hip pain	Diagnoses other than septic arthritis or transient synovitis, history of trauma, later development of rheumatological disease, avascular necrosis of femoral head or Legg-Calvé-Perthes disease	Culture-positive (positive synovial fluid culture) or culture-negative (macroscopic pus on joint aspiration and synovial fluid WBC count >50 000×10⁹ cells/L)	Synovial fluid <50 000×10⁹ cells/L with negative culture, symptom resolution with conservative treatment, no recurrence documented in medical records	38.9%	Joint effusion on MRI Soft tissue signal intensity alterations on MRI Bone marrow signal intensity alterations on MRI	Yes, due to patient selection and study flow and timing	Yes, due to patient selection and index test

Continued

Table 1 Continued

Study	Location	SA cases/ total patients	Mean age, years	Study design	Inclusion criteria	Exclusion criteria	Case selection	Control selection	Prevalence of SA, %	Variables examined	Risk of bias?	Applicability concerns to ED setting?
Deanehan et al[17]	USA	17/373	8	Retrospective study	Presented with knee monoarthritis	Recent history of significant knee trauma, knee surgery within the past 30 days, previous arthritis of any joint, history of rheumatological disease, immunocompromised, multiple joint involvement, knee cellulitis or other overlying infection or critical illness (defined as hypotension requiring vasoactive medications or respiratory distress requiring assisted ventilation)	Positive synovial fluid culture or synovial fluid pleocytosis (synovial WBC count >40 000 cells/mL) with positive blood culture result	Documented history of erythema migrans rash, clinical manifestations of Lyme disease with a positive Lyme serology (ie, western blot analysis) and patients with a negative synovial fluid and blood culture, normal synovial WBC count and negative Lyme serology	4.6%	Synovial fluid WBC count >40 000 cells/mL, >50 000 cells/ mL, >75 000 cells/mL, >1 00 000 cells/ mL	Yes, due to study flow and timing	Yes, due to patient selection
Total		560/2672										

*Maximum possible number of eligible patients; some clinical parameters had missing data for specific parameters.
†Included osteomyelitis patients.
CRP, C reactive protein; ED, emergency department; ESR, erythrocyte sedimentation rate; IBD, inflammatory bowel disease; ROM, range of motion; SA, septic arthritis; WBC, white blood cell.

Table 2 Sensitivities, specificities and likelihood ratios for findings on history and examination

