# The diagnostic test accuracy of serum Procalcitonin compared to C-reactive protein for bone and joint infection in children and adolescents: A Systematic Review and Meta-analysis.

Dr Brett Kingsley Ritchie, BSc (Hons), MB.BS

Master of Clinical Science

JBI, Faculty of Health and Medical Sciences, The University of Adelaide, Adelaide, Australia

December 2020

# Table of Contents

# Abstract

## Objective:

To synthesise the best available evidence for the diagnostic test accuracy of measurement of serum procalcitonin compared to serum C-reactive protein for suspected osteomyelitis and septic arthritis in hospitalised children and adolescents.

## Introduction:

Measurement of serum C-reactive protein remains a routine investigation for the diagnosis of osteoarticular infection in children and adolescents. Measurement of serum procalcitonin has been shown to outperform C-reactive protein in adults with osteomyelitis and septic arthritis. Before procalcitonin can be considered as a potential replacement or additional test in children and adolescents, a systematic review and meta-analysis targeting this population is needed.

## Inclusion criteria:

Original studies reporting on the diagnostic accuracy of procalcitonin and/or C-reactive protein in children and adolescents aged between one month and 18 years admitted to hospital with suspected osteoarticular infection were included, compared to at least one reference test. The reference test was defined as positive culture or polymerase chain reaction confirmation of a pathogen from blood and/or bone biopsy and/or joint fluid aspirate and/or at least two of the following: 1) purulent material from sterile site; 2) positive radiological findings consistent with osteoarticular infection; 3) symptoms and signs consistent with osteomyelitis and/or septic arthritis.

## Methods:

JBI methodology for systematic reviews of diagnostic test accuracy was employed. Information was sourced from four databases; MEDLINE, Embase, Cochrane Central Register of Controlled Trials, and Web of Science and four grey literature sources; Mednar, OpenGrey, Google scholar and ProQuest Dissertations and Thesis. Only studies published in English were considered. The methodological quality of selected studies was formally evaluated and sensitivity and specificity data were extracted and 95% confidence intervals determined. Meta-analysis was performed to estimate summary points using a bivariate model and generate a hierarchial summary receiver operating characteristic curve (HSROC) with global measures of test accuracy performance including likelihood ratio and diagnostic odds ratio. A narrative summary was provided where meta-analysis was not feasible.

## Results:

Eight out of 3086 studies were included in the final analysis. Four of these studies used a common CRP test threshold of 20mg/L for septic arthritis cases only. At this threshold the estimated pooled sensitivity of C-reactive protein was 0.86 (0.68-0.96) and the pooled specificity 0.9 (0.83-0.94). Using

a HSROC model from six studies including all osteoarticular infections, the diagnostic odds ratio for C-reactive protein was estimated to be 39.4 (14.8-104.9) with a positive likelihood ratio 5.3 (2.3-11.9) and negative likelihood ratio 0.1 (0.07-0.2). There were insufficient studies from this review to statistically evaluate the diagnostic accuracy performance of procalcitonin using meta-analysis.

## Conclusion:

We have synthesised the best available evidence to evaluate the diagnostic test accuracy of serum measurement of procalcitonin and C-reactive protein in children and adolescents with suspected osteomyelitis and septic arthritis. Clinicians should continue to measure serum C-reactive protein as the preferred inflammatory marker in this setting and await more evidence before incorporating procalcitonin routinely into their diagnostic test strategy for this specific setting.

# Declaration

I, Brett Kingsley Richie, certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where this has been referenced in the text. In addition, I certify that no part of this work will, in the future be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without prior approval of The University of Adelaide, and where applicable, any partner institution responsible for the joint-award of this degree.

I give permission for the digital version of my thesis to be made available on the web via the University digital research repository, the library search and also through the web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Dr Brett Kingsley Ritchie
5th December 2020

# Acknowledgement

I would like to thank my supervisors, Dr Kylie Porritt, Ms Tania Marin and Associate Professor Nicole Williams for their constant guidance, support and encouragement during my candidature. Their close collaboration provided a challenging and thoroughly enjoyable learning environment.
I would also like to sincerely thank Dr Soumya Soumya as my second reviewer for critical appraisal, librarian Ms Natalie Dempster for her assistance and direction with the search strategies and Dr Suzanne Edwards for providing statistical advice.

Finally, I would like to dedicate this thesis to my wife Petra and my daughters Meg and Poppy for their love, support and inspiration to keep me going.

# 1 Introduction

Chapter one provides an outline of the thesis structure and context of this review. The research question is stated and the methodological basis for the research approach implemented is discussed. Finally, the rationale for performing this systematic review is provided along with the researcher's background and expertise in this area. Based on the work from this thesis a systematic review protocol has been published and the findings from the completed systematic review has been accepted for publication.[1]

## 1.1 Structure of the thesis

This thesis has four chapters. Chapter one identifies the research problem, proposes the research question and then provides a pathway forward. Chapter two gives a detailed account of the methodology and methods employed by the researcher and in chapter three the results are presented. In chapter four the significance of the findings and implications in advancing scientific knowledge in this field is discussed.