Risk factors	Study	Prevalence of SA, %	Sensitivity (95% CI)	Specificity (95% CI)	Likelihood ratio (95% CI) Positive (LR+)	Negative (LR−)
History						
Male	Deanehan et al[16]	2.7	0.54 (0.25 to 0.81)	0.38 (0.33 to 0.42)	0.9 (0.5 to 1.4)	1.2 (0.7 to 2.2)
	Del Becarro et al[22]	28.8	0.68 (0.51 to 0.82)	0.41 (0.31 to 0.52)	1.2 (0.9 to 1.5)	0.8 (0.5 to 1.3)
	Heyworth et al[12]	32.6	0.60 (0.32 to 0.84)	0.42 (0.25 to 0.61)	1.0 (0.6 to 1.7)	1.0 (0.5 to 2.0)
	Jung et al[23]	21.8	0.67 (0.46 to 0.83)	0.27 (0.18 to 0.37)	0.9 (0.7 to 1.2)	1.2 (0.7 to 2.3)
	Kocher et al[13]	48.8	0.50 (0.39 to 0.61)	0.34 (0.24 to 0.45)	0.8 (0.6 to 1.0)	1.5 (1.0 to 2.1)
	Kocher et al[14]	33.1	0.49 (0.35 to 0.63)	0.29 (0.21 to 0.39)	0.7 (0.5 to 0.9)	1.8 (1.2 to 2.6)
	Luhmann et al[15]	28.5	0.53 (0.38 to 0.68)	0.27 (0.19 to 0.36)	0.7 (0.6 to 1.0)	1.7 (1.1 to 2.6)
	Sultan et al [21]	5.2	0.60 (0.15 to 0.95)	0.31 (0.22 to 0.42)	0.9 (0.4 to 1.8)	1.3 (0.4 to 3.8)
History of previous antibiotic use	Baldwin et al[19]	25.9	0.22 (0.12 to 0.37)	0.94 (0.88 to 0.97)	3.5 (1.5 to 7.9)	0.8 (0.7 to 1.0)
	Caird et al[18]	75.6	0.32 (0.17 to 0.51)	0.80 (0.44 to 0.97)	1.6 (0.4 to 6.2)	0.9 (0.6 to 1.3)
	Deanehan et al[16]	2.6	0.17 (0.02 to 0.48)	0.89 (0.86 to 0.92)	1.5 (0.4 to 5.5)	0.9 (0.7 to 1.2)
	Kocher et al[13]	48.8	0.24 (0.16 to 0.35)	0.90 (0.81 to 0.95)	2.3 (1.1 to 4.8)	0.8 (0.7 to 1.0)
	Kocher et al [14]	33.1	0.12 (0.04 to 0.24)	0.83 (0.74 to 0.89)	0.7 (0.3 to 1.6)	1.1 (0.9 to 1.2)
	Luhmann et al[15]	28.5	0.32 (0.19 to 0.47)	0.82 (0.74 to 0.89)	1.8 (1.0 to 3.2)	0.8 (0.7 to 1.0)
Lyme season	Deanehan et al[16]	2.7	0.38 (0.14 to 0.68)	0.56 (0.51 to 0.61)	0.9 (0.4 to 1.8)	1.1 (0.7 to 1.7)
Previous healthcare visit	Luhmann et al[15]	28.5	0.64 (0.49 to 0.77)	0.55 (0.46 to 0.64)	1.4 (1.1 to 1.9)	0.7 (0.4 to 1.0)
History of recent illness	Baldwin et al[19]	28.5	0.24 (0.13 to 0.39)	0.87 (0.80 to 0.92)	1.9 (1.0 to 3.7)	0.9 (0.7 to 1.0)
	Deanehan et al [16]	2.8	0.46 (0.19 to 0.75)	0.83 (0.79 to 0.86)	2.7 (1.5 to 5.1)	0.7 (0.4 to 1.1)
History of tick bite	Deanehan et al[16]	1.7	0.00 (0.00 to 0.52)	0.86 (0.82 to 0.90)	NA	1.2 (1.1 to 1.2)
History of chills	Kocher et al[13]	48.8	0.11 (0.05 to 0.20)	1.00 (0.96 to 1.00)	NA	0.9 (0.8 to 1.0)
	Kocher et al[14]	33.1	0.04 (0.00 to 0.13)	1.00 (0.96 to 1.00)	NA	1.0 (0.9 to 1.0)
History of joint pain	Baldwin et al[19]	25.9	0.98 (0.89 to 1.00)	0.00 (0.00 to 0.03)	1.0 (0.9 to 1.0)	NA
	Deanehan et al[16]	2.8	1.00 (0.75 to 1.00)	0.05 (0.03 to 0.07)	1.1 (1.0 to 1.1)	NA
Examination						
Fever documented, but not defined	Baldwin et al[19]	25.9	0.80 (0.66 to 0.90)	0.59 (0.51 to 0.68)	2.0 (1.5 to 2.5)	0.3 (0.2 to 0.6)
	Deanehan et al[16]	2.8	0.77 (0.46 to 0.95)	0.73 (0.68 to 0.77)	2.8 (2.0 to 3.9)	0.3 (0.1 to 0.9)
	Heyworth et al[12]	32.6	0.27 (0.08 to 0.55)	0.94 (0.79 to 0.99)	4.1 (0.9 to 20.1)	0.8 (0.6 to 1.1)
Temperature ≥37.0°C	Jung et al[23]	21.8	0.78 (0.58 to 0.91)	0.97 (0.91 to 0.99)	25.2 (8.1 to 78.0)	0.2 (0.1 to 0.5)
Temperature ≥37.5°C	Del Beccaro et al[22]	29.9	0.66 (0.49 to 0.80)	0.70 (0.59 to 0.79)	2.2 (1.5 to 3.2)	0.5 (0.3 to 0.8)
	Kuda et al[24]	27.3	1.00 (0.66 to 1.00)	0.67 (0.45 to 0.84)	3.0 (1.7 to 5.3)	NA
Temperature ≥38.0°C	Del Beccaro et al[22]	29.9	0.45 (0.29 to 0.62)	0.85 (0.76 to 0.92)	3.1 (1.7 to 5.7)	0.7 (0.5 to 0.9)
	Faesch et al[11]	14.2	0.60 (0.45 to 0.74)	0.85 (0.80 to 0.89)	4.0 (2.8 to 5.7)	0.5 (0.3 to 0.7)
	Sultan et al[21]	5.2	1.00 (0.48 to 1.00)	0.90 (0.82 to 0.95)	10.1 (5.4 to 18.8)	NA
Temperature ≥38.5°C	Caird et al[18]	70.8	0.44 (0.27 to 0.62)	1.00 (0.77 to 1.00)	NA	0.6 (0.4 to 0.8)
	Kocher et al[13]	48.8	0.82 (0.72 to 0.89)	0.92 (0.84 to 0.97)	10.0 (4.9 to 20.6)	0.2 (0.1 to 0.3)
	Kocher et al[14]	33.1	0.61 (0.46 to 0.74)	0.76 (0.66 to 0.84)	2.5 (1.7 to 3.8)	0.5 (0.4 to 0.7)
	Kuda et al[24]	27.3	0.78 (0.40 to 0.97)	0.92 (0.73 to 0.99)	9.3 (2.4 to 36.8)	0.2 (0.1 to 0.8)
	Luhmann et al[15]	28.5	0.72 (0.57 to 0.84)	0.65 (0.56 to 0.74)	2.1 (1.5 to 2.8)	0.4 (0.3 to 0.7)
	Sultan et al[21]	5.2	0.80 (0.28 to 0.99)	0.96 (0.86 to 0.99)	18.2 (6.4 to 52.2)	0.2 (0.0 to 1.2)
Joint tenderness	Baldwin et al[19]	25.9	0.73 (0.59 to 0.85)	0.94 (0.88 to 0.97)	11.4 (5.9 to 22.0)	0.3 (0.2 to 0.5)

Continued

Table 2 Continued

Risk factors	Study	Prevalence of SA, %	Sensitivity (95% CI)	Specificity (95% CI)	Likelihood ratio (95% CI) Positive (LR+)	Negative (LR–)
Limited range of motion	Deanehan et al[16]	2.8	0.92 (0.64 to 1.00)	0.30 (0.26 to 0.34)	1.3 (1.1 to 1.6)	0.3 (0.0 to 1.7)
Joint warmth	Baldwin et al[19]	25.9	0.82 (0.68 to 0.91)	0.56 (0.48 to 0.65)	1.9 (1.5 to 2.4)	0.3 (0.2 to 0.6)
	Deanehan et al[16]	3.1	0.83 (0.52 to 0.98)	0.40 (0.35 to 0.45)	1.4 (1.1 to 1.8)	0.4 (0.1 to 1.5)
Non-weight bearing	Baldwin et al[19]	25.9	0.53 (0.38 to 0.67)	0.75 (0.67 to 0.82)	2.1 (1.4 to 3.1)	0.6 (0.5 to 0.9)
	Caird et al[18]	70.8	0.91 (0.76 to 0.98)	0.29 (0.08 to 0.58)	1.3 (0.9 to 1.8)	0.3 (0.1 to 1.2)
	Deanehan et al[16]	2.7	0.33 (0.10 to 0.65)	0.76 (0.72 to 0.80)	1.4 (0.6 to 3.2)	0.9 (0.6 to 1.3)
	Heyworth et al[12]	32.6	0.93 (0.68 to 1.00)	0.19 (0.07 to 0.37)	1.2 (0.9 to 1.4)	0.3 (0.1 to 2.6)
	Kocher et al[13]	48.8	0.95 (0.88 to 0.99)	0.65 (0.54 to 0.75)	2.7 (2.0 to 3.7)	0.1 (0.0 to 0.2)
	Kocher et al[14]	33.1	0.84 (0.71 to 0.93)	0.51 (0.41 to 0.61)	1.7 (1.4 to 2.2)	0.3 (0.2 to 0.6)
	Luhmann et al[15]	28.5	0.81 (0.67 to 0.91)	0.31 (0.23 to 0.41)	1.2 (1.0 to 1.4)	0.6 (0.3 to 1.2)
	Sultan et al[21]	5.2	0.60 (0.15 to 0.95)	0.71 (0.61 to 0.80)	2.1 (1.0 to 4.6)	0.6 (0.2 to 1.7)

elevated WBC count varied between studies, ranging from $\geq 11.0 \times 10^9$/L to $\geq 15.0 \times 10^9$/L. Irrespective of the definition used, the presence or absence of an elevated WBC count did not significantly change the odds of septic arthritis.

The performance of an elevated ESR count was evaluated in seven studies.[12 18 20–24] The definition of an elevated ESR varied from ESR ≥ 20 mm/hour to ESR ≥ 75 mm/hour. Positive LR for an elevated ESR ranged from 1.2 to 12, with negative LR ranging from 0.1 to 0.9.

The definition of an elevated CRP level also varied, ranging from CRP ≥ 7 mg/L to ≥ 105 mg/L.[12 18 20 23] Positive LR for an elevated CRP ranged from 1.2 to 12.3, with negative LR ranging from 0.1 to 0.7.

A single prospective study of 339 children presenting with non-traumatic decreased range of motion of a skeletal segment found that a procalcitonin ≥ 0.5 ng/mL was highly specific (0.97; 95% CI 0.94 to 0.99) for the diagnosis of septic arthritis, but had a very low sensitivity (0.13; 95% CI 0.05 to 0.25).[11]

A synovial WBC count $\geq 100\,000$ cells/µL has traditionally been used as one of the diagnostic criteria for septic arthritis. Conversely, a synovial WBC count $\leq 25\,000$ cells/µL is generally seen to exclude septic arthritis. However, two studies examining a range of synovial WBC counts between 25 000 and 100 000 cells/µL did not identify a cut-off with a clinically useful positive or negative LR.[12 17]

Radiographic variables
Six studies evaluated the role of plain radiographs for the diagnosis of septic arthritis in the child presenting with non-traumatic limp, but had inconsistent findings (table 5).[13 14 19 22 23 28]

A single study of 30 children found that the presence of an ultrasonographic effusion moderately increased the risk of septic arthritis (LR+ 8.4; 95% CI 4.1 to 17.1), while the absence of an ultrasonographic effusion moderately reduced the risk of septic arthritis (LR– 0.2; 95% CI 0.1 to 0.3).[25]

MRI was evaluated in two studies, which together examined 41 patients, 16 of whom had with septic arthritis.[26 27] Various findings were assessed, with a range of diagnostic utility.

Separate summary receiver operator curves are presented for fever, features on history, features on examination, WBC count, CRP level, ESR level, synovial fluid findings, X-ray, ultrasound and MRI in online supplemental materials 3–11.

Clinical risk prediction tools
Our review identified two multivariate clinical risk prediction tools for septic arthritis among children presenting with non-traumatic limp that have had their validity assessed in populations separate to their derivation sample.

The four components of prediction tool by Kocher et al[13] are non-weight-bearing status, fever, raised WBC count ($\geq 12 \times 10^9$/L) and raised ESR (≥ 40 mm/hour) (table 3).[14 15 18 21] Caird et al[18] derived an alternative prediction tool that included a fifth parameter, CRP ≥ 20 mg/L, in addition to original four-predictor model by Kocher et al (table 3).

Notably, the performances of both clinical risk prediction tools are significantly worse in external validation studies (table 3). In original derivation study by Kocher et al, the predicted probability of a child presenting with non-traumatic limp having septic arthritis was 99.6% when all four components were present. However, when this clinical risk prediction tool was applied in two external validation studies, the predicted probability of septic arthritis ranged from 58.1% to 93%. The area under the receiver-operator curve was also notably lower in both studies compared with original derivation study by Kocher et al (0.80[15] and 0.86,[14] respectively compared with 0.96).