## 1.2 Context of the review

Choosing Wisely® is an initiative of the American Board of Internal Medicine Foundation launched in 2012 that aims to encourage meaningful conversations between health care practitioners and clients to ensure that appropriate and quality of care in being provided.[2] Specifically the goals are to promote interventions and investigations that are supported by evidence, not duplicative of other tests or procedures already received, free from harm and truly necessary. Choosing Wisely campaigns now exist throughout many countries.[2] Choosing Wisely Australia® is an initiative of the Australian National Prescribing Service Limited in partnership with Australia's health professional colleges, societies and associations.[3] A feature of this national campaign is to enable doctors to start important conversations about the use of laboratory tests in medicine with an emphasis on utilising evidence to demonstrate the benefits of, and/or highlight potential risks from, the ordering of medical tests. The Royal Australian College of Physicians, division of Paediatrics and Child Health have already commenced work on this by providing their fellows with recommendations around a small number of tests for children and adolescents. The Australasian Society for Infectious Diseases has also made a smaller contribution for infectious diseases testing in both adults and children. The recommendations from these professional organisations aim to be evidence-based, developed and approved using a transparent process and focused on tests that are cost-effective and/or are used frequently across our healthcare sector.

While healthcare professionals have an important responsibility to ensure the tests they request are evidence-based and add value to their patient's healthcare journey, the consumer can also play a significant role in this process. Engagement of consumers is to be encouraged by asking doctors for more information in five key areas: the need for any test; the risks associated with the test; the

availability of simpler or safer options; consideration of not having the test; and where relevant, the financial cost of the test. In Australia medical tests that are not listed on the Medicare Benefits Schedule are ineligible for government rebate making the cost of performing these more expensive to the individual or health care organisation.

The Choosing Wisely Australia® campaign is in its early stages. It is too soon to determine if it will be successful in improving the patient experience of care by avoiding unnecessary tests and/or reducing financial cost to our healthcare system. Increasing awareness of this campaign and compliance with the key messages for both healthcare professionals and consumers will be essential for widespread implementation.

For clinical doctors working in the public hospital system, the ordering of tests for prognosis, diagnosis and monitoring of treatment forms an integral part of their daily work. For doctors who are required to make rapid clinical decisions based on test results, the availability of onsite 24-hour laboratories processing these urgent tests is essential.

At the Women's and Children's Hospital (WCH) in Adelaide, laboratory services are delivered by the South Australian government pathology service (SA Pathology®) which provides both onsite and local laboratories for patients.

In 2018, the WCH introduced and implemented a hospital wide "early sepsis recognition" program aimed to improve both the diagnosis and treatment of sepsis in children and women. As part of this program, a clinical protocol recommended doctors request the serum measurement of a new biological marker to aid in the diagnosing of sepsis, called procalcitonin (PCT). This new biomarker test has shown to be more sensitive and more specific for sepsis and serious bacterial infections compared to other routine inflammatory markers in use, including C-reactive protein (CRP). The recommendation for use of this test was based on published data from systematic reviews and meta-analyses.[4-9] SA Pathology® agreed to provide this laboratory service using a rapid kit test at an estimated cost of approximately 10 times that of CRP. While the aim of the protocol was to restrict the use of PCT in patients with sepsis or suspected sepsis only, an unexpected and incremental increase in demand emerged from clinicians. It became apparent that PCT was being requested for infective conditions other than for patients with suspected sepsis. Without a hospital-wide governance process in place to monitor and restrict laboratory testing, combined with a lack of awareness of the messages from the Choosing Wisely Australia® initiative, PCT was being requested as an additional test for patients presenting across a wide spectrum of suspected paediatric infections.

The clinical impression by the author is that children and adolescents presenting to hospital with suspected bone and joint infections represented one such cohort of patients where PCT was being requested outside the recommended sepsis protocol. The rationale for requesting PCT seems biologically plausible given the diagnosis of osteoarticular infections can be difficult and this test could

potentially improve the post-test probability of having an infection based on the evidence from its use in sepsis. Pressures on doctors resulting from any unnecessary delays in the diagnosis and/or the administration of effective antibiotic therapy may also be an important driver of this requesting behaviour given these infections can lead to permanent and severe sequelae

As there is currently no single reliable and or readily available non-invasive test to confirm or exclude bone and joint infection, clinicians usually perform a cascade of diagnostic investigations. This usually begins with the least invasive and lowest risk tests being performed first and then progressing to more invasive tests and procedures to either confirm or exclude the diagnosis. Included in the preliminary investigation panel of routine blood tests is CRP.[10]
The test accuracy of CRP in bone and/or joint infection has been recently established by two published systematic reviews and meta-analyses but in predominately adult subjects. Both of these studies suggest PCT may outperform CRP as a diagnostic test.[11, 12] While this is encouraging for potential translation to the paediatric setting, to date there has not been a comprehensive and rigorous evaluation specifically in children and adolescents to guide clinicians as to the diagnostic accuracy and utility of PCT. It is therefore important to examine and carefully evaluate the evidence behind the local practice of ordering PCT in bone and joint infection in children and adolescents, before providing any formal recommendations back to WCH clinicians.