Table 3 Sensitivities, specificities and likelihood ratios for laboratory findings

Serum laboratory values	Study	Prevalence of SA, %	Sensitivity (95% CI)	Specificity (95% CI)	Likelihood ratio (95% CI) Positive (LR+)	Negative (LR−)
White blood cell count (WBC)						
WBC bands ≥350/mm³	Del Beccaro et al[22]	31.2	0.53 (0.35 to 0.70)	0.75 (0.64 to 0.84)	2.1 (1.3 to 3.4)	0.6 (0.4 to 0.9)
≥11.0×10⁹/L	Jung et al[23]	21.8	0.74 (0.54 to 0.89)	0.94 (0.87 to 0.98)	12.0 (5.4 to 26.8)	0.3 (0.2 to 0.5)
≥12.0×10⁹/L	Caird et al[18]	70.8	0.50 (0.32 to 0.68)	0.71 (0.42 to 0.92)	1.8 (0.7 to 4.3)	0.7 (0.4 to 1.1)
	Heyworth et al[12]	32.6	0.47 (0.21 to 0.73)	0.55 (0.36 to 0.73)	1.0 (0.5 to 2.0)	1.0 (0.6 to 1.7)
	Sultan et al[21]	5.2	0.40 (0.05 to 0.85)	0.81 (0.72 to 0.89)	2.1 (0.7 to 6.8)	0.7 (0.4 to 1.5)
≥15.0×10⁹/L	Del Beccaro et al[22]	29.2	0.26 (0.13 to 0.43)	0.84 (0.75 to 0.91)	1.6 (0.8 to 3.3)	0.9 (0.7 to 1.1)
	Kuda et al[24]	27.3	0.78 (0.40 to 0.97)	0.75 (0.53 to 0.90)	3.1 (1.4 to 6.8)	0.3 (0.1 to 1.0)
Erythrocyte sedimentation rate (ESR)						
≥20 mm/hour	Del Beccaro et al[22]	29.2	0.79 (0.63 to 0.90)	0.72 (0.61 to 0.81)	2.8 (1.9 to 4.0)	0.3 (0.2 to 0.6)
	Jung et al[23]	21.8	0.93 (0.76 to 0.99)	0.59 (0.48 to 0.69)	2.3 (1.7 to 2.9)	0.1 (0.0 to 0.5)
	Kuda et al[24]	27.3	1.00 (0.66 to 1.00)	0.75 (0.53 to 0.90)	4.0 (2.0 to 8.0)	NA
≥25 mm/hour	Levine et al[20]	31.5	0.92 (0.79 to 0.98)	0.22 (0.14 to 0.33)	1.2 (1.0 to 1.4)	0.3 (0.1 to 1.1)
≥30 mm/hour	Del Beccaro et al[22]	29.2	0.71 (0.54 to 0.85)	0.86 (0.77 to 0.92)	5.0 (2.9 to 8.7)	0.3 (0.2 to 0.6)
≥40 mm/hour	Caird et al[18]	70.8	0.56 (0.38 to 0.73)	0.86 (0.57 to 0.98)	3.9 (1.1 to 14.6)	0.5 (0.3 to 0.8)
	Heyworth et al[12]	32.6	0.53 (0.27 to 0.79)	0.58 (0.39 to 0.75)	1.3 (0.7 to 2.4)	0.8 (0.4 to 1.5)
	Jung et al[23]	21.8	0.74 (0.54 to 0.89)	0.94 (0.87 to 0.98)	12.0 (5.4 to 26.8)	0.3 (0.2 to 0.5)
	Sultan et al[21]	5.2	0.40 (0.05 to 0.85)	0.88 (0.79 to 0.84)	3.3 (1.0 to 11.1)	0.7 (0.3 to 1.4)
≥50 mm/hour	Kuda et al[24]	27.3	0.67 (0.30 to 0.93)	0.92 (0.73 to 0.99)	8.0 (2.0 to 32.6)	0.4 (0.1 to 0.9)
≥75 mm/hour	Levine et al[20]	31.5	0.26 (0.13 to 0.42)	0.87 (0.78 to 0.93)	2.0 (0.9 to 4.3)	0.9 (0.7 to 1.1)
C reactive protein (CRP)						
≥7 mg/L	Heyworth et al[12]	31	1.00 (0.75 to 1.00)	0.14 (0.04 to 0.32)	1.2 (1.0 to 1.3)	NA
≥10 mg/L	Jung et al[23]	21.8	0.89 (0.71 to 0.98)	0.93 (0.86 to 0.97)	12.3 (6.0 to 25.5)	0.1 (0.0 to 0.4)
	Levine et al[20]	29.3	0.90 (0.76 to 0.97)	0.29 (0.20 to 0.39)	1.3 (1.1 to 1.5)	0.4 (0.1 to 1.0)
≥20 mg/L	Caird et al[18]	70.8	0.85 (0.69 to 0.95)	0.71 (0.42 to 0.92)	3.0 (1.3 to 6.9)	0.2 (0.1 to 0.5)
	Sultan et al[21]	5.2	0.60 (0.15 to 0.95)	0.90 (0.82 to 0.95)	6.1 (2.4 to 15.6)	0.4 (0.2 to 1.3)
≥50 mg/L	Levine et al[20]	29.3	0.67 (0.50 to 0.81)	0.67 (0.57 to 0.76)	2.0 (1.4 to 2.9)	0.5 (0.3 to 0.8)
≥105 mg/L	Levine et al[20]	29.3	0.41 (0.26 to 0.58)	0.85 (0.76 to 0.92)	2.8 (1.5 to 5.1)	0.7 (0.5 to 0.9)
Serum procalcitonin						
≥0.5 ng/mL	Faesch et al[11]	14.2	0.13 (0.05 to 0.25)	0.97 (0.94 to 0.99)	4.0 (1.5 to 10.8)	0.9 (0.8 to 1.0)
Synovial fluid WBC count						
≥40 000 cells/µL	Deanehan et al[17]	4.6	0.65 (0.38 to 0.86)	0.46 (0.40 to 0.51)	1.2 (0.8 to 1.7)	0.8 (0.4 to 1.5)
50 000 cells/µL	Deanehan et al[17]	4.6	0.41 (0.18 to 0.67)	0.59 (0.54 to 0.64)	1.0 (0.6 to 1.8)	1.0 (0.7 to 1.5)
≥75 000 cells/µL	Deanehan et al[17]	4.6	0.29 (0.10 to 0.56)	0.79 (0.74 to 0.83)	1.4 (0.6 to 3.0)	0.9 (0.7 to 1.2)
≥100 000 cells/µL	Deanehan et al[17]	4.6	0.24 (0.07 to 0.50)	0.89 (0.85 to 0.92)	2.2 (0.9 to 5.3)	0.9 (0.7 to 1.1)
25 000–50 000 cells/mm³	Heyworth et al[12]	32.6	0.27 (0.08 to 0.55)	0.39 (0.22 to 0.58)	0.4 (0.2 to 1.1)	1.9 (1.1 to 3.2)
50 000–75 000 cells/mm³	Heyworth et al[12]	32.6	0.73 (0.45 to 0.92)	0.61 (0.42 to 0.78)	1.9 (1.1 to 3.2)	0.4 (0.2 to 1.1)

NA, not available.

A similar observation was seen when the external validity of prediction tool by Caird et al was assessed in follow-up study by Sultan et al,[21] with a comparable reduction noted in the predicted probability of a child having septic arthritis when all five predictors were present in a new population group (60%[21] compared with 98%[18]).

Table 4 Sensitivities, specificities and likelihood ratios for imaging findings

Radiographic findings	Study	Prevalence of SA, %	Sensitivity (95% CI)	Specificity (95% CI)	Likelihood ratio (95% CI)	
					Positive (LR+)	Negative (LR–)
X-ray						
Radiographic effusion	Baldwin et al[19]	25.9	0.90 (0.78 to 0.97)	0.05 (0.02 to 0.10)	1.0 (0.9 to 1.1)	2.0 (0.7 to 6.1)
	Del Beccaro et al[22]	30.2	0.84 (0.69 to 0.94)	0.58 (0.47 to 0.68)	2.0 (1.5 to 2.7)	0.3 (0.1 to 0.6)
	Jung et al[23]	21.8	1.00 (0.87 to 1.00)	0.23 (0.15 to 0.32)	1.3 (1.2 to 1.4)	NA
	Kocher et al[13]	48.8	0.77 (0.66 to 0.85)	0.62 (0.51 to 0.72)	2.0 (1.5 to 2.7)	0.4 (0.3 to 0.6)
	[14]Kocher et al	33.1	0.14 (0.06 to 0.26)	0.89 (0.82 to 0.95)	1.3 (0.5 to 3.1)	1.0 (0.9 to 1.1)
	Strouse et al[28]	36.4	0.13 (0.00 to 0.53)	0.79 (0.49 to 0.95)	0.6 (0.1 to 4.7)	1.1 (0.8 to 1.6)
Ultrasound						
Ultrasonographic effusion	Zamzam et al[25]	46.4	0.86 (0.75 to 0.94)	0.90 (0.80 to 0.96)	8.4 (4.1 to 17.1)	0.2 (0.1 to 0.3)
Debris within effusion	Strouse et al[28]	40.7	0.73 (0.39 to 0.94)	0.63 (0.35 to 0.85)	1.9 (0.9 to 4.0)	0.4 (0.2 to 1.2)
Difference in power Doppler signal between two hips	Strouse et al[28]	36.7	0.27 (0.06 to 0.61)	1.00 (0.82 to 1.00)	NA	0.7 (0.5 to 1.0)
MRI						
Grade 3 joint effusion	Kwack et al[27]	38.9	0.57 (0.18 to 0.90)	0.36 (0.11 to 0.69)	0.9 (0.4 to 2.0)	1.2 (0.4 to 3.8)
	Lee et al[26]	39.1	0.89 (0.52 to 1.00)	0.29 (0.08 to 0.58)	1.2 (0.8 to 1.9)	0.4 (0.1 to 3.0)
Low signal intensity in fat-suppressed gadolinium-enhanced T1-weighted coronal MRI, Decreased perfusion to femoral head	Kwack et al[27]	38.9	0.86 (0.42 to 1.00)	0.82 (0.48 to 0.98)	4.7 (1.3 to 17.1)	0.2 (0.0 to 1.1)
Low signal intensity in bone marrow on T1-weighted and high signal intensity on fat-suppressed T2-weighted images	Kwack et al[27]	38.9	0.43 (0.10 to 0.82)	1.00 (0.72 to 1.00)	NA	0.6 (0.3 to 1.1)
	Lee et al[26]	39.1	0.89 (0.52 to 1.00)	1.00 (0.77 to 1.00)	NA	0.1 (0.0 to 0.7)
Signal intensity alterations in soft tissue; 'poorly defined areas on high signal intensity on fat-suppressed T2-weighted images'	Kwack et al[27]	38.9	0.71 (0.29 to 0.96)	0.64 (0.31 to 0.89)	2.0 (0.8 to 4.9)	0.5 (0.1. 1.6)
	Lee et al[26]	39.1	0.89 (0.52 to 1.00)	0.29 (0.08 to 0.58)	1.2 (0.8 to 1.9)	0.4 (0.1 to 3.0)
Enhancing thick rim of inflamed synovial membrane on fat-suppressed contrast-enhanced images	Kwack et al[27]	38.9	0.57 (0.18 to 0.90)	0.55 (0.23 to 0.83)	1.3 (0.5 to 3.1)	0.8 (0.3 to 2.2)