The timing of the observed increase in PCT testing at the WCH provided an opportunity to engage and support the national Choosing Wisely initiative at an organisation level and importantly focus clinicians' attentions towards improving laboratory stewardship with the review, assessment and approval process of new diagnostic tests.

## 1.3 Aim

The primary aim of this thesis is to synthesise and analyse the best available evidence for the diagnostic test accuracy of PCT compared to CRP for suspected acute osteomyelitis and septic arthritis in hospitalised children and adolescents.

## 1.4 Osteoarticular infections in children

In children and adolescents, bacteria can infect the bone and joints through a rich blood supply resulting in an acute inflammatory response termed acute haematogenous osteomyelitis (OM) when confined to bone, or septic arthritis if localised to the joint.[13] In developed countries, OM rates vary from two to 13 per 100,000 children and is more common than septic arthritis.[14, 15] Osteomyelitis affects twice as many boys as girls, while the male to female ratio for septic arthritis is 1.2/1.[16, 17] In children and adolescents aged 16 years and under, the median age for OM is four years and one and a half years for septic arthritis.[18] Although any bone or joint can be infected, OM and septic arthritis more commonly affect the lower limbs in children.[13]

The source of bacteria leading to infection is usually not clinically evident, suggesting colonization of the skin; however, oral or respiratory mucosa are the most likely portal of entry.[19] Common pathogens identified in OM and septic arthritis in children and adolescents are *Staphylococcus aureus, Kingella kingae, Streptococcus pneumonia and Streptococcus pyogenes*, depending upon age, risk factors and geographical location.[10, 14, 19, 20] Once infection is established in the bone or joint, a prompt cascade of inflammatory responses ensues, both locally and systemically, with release of various inflammatory factors.[21, 22] Most children who receive early and appropriate therapy for OM and septic arthritis have no long-term sequelae.[13, 23] A delay in treatment or inappropriate antibiotic choice however, can result in complications including disruption to bone growth, angular deformity of bone, cartilage destruction and unstable articular surfaces of the joints all of which can result in permanent damage.[13, 23]

The clinical manifestations of OM can vary and frequently include acute, persistent and increasing pain over the affected bone, reduced weight-bearing, decreased range of movement, and fever.[23] Swelling or redness of the soft tissue over the affected bone may also be seen.[13] Features of septic arthritis overlap with the above presentation findings in addition to specific swelling and erythema of the joint and pain on movement of the isolated joint.[19]

At present there is no single error free diagnostic test that can confirm or rule out OM and or septic arthritis irrespective of the patient's age.[14] Diagnosis continues to require a combination of careful history, thorough physical examination, radiological imaging, laboratory blood tests and microbiological sampling for suspected bacterial pathogens.[14] While conventional x-rays are the first imaging study performed, they are frequently negative for both septic arthritis and early OM.[14] Classic bone changes of OM involving long bones may not be apparent on x-ray until 10-21 days of infection.[21] Ultrasound can detect excess fluid in large joints such as the hip with a sensitivity of 95% and identify collections in subperiosteal areas of bone, if present, with a sensitivity of 55%.[19, 23, 24] Magnetic resonance imaging (MRI) is now the most accurate imaging tool for OM with excellent sensitivity (98%) and specificity (92%).[25] It is however relatively expensive, not universally available, and requires sedation or anaesthetic for younger children.[23]

Laboratory blood tests commonly requested include examination of the total and differential white blood cell (WBC) count, erythrocyte sedimentation rate and CRP. For septic arthritis, microbiological sampling of synovial fluid is routinely collected for cellular microscopy, Gram stain, culture and bacterial polymerase chain reaction (PCR). In children this is usually performed during therapeutic treatment with joint drainage under general anaesthesia.[10]

For uncomplicated OM in children and adolescents however, bone aspiration and/or biopsy to identify bacteria is not routinely performed as it is invasive with associated risks and may not have any therapeutic benefit.[10] Infection in both OM and septic arthritis is usually haematogenous spread in paediatric patients and blood cultures are often collected, with a causative pathogen identified by 48 hours in 10% to 40% of cases.[10, 16]