NA, not available.

DISCUSSION

The overall quality of the current literature used to inform the evaluation of a child presenting with acute non-traumatic limp to the ED is relatively low. It is difficult to apply the literature to children presenting to the ED with limp, due to spectrum bias, inconsistent definitions between studies, unclear temporal relationship between diagnostic tests and gold-standard diagnosis and poor performance of clinical risk prediction tools in external validation studies.

Table 5 Sensitivities, specificities and likelihood ratios for clinical risk prediction tools

	Kocher*et al*[13]	Kocher*et al*[14]	Luhmann*et al*[15]	Caird*et al*[18]	Sultan aand Hughes[21]
Study characteristics					
Study design	Retrospective observational	Prospective observational	Retrospective observational	Prospective observational	Retrospective observational
Study type	Derivation	Validation	Validation	Derivation	Validation
Study population size	168	154	165	48	96
% of patients with septic arthritis	48.8	33.1	28.5	70.8	5.2
Non-weight bearing	✓	✓	✓	✓	✓
Fever (≥38.5°C)	✓	✓	✓	✓	✓
White cell count (≥12.0×10^9/L)	✓	✓	✓	✓	✓
Erythrocyte sedimentation rate >40 mm/hour	✓	✓	✓	✓	✓
C reactive protein >20 mg/L				✓	✓
Area under receiver-operator curve	0.96	0.86	0.80		
If 0 predictors present	<0.2%	2%		17%	2.3%
If 1 predictors present	3%	9.5%		37%	5%
If 2 predictors present	40%	35%		62%	11%
If 3 predictors present	93%	73%		83%	22%
If 4 predictors present	99.6%	93.0%	59.1%	93%	39%
If 5 predictors present				98%	60%

Spectrum bias occurs when differences exist in the prevalence of septic arthritis between the study and clinical setting. The prevalence of septic arthritis among children with non-traumatic limp varied from 5.2%[21] to 75.6%[18] in hospital inpatient or joint aspiration studies, while it was only 3% in an ED-based study population in a Lyme disease-endemic area.[16] In high prevalence settings, the diagnostic performance of the evaluated clinical, biochemical and radiographic variables are likely to be an overestimate of their actual performance when applied to a more general ED population group.

It was difficult to determine the timing between assessment for specific variables and the time that diagnostic joint aspiration was performed to exclude or confirm a diagnosis of septic arthritis among children presenting with non-traumatic limp. For example, when considering the parameter 'fever', the timing of the fever was unclear (ie, pre-arrival, initial ED assessment or during their hospital inpatient stay). Furthermore, the timing between onset of symptoms and diagnostic testing was also indeterminable. This factor can have important ramifications for the diagnostic values of tests, such as CRP; for example, a low CRP after 3 days of symptoms may be expected to carry a higher sensitivity for septic arthritis than a low CRP 3 hours after symptom onset.[29]

A number of important variables did not have consistent definitions between studies. Different thresholds were used to define an 'abnormal' finding for fever, WBC count, CRP and ESR. This consequently leads to fragmentation of the aggregated data and smaller number of studies examining each definition.

Although appealing, the use of clinical risk prediction tools remains contentious. Key limitations of clinical risk prediction tools include poor performance in external validation studies, and the application of the tool originally to highly selected populations. This may be due to small sample sizes, a limited number of septic arthritis cases and differences in the prevalence of septic arthritis in derivation and validation samples.[30] With relatively poor performance in external validation studies, it is difficult to justify the application of these tools to the ED setting.

No studies have assessed the application of clinical risk prediction tools to children presenting with acute non-traumatic limp to the ED. Kocher *et al,*[13] Luhmann *et al*[15] and Caird *et al*[18] specifically examined children who underwent joint aspiration, while Sultan *et al*[21] examined children who were admitted as an inpatient for an unclear cause of limp. As the prevalence of septic arthritis was relatively higher in study populations by Kocher *et al*[14] and Caird *et al,*[18] the performance of both is likely to be lower when applied to the lower-risk ED population. Additionally, previously published prediction tools have excluded children who had 'later development of rheumatological disease', 'later development of Legg-Calvé-Perthes disease' and 'associated proximal femoral osteomyelitis'. This further undermines their applicability to the ED setting, where this information would not be available at the time of assessment.

Recommendations for future research

Based on our review of the literature, considerable uncertainty remains regarding the optimal approach to evaluating for septic arthritis among children presenting with acute non-traumatic limp to the ED. Future studies should carefully address issues such as explicit definitions for index tests (ie, predictor variables), reporting the interval between the index test and reference standard, blinding outcome assessors to pertinent clinical and biochemical data and reporting follow-up after initial workup for those not undergoing a definitive diagnostic procedure.[31] Incorporation of relevant elements of the QUADAS-2 tool and Standards for Reporting of Diagnostic Accuracy Study guidelines would likely lead to considerable improvement in future studies.[8]

Spectrum bias, variable performance and very select study populations have hampered the widespread uptake of existing clinical risk prediction tools. In order to accurately evaluate the diagnostic performance of clinical or investigation findings and clinical risk prediction tools in the ED setting, it is critical to conduct prospective observational studies. By recruiting children presenting with acute non-traumatic limp before a diagnosis of septic arthritis is established, research investigators would be able to prospectively collect clinical and basic investigation data at the time of initial ED assessment.

While the reference standard for children with septic arthritis has traditionally been a diagnostic joint aspiration with synovial fluid analysis, we acknowledge that such data would not be readily available on all patients who present with an acute non-traumatic limp to the ED. In clinical practice, very few children undergo a diagnostic joint aspiration and most children are managed conservatively.

The natural course of septic arthritis is such that long-term morbidity is likely to occur within days to weeks without prompt joint washout and antibiotic therapy, while non-emergent disorders are likely to self-resolve with time.[5] As such, we suggest considering the use of a follow-up assessment to determine the presence of persistent symptoms, the need for further diagnostic testing and/or invasive procedures for prospective ED-based studies seeking to evaluate for septic arthritis among children presenting with acute non-traumatic limp.

Limitations

We searched only two electronic sources and limited our analysis to English-language articles from 1966 to June 2019. As such, we may have missed older manuscripts or non-English research reports. However, a rcent analysis suggests that the use of additional databases for most systematic reviews may be low yield.[32] Furthermore, with changes to epidemiology (increasing prevalence of methicillin-resistant *Staphylococcus aureus*) and vaccine coverage (*Streptococcus pneumoniae* and *Haemophilus influenzae*) over recent years, it is unclear how applicable older studies are to current clinical practice. Our review

identified that the overall quality of the literature relevant to paediatric septic arthritis in the ED setting was poor to moderate, resulting in summary estimates that may be biased by several confounders.

Due to the limited number of ED-based studies in this area, children from other populations (ie, children who were admitted as an inpatient for exclusion of septic arthritis or who underwent a joint aspiration) were also included in this review. Pooled estimates of the data could not be performed because of the clinical heterogeneity between studies (ie, different settings, prevalence rates of septic arthritis and exclusion criteria).[33]

As previously discussed, our study objective sought to evaluate for septic arthritis among children presenting with lower limb complaints. Safe differentiation of children with septic arthritis from those with non-emergent disorders, such as transient synovitis, remains a diagnostic challenge for emergency physicians. We acknowledge that septic arthritis can present in other ways, such as the febrile or 'generally unwell' child, however, such studies that evaluated these presenting complaints were beyond the focus of our study objective and consequently not included in our analysis. Thus, the diagnostic utility of historical, biochemical and radiographic parameters for septic arthritis reported in our study cannot be extrapolated to children with non-lower limb presenting complaints.

CONCLUSION

No consensus currently exists on the optimal approach to evaluating for septic arthritis among children with acute non-traumatic limp, and commonly used clinical prediction tools appear unreliable. Clinical laboratory and imaging findings have, to date, been reported using varying definitions and cut-offs, and applied in differing study populations.

The presence or absence of joint tenderness, and an effusion on ultrasound appear to be useful, however, needs to be confirmed in future studies. Spectrum bias and overall poor-to-moderate quality study design limits applicability of currently available research to the acute ED setting. Existing clinical risk prediction tools are hampered by poor performance in external validation studies and very select study populations.

Differentiating children with septic arthritis from non-emergent disorders of non-traumatic limp remains a key diagnostic challenge for emergency physicians. There is a need for clinical risk prediction tools to be prospectively derived and validated in ED-based study populations.

Author affiliations
[1]Monash Health, Melbourne, Victoria, Australia
[2]School of Clinical Sciences at Monash Health, Monash University, Melbourne, Victoria, Australia
[3]Department of Paediatrics and Department of Paediatric Rheumatology, Monash Children's Hospital, Melbourne, Victoria, Australia
[4]Department of Paediatrics, School of Clinical Sciences, Monash University, Melbourne, Victoria, Australia

[5]Austin Health, Melbourne, Victoria, Australia
[6]Pediatric Emergency Department, Monash Medical Centre, Emergency Service, Monash Health, Melbourne, Victoria, Australia
[7]Emergency Research, Murdoch Children's Research Institute, Parkville, Victoria, Australia