While there appears to be variation in the definitions of reference standards for OM and septic arthritis in children and adolescents in the published literature, the isolation of a microorganism from the bone, joint and or blood (with a compatible clinical or radiological syndrome), has been proposed as the gold standard in a recent international paediatric bone and joint infection guideline.[10] This definition is consistent with other published findings.[26-28] However, routinely applying this gold standard for all suspected cases may be associated with increased risk from potential harmful operative procedures, may be more expensive, or not readily available and may not improve the diagnostic yield.[29] Therefore, by using less stringent criteria for the gold standard test by excluding isolation of an accepted pathogen for OM, this approach better reflects actual clinical practice, see section 2.3.3 for details.[18, 30, 31]

While OM and septic arthritis can be found as separate conditions in the published literature, from a clinical perspective concurrent infection involving adjacent bone and joint is well described in children and adolescents, and investigators often evaluate them together as a single condition.[13, 18, 19, 28] Studies using MRI suggest concomitant bone and joint infection may occur far more frequently than clinically suspected and in more than one third of paediatric cases.[32, 33]

It is important for clinicians to make timely decisions regarding the diagnosis and treatment of OM and septic arthritis given the potential for serious sequelae. While essential, the clinical history and examination findings cannot always reliably differentiate infective from non-infective causes of bone and joint pathology in children. Furthermore, the initial radiological imaging requested may be normal, and invasive biopsy procedures for specimens when performed take time to process. Specimens taken for microbiological culture of pathogens are positive only in approximately half of cases and results can take up to 24 hours or longer.[34 35] Identification of pathogens using Polymerase Chain Reaction (PCR) techniques in the microbiology laboratory is a more rapid process, but not universally available to most clinicians for bone and joint infections.

Total WBC count however is a rapid and easy test to perform either manually or by automated counter. Unfortunately it is an unreliable marker of infection as WBCs are elevated in only 36% of children with OM, and 30% with septic arthritis can have a normal WBC count.[14, 21, 36] Consequently, clinicians rely on more accurate non-invasive and rapid laboratory investigations such as CRP which is routinely performed.[16, 27]

## 1.5 The biology and use of C-reactive protein serum measurement

Measurement of serum CRP is one of the most frequently requested inflammatory biomarker tests in paediatric clinical practice.[37] It was discovered from human serum during the acute phase of streptococcal pneumonia infection in 1930 and received this name because it reacted with the C-polysaccharide of pneumococcus bacteria to form a precipitate.[38] CRP was the first acute phase protein to be described and is a sensitive systemic marker of tissue damage, infection and inflammation.[39] It is a highly conserved plasma protein synthesized by hepatocytes in the liver and is primarily induced by proinflammatory cytokines released locally at the site of the tissue damage,

infection and/or inflammation but also from systemic release.[39] Interleukin-6 appears to be the main cytokine regulator driving transcription of CRP protein and is reinforced by other pro-inflammatory cytokines including IL-1 *B* and tumour necrosis factor.[40]

Human CRP once released into the circulation, binds with highest affinity to extrinsic phosphocholine residues expressed on cell wall surfaces of bacteria, fungi and parasites as well as plant products. [39] When CRP is bound to these macromolecular ligands it is recognised by C1q and activates the classical complement pathway via C3 leading to opsonisation and phagocytosis of invading pathogens.[39] Both immunoglobulin G and FcγR receptors expressed on numerous cell surfaces including macrophages and lymphocytes can recognise CRP and this aggregation may trigger proinflammatory mediators with further immune cell recruitment, modulation and activation.[40]

In healthy young adults the median serum concentration of CRP from one study was 0.8 mg/L, with a 99$^{th}$ percentile of 10mg/L.[39] Reporting of the 97.5th percentile for measurement of serum CRP in children ranged from 2.8 to 6.3 mg/L in two to 10.9 year old boys and girls.[41] A higher 97.5th percentile level of 11.3 mg/L was reported in a Swedish study in children between the age of 6 months and 18 years.[42]

Following an acute phase stimulus, CRP rises 4-6 hours later duplicating every eight hours from baseline until peaking at 36-50 hours.[38] The plasma half-life is 19 hours and is constant under all conditions of health and disease so that the sole determinant of CRP concentration is the rate of synthesis.[39]

Levels of CRP in the plasma rise in proportion to the intensity of the pathological or inflammatory insult.[43] Markedly, elevated levels are more often associated with bacterial infections rather than viral aetiologies and CRP levels can reach up to 1,000 fold.[44] Elevated CRP levels however are not specific and found in non-infectious states including autoimmune disorders, inflammatory diseases, malignancy, trauma, burns, and myocardial infarction.[45] When the infection stimulus for CRP production falls and ceases following appropriate therapeutic interventions such as antibiotic therapy and/or operative treatment, the serum concentration falls at almost the rate of plasma CRP clearance.[39]

C-reactive protein has been used to assist with the diagnosis of bacterial osteoarticular infections in children and adolescents for decades.[36] A systematic review of haematogenous acute and subacute paediatric osteomyelitis from 2012, found CRP was abnormally elevated on presentation in 80.5% cases and 100% with concomitant septic arthritis; however, no formal meta-analysis on the test characteristics of CRP was performed.[23] For isolated septic arthritis, CRP has shown to be a strong biomarker for diagnosing septic arthritis of the hip in the paediatric setting, while others have suggested the value of CRP measurement may be better as a negative predictor (that is a “negative” result for this test may be more informative to exclude infection rather than to confirm it with a positive

result).[27, 46] More recently, the usefulness of CRP in diagnosing long bone osteomyelitis in children has been questioned.[47]

The popularity of CRP measurement to assist in the diagnosis of bone and joint infections in children and adolescents without robust systematic analysis of its test performance characteristics, may be attributed to familiarity from use, inexpensive cost, rapid turnaround time, predictable kinetics and possible bias given its usefulness for prognosis, monitoring treatment response and guiding the timing of intravenous to oral antibiotic stepdown therapy.[18, 28]

While the general consensus seems to favour inclusion of serum measurement of CRP in the investigation of osteoarticular infections, a systematic review and meta-analysis focusing on the overall diagnostic accuracy in children and adolescents will contribute towards a better understanding of the benefits and limitations of this test and clinical utility in the paediatric setting. This information will also facilitate formal comparison with more novel inflammatory markers of bacterial infection in this setting including serum measurement of PCT.

## 1.6 The biology and use of serum Procalcitonin measurement

Procalcitonin is a 116 amino acid peptide and precursor of the hormone calcitonin and has recently emerged as more specific biomarker for bacterial infection in the paediatric setting.[48] The precise role which PCT plays in the immune response to inflammation is unknown. One study has shown that PCT is a potent amplifier of the inflammatory cascade by increasing expression of surface markers on leukocytes, increasing cytokines and augmenting nitric oxide production.[49] Other studies however have demonstrated that it may neutralise bacterial lipopolysaccharides and reduce both the phagocytic activity of neutrophils and their ability to migrate towards chemo-attractants.[50, 51]

Production of PCT can be activated in all parenchymal tissues in response to bacterial, fungal and some parasite infections mediated by cytokines interleukin-6, tumour necrosis factor and interleukin-1*B.*[48] Conversely, PCT is attenuated by interferon-$\gamma$ primarily released in response to viral infection.[48]

Procalcitonin synthesis is detectable in serum approximately two hours after onset of infection and increases promptly upon stimulation via cytokine production.[52] Peak PCT levels occur within 48 hours and can increase to more than 400 times baseline levels.[53 54] Procalcitonin has a half-life of approximately 24 hours [48] and levels also correlate with the extent and severity of inflammation.[55, 56] The highest serum levels of PCT however are usually seen in acute systemic bacterial infection and sepsis.[57]

Procalcitonin measurement therefore has some advantages over CRP with a more rapid kinetic profile, increasing earlier in response to infection, peaking quicker and showing faster decline with resolution of infection.[48] Serum PCT levels in healthy individuals are exceedingly low and rise in response to the severity of systemic infection; early and or more localised infections however may

result in smaller increments of PCT elevation.[48, 58] Levels of PCT can also be elevated in other non-infective conditions including the first 48 hours of neonatal life, following major surgery, trauma, burns, and cardiac shock.[58]

Quantification of serum PCT can be performed easily using commercially available automated assays or more rapid point of care test kits.[48] The concentration of PCT from the serum of healthy individuals is typically less than 0.1 ug/L.[48, 52, 59] Numerous algorithms have been published using serum PCT measurement across different threshold values (0.05ng/mL-9.33ng/mL) to assist clinicians in the diagnosis of various infectious conditions.[52, 53, 59-61]

In the paediatric published literature, evidence obtained from systematic reviews and meta-analyses support the use of PCT as a routine test where serious bacterial infections may be suspected, but without an apparent source,[8, 9, 62, 63] with PCT outperforming CRP.[8, 63] Procalcitonin has also demonstrated similar advantages in neonatal sepsis showing better diagnostic accuracy than CRP for bacterial infection.[64, 65] Use of serum PCT measurement in other paediatric infections is expanding and includes urinary tract infection, pneumonia, febrile neutropenia as well as bone and joint infections.[58]

## 1.7 Statement of the review question

What is the diagnostic test accuracy performance of serum measurement of PCT compared to CRP in patients aged between one month and 18 years, admitted to hospital with suspected acute osteoarticular infection?