ORCID iD
Simon Craig http://orcid.org/0000-0003-2594-1643

REFERENCES

1 Fischer SU, Beattie TF. The limping child: epidemiology, assessment and outcome. *J Bone Joint Surg Br* 1999;81:1029–34.
2 Jain N, Sah M, Chakraverty J, *et al*. Radiological approach to a child with hip pain. *Clin Radiol* 2013;68:1167–78.
3 Herman MJ, Martinek M. The Limping child. *Pediatr Rev* 2015;36:184–97.
4 Altemeier WA. Pediatric orthopedic infections: missed diagnoses. *Pediatr Ann* 1999;28:718–20.
5 Tse SML, Laxer RM. Approach to acute limb pain in childhood. *Pediatrics in Review* 2006;27:170–80.
6 Beach R. Minimally invasive approach to management of irritable hip in children. *The Lancet* 2000;355:1202–3.
7 McCanny PJ, McCoy S, Grant T, *et al*. Implementation of an evidence based guideline reduces blood tests and length of stay for the limping child in a paediatric emergency department. *Emergency Medicine Journal* 2013;30:19–23.
8 Whiting PF, Rutjes AWS, Westwood ME, *et al*. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36.
9 Cohen J. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychol Bull* 1968;70:213–20.
10 Willis BH. Spectrum bias--why clinicians need to be cautious when applying diagnostic test studies. *Fam Pract* 2008;25:390–6.
11 Faesch S, Cojocaru B, Hennequin C, *et al*. Can procalcitonin measurement help the diagnosis of osteomyelitis and septic arthritis? A prospective trial. *Ital J Pediatr* 2009;35:33.
12 Heyworth BE, Shore BJ, Donohue KS, *et al*. Management of pediatric patients with synovial fluid white Blood-Cell counts of 25,000 to 75,000 Cells/mm3 after aspiration of the hip. *J Bone Joint Surg Am* 2015;97:389–95.
13 Kocher MS, Zurakowski D, Kasser JR. Differentiating between septic arthritis and transient synovitis of the hip in children: an evidence-based clinical prediction algorithm. *J Bone Joint Surg Am* 1999;81:1662–70.
14 Kocher MS, Mandiga R, Zurakowski D, *et al*. 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;86:1629–35.
15 Luhmann SJ, Jones A, Schootman M, *et al*. Differentiation between septic arthritis and transient synovitis of the hip in children with clinical prediction algorithms. *J Bone Joint Surg Am* 2004;86:956–62.
16 Deanehan JK, Kimia AA, Tan Tanny SP, *et al*. Distinguishing Lyme from septic knee monoarthritis in Lyme disease-endemic areas. *Pediatrics* 2013;131:e695–701.
17 Deanehan JK, Nigrovic PA, Milewski MD, *et al*. Synovial fluid findings in children with knee monoarthritis in Lyme disease endemic areas. *Pediatr Emerg Care* 2014;30:16–19.
18 Caird MS, Flynn JM, Leung YLEO, *et al*. Factors distinguishing septic arthritis from transient synovitis of the hip in children. *J Bone Joint Surg Am* 2006;88:1251–7.
19 Baldwin KD, Brusalis CM, Nduaguba AM, *et al*. Predictive factors for differentiating between septic arthritis and Lyme disease of the knee in children. *J Bone Joint Surg Am* 2016;98:721–8.
20 Levine MJ, McGuire KJ, McGowan KL, *et al*. Assessment of the test characteristics of C-reactive protein for septic arthritis in children. *J Pediatr Orthop* 2003;23:373–7.
21 Sultan J, Hughes PJ. Septic arthritis or transient synovitis of the hip in children: the value of clinical prediction algorithms. *J Bone Joint Surg Br* 2010;92:1289–93.
22 Del Beccaro MA, Champoux AN, Bockers T, *et al*. Septic arthritis versus transient synovitis of the hip: the value of screening laboratory tests. *Ann Emerg Med* 1992;21:1418–22.
23 Jung ST, Rowe SM, Moon ES, *et al*. Significance of laboratory and radiologic findings for differentiating between septic arthritis and transient synovitis of the hip. *J Pediatr Orthop* 2003;23:368–72.
24 Kuda E, Al Wahadneh A. The rational use of a local scoring system to differentiate between septic arthritis & transient synovitis in children. *Qatar Med J* 2000;2000:16.
25 Zamzam MM. The role of ultrasound in differentiating septic arthritis from transient synovitis of the hip in children. *J Pediatr Orthop B* 2006;15:418–22.
26 Lee SK, Suh KJ, Kim YW, *et al*. Septic arthritis versus transient synovitis at MR imaging: preliminary assessment with signal intensity alterations in bone marrow. *Radiology* 1999;211:459–65.
27 Kwack K-S, Cho JH, Lee JH, *et al*. Septic arthritis versus transient synovitis of the hip: gadolinium-enhanced MRI finding of decreased perfusion at the femoral epiphysis. *American Journal of Roentgenology* 2007;189:437–45.
28 Strouse PJ, DiPietro MA, Adler RS. Pediatric hip effusions: evaluation with power Doppler sonography. *Radiology* 1998;206:731–5.
29 Putto A, Ruuskanen O, Meurman O, *et al*. C reactive protein in the evaluation of febrile illness. *Arch Dis Child* 1986;61:24–9.
30 Uzoigwe CE. Another look: is there a flaw to current hip septic arthritis diagnostic algorithms? *Clin Orthop Relat Res* 2014;472:1645–51.
31 Kaji AH, Schriger D, Green S. Looking through the retrospectoscope: reducing bias in emergency medicine chart review studies. *Ann Emerg Med* 2014;64:292–8.
32 Hartling L, Featherstone R, Nuspl M, *et al*. The contribution of databases to the results of systematic reviews: a cross-sectional study. *BMC Med Res Methodol* 2016;16:127.
33 Kim KW, Lee J, Choi SH, *et al*. Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical Researchers-Part I. General guidance and tips. *Korean J Radiol* 2015;16:1175–87.