## 1.8 Overview of knowledge synthesis

Knowledge synthesis is the assembly of parts into a new whole, that organises, interprets the concepts, connections, controversies, and constraints of a body of literature filling in gaps, and generating new insights, perspectives, directions and novel explanations about a defined research topic.[66] Traditionally this would involve an authoritative expert in a specialised field publishing a narrative review to provide recommendations supported by subjectively chosen evidence.[67] In the 1960's knowledge syntheses were common within social sciences, education and psychology, however the widespread application and demand for evidence synthesis in medicine soon followed.[68]

While many models of knowledge synthesis exist, in the healthcare setting systematic reviews evolved focused on rigorous methodology originally developed by the Cochrane Collaboration. This synthesis consists of a clearly formulated question and the use of systematic and explicit methods to identify, select, critically appraise and extract and analyse data from relevant research.[69] What fundamentally sets systematic reviews apart from the traditional narrative reviews are the formal

methods aimed to minimise risk of bias and increase the reliability and accuracy of the conclusions drawn.[67]

Systematic reviews are the gold standard to search for, collate, critique and summarise the best available evidence regarding a clinical question.[70] They are also regarded as the pillar of evidence-based health care.[70] There are a number of fundamental steps in performing a systematic review and these have been well described in the published literature.[71] Determining the appropriate type of systematic review first requires careful consideration of the research question. Formulating the relevant questions typically found in an effectiveness review requires a structured approach and includes basic elements such as the population(s) of interest, the intervention(s), the comparison(s) and the outcome (s).[72] To improve the value of the systematic review, the study question ideally should be clearly and completely defined and should be clinically important with sufficient uncertainty or debate in the literature.[73]

Once the need for a systematic review has been established, the next step is to define the objectives using a priori protocol specifying the explicit methods that will be used to guide this process. This includes a detailed description of the methods used to search, retrieve, appraise the literature and how the data is to be extracted, synthesised and analysed. This should follow the Preferred Reporting Items for Systematic reviews and Meta-analysis Protocols (PRISMA-P) guidelines.[74] The purpose of these guidelines is to decrease the likelihood of bias from post hoc changes and analyses and selective reporting in the review.[69] They also help to facilitate reproducibility of the findings.[67] Publication of the protocol in a peered review journal and registration of the protocol in the International Prospective Register of Systematic Reviews (PROSPERO) data base increases transparency and reduces unnecessary duplication of any work already being done on this topic.[67]

For systematic reviews, the aim of the search strategy is to find all of the available evidence that can be used to answer the particular question proposed.[75] Therefore the search strategy should be as comprehensible and as inclusive as feasible. For clinical questions this usually includes searches of all of the major electronic data bases for both published literature as well as the unpublished grey literature to help minimise publication bias.[69] The search strategies recorded should be replicable to allow updating and repeating by other researchers.

The next step is the selection of relevant studies to address the research question. This is usually conducted by first screening by title and abstract of the citations and then a second screen is performed in detail by examining the full text articles using predetermined criteria for eligibility and inclusion. Reasons for excluding full text studies are documented for transparency and reproducibility.

Included studies are assessed by two reviewers to evaluate the methodological quality of each study. The validity of the results of the systematic review will greatly depend on the risk of bias in the individual studies.[69] This process is standardised using published validated tools depending upon the

type of systematic review.[67, 73] Quality assessment using the instruments available however still requires subjective judgment by the assessor and disagreements may require a further reviewer to assist in reaching a conclusion.[76] The results of quality assessment are best summarised in table or graphical format to provide readers with an overall impression of the quality of the data and help assess the effect of any bias on the results.[75]

Finally, the data is extracted and synthesised. All data to be extracted from individual studies should be clearly defined in the protocol and can be recorded using specific data extraction tools.[69] This information can then be abstracted into a single database. Synthesis and analysis of the evidence is dependent upon the question being asked and the variables collected. For quantitative data this may or may not include meta-analysis.

Meta-analysis is a statistical method for combining the results of multiple independent primary studies that address the same question to produce a pooled or summary quantitative estimate of the outcome of interest.[77] A key benefit of this approach is the aggregation of quantitative data which may lead to higher statistical power and a more robust point estimate with greater precision than is possible from the measure derived from any individual study (providing the studies are of equal level of evidence). [78] If meta-analysis is not found to be feasible, most commonly from heterogeneity and/or studies being thought to be too different, then a structured systematic review process may still provide a more objective appraisal of the available evidence than a traditional narrative review. [79]

In addition to the statistical advantages provided, a systematic review with a meta-analysis approach may better establish the generalisability of scientific findings as multiple studies provides an interpretive context not available from single studies.[80] Furthermore consistency and inconsistency is more easily observed across individual studies using this approach.[78] High quality systematic reviews can also define the boundaries of what is known and gaps in current understanding that may help determine the direction of future research.[81]

Systematic reviews and meta-analysis of diagnostic test accuracy aim to summarise the test accuracy performance of an index test(s) with a reference or gold standard. Diagnostic test outcome measures relevant to clinical practice usually include sensitivity, specificity, likelihood values and receiver operating characteristic curves. This information is important to clinicians and other healthcare professionals to make informed decisions regarding the optimum test use in practice.[82] While the approach outlined above remains this same for diagnostic test accuracy reviews, there are some unique study characteristics that require a modified approach. For this reason the Preferred Reporting Items for Systematic reviews and Meta-analysis (PRISMA) statement was extended and published as the PRISMA-DTA to help guide researchers improve the quality and consistency of diagnostic test reviews.[71] PRISMA-DTA addresses several detailed and specific test related requirements that are not included in the PRISMA statement.

## 1.9 Rationale for this systematic review/meta-analysis

The rationale for evaluating both index tests in this systematic review and meta-analysis was to provide clinicians with evidence regarding the performance of the routine measurement of CRP compared with measurement of PCT. This would be an essential prerequisite for any future recommendation around the use of these diagnostic tests in practice at the WCH and other similar health care settings. As OM and septic arthritis frequently co-exist in children and adolescents, it was considered clinically relevant to evaluate the diagnostic accuracy of PCT for both conditions into a single systematic review and meta-analysis.

While consideration was given to conducting a diagnostic study to compare both diagnostic tests directly in our local health setting, a systematic review was considered to be a quicker and less costly pathway to achieve the desired outcome.[78] Even if an additional single diagnostic accuracy study could be performed for CRP and PCT, the results seen may not be representative of the overall test performance and the recommendations may be erroneous. [83] Causes of variation when evaluating sensitivity and specificity of diagnostic test outcomes has been well documented and includes, random chance, differences in thresholds for test positivity, deficiencies in design and conduct, diversity of settings and population, and unexplained variation.[76] Biased results can lead to poor recommendations regarding test accuracy performance and may affect patient outcomes. [83] Undertaking a systematic review with meta-analysis approach may better establish the generalisability of the performance of both index tests and identify the direction and magnitude of outcome measures, providing of course there are sufficient studies to be included. It may also help identify inconsistencies or conflicts from observed any outcome measures.[76]

Based on the research question proposed, a systematic review of diagnostic test accuracy was considered to be the most appropriate methodology to evaluate the test performance outcomes of both CRP and PCT.[70]

A preliminary search of MEDLINE, Cochrane Database of Systematic Reviews, and the JBI Database of Systematic Reviews and Implementation Reports was conducted as a first step to identify any previously published systematic review with meta-analyses on this topic. This resulted in identifying two systematic reviews with meta-analyses examining the diagnostic performance of PCT in osteoarticular infections. In both of these studies the authors conclusions favoured PCT over CRP as a diagnostic test.[11, 12] In the first published study from 2013 examining bone and joint infections, of the seven studies evaluated, only two included children, with conflicting results.[11] While the investigators included all children, variability in and compliance with the gold standard or reference test was observed. Furthermore, errors associated with the referencing numbers in citing of the paediatric papers permitted the possibility of misinterpretation of their overall findings.

In the second published systematic review and meta-analysis from 2017, ten studies examining septic arthritis cases only were included in the final analysis to assess the diagnostic accuracy of PCT.[12] The same two paediatric papers as reported in the initial review above were included, and one of these papers was incorrectly classified and included in the statistical analysis as an adult study. Furthermore, inconsistencies regarding the inclusion study flow chart and concerns of a limited search strategy, potentially biased the outcomes of their findings.[84] Both of these systematic reviews searched PubMed, Embase and the Cochrane Library, but a grey literature search was not performed. On further searching, a proposed systematic review of serum PCT levels as a diagnostic marker of joint infection was found registered in PROSPERO.[85] Upon contact with the registering author by email, this review had not and was unlikely to progress.

A subsequent but limited exploratory search for individual publications in the paediatric literature revealed fewer PCT diagnostic test accuracy studies with smaller patient numbers compared to CRP. Furthermore, most of the studies did not evaluate both CRP and PCT directly to the reference test in a single patient. The most feasible pathway forward to increase sample sizes and optimise the power and precision of any future estimates of diagnostic test accuracy of both tests was considered to be a new systematic review with meta-analysis.[81] In addition emerging network meta-analysis techniques were considered to be potentially useful by allowing studies using only one index test to be included and compared indirectly to the reference standard. It was ascertained that providing trials included were sufficiently similar, using this methodology should further increase the probability of including more eligible studies into the final formal pooled meta-analysis.[86]

## 1.10 Researcher's experience in this field

The researcher is a paediatric infectious diseases specialist, currently working within the Infectious Diseases Department at the WCH in Adelaide, South Australia. Between 20% and 25% of his current inpatient consultative workload involves working with orthopaedic specialists in the microbiological diagnosis and antimicrobial management of acute osteoarticular infections in children and adolescents admitted to hospital. As this type of clinical work involves frequent use of routine tests, it is important that both disciplines remain up to date and consistent with an agreed approach and where possible use the best available evidence on which to make recommendations regarding practice change. While the researcher has not published in this specific field, he has accumulated sufficient experience in this area to provide a level of understanding and clinical perspective for such a review. Having had exposure to both CRP and more recently PCT testing in our facility, the results from this systematic review should have direct relevance to all clinicians working within similar healthcare settings.

## 1.11 Conclusion

Clinicians have a responsibility to ensure their request for performing measurement of PCT from an individual patient is evidence-based for their specific indication. There is sufficient uncertainty as to the value of measuring PCT in children and adolescents presenting to hospital with osteoarticular disease. Examination of the available scientific literature to answer the research question may be best presented using systematic review methodology.

# 2 Systematic review protocol

## 2.1 Introduction

Chapter two provides the protocol framework for this systematic review. A detailed description of the essential inclusion criteria for study selection is first provided. This is followed by the search strategy steps employed in accordance with the JBI methodology for systematic reviews of diagnostic test accuracy and documentation of the information sources searched.[82] The process of study selection, assessment of study quality and data extracted is then provided. Finally, the methods for data synthesis and analysis, and for assessing certainty of the findings will be described.
The process follows an a priori protocol was published in JBI Evidence Synthesis and registered in PROSPERO, number CRD42019140276. [1]

## 2.2 The review objective

This objective of this systematic review was to identify, evaluate, extract and synthesise the best available scientific evidence to determine the diagnostic test accuracy of serum measurement of PCT compared to CRP in children and adolescents aged between one month and 18 years, admitted to hospital with suspected osteoarticular infection.

## 2.3 Inclusion criteria

### 2.3.1 Participants

Participants were children aged one month to 18 years admitted to hospital for suspected bone and/or joint infection. Studies reporting on both adults and children and where the data were reported separately for each group, were included. Patients with chronic osteoarticular infection and/or who were immunocompromised for any reason were excluded.

### 2.3.2 Index Tests

Original studies that measured, reported on and evaluated the diagnostic accuracy of both PCT and CRP index tests in the same group of participants and studies that tested only one biomarker were evaluated. There was no limitation based on the test methodologies used for the different commercially available index tests.

### 2.3.3 Reference Test

This review considered studies that defined the reference standard for OM and/or septic arthritis as:

a) a positive bacterial culture or PCR conformation of an accepted pathogen from blood, biopsy or aspirate and/or
b) having at least two of the following criteria:
    i) purulent material from biopsy or aspirate specimen,
    ii) positive radiological findings consistent with osteoarticular infection, or

iii) signs/symptoms consistent with OM/septic arthritis

### 2.3.4 Diagnosis of Interest

Studies where the diagnosis of interest was acute presumed haematogenous OM and/or septic arthritis were considered in this review.

### 2.3.5 Types of studies

This review considered any quantitative design study published in the English language that examined the diagnostic accuracy of serum measurement of PCT and or CRP for OM or septic arthritis where the participant had one or both index tests and reference test performed.

### 2.3.6 Types of outcomes

Any English language study that quantified the diagnostic test accuracy of measurement of PCT and/or CRP (sensitivity and specificity) compared to the reference standard for OM and septic arthritis was considered.

## 2.4 Review Methods

This systematic review was conducted in accordance with the JBI methodology for systematic reviews of diagnostic test accuracy and also compliant with the recommendations outlined in the PRISMA-P and PRISMA-DTA guidelines.[74, 82, 87]

## 2.5 Search strategy

A comprehensive three-step search strategy was first performed on the 30th July 2019, to locate both published and unpublished studies written in English. The initial limited search (stage 1) was undertaken in MEDLINE (PubMed), followed by an analysis of text words contained in the title and abstracts, and of the index terms used to describe articles. This informed the development of a comprehensive search strategy that was adapted for each additional data base searched (stage 2). This search strategy is shown in Appendix 1. The reference lists of all studies selected for critical appraisal were also screened for any additional eligible studies not previously identified (stage 3). Unpublished studies and grey literature were also included.

### 2.5.1 Information sources

The databases and platforms searched included: MEDLINE (PubMed), Embase (Ovid®), Cochrane Central Register of Controlled Trials (Ovid®) and Web of Science. Sources of unpublished studies and grey literature searched included Mednar, OpenGrey, ProQuest Dissertations and Theses and Google Scholar. The search statements used for each database are detailed in Appendix 1.

### 2.5.2 Study selection

Following completion of the search strategy, all identified citations were collated and uploaded into EndNote version X9 (Clarivate Analytics, PA, USA) and any duplicates removed. Titles and abstracts of individual studies were initially screened for suitability against the inclusion and exclusion criteria for this systematic review. All potentially relevant studies were then retrieved in full text. All full text articles were formally assessed in detail for eligibility against all of the inclusion/exclusion criteria in section 2.2. Reasons for exclusion of any full text study that did not meet the inclusion criteria were recorded and are reported in Appendix 2.

### 2.5.3 Assessment of Methodological Quality

Citations of selected studies after the full text review, were imported into the JBI System for the Unified Management, Assessment and Review of Information (JBI SUMARI). Two independent reviewers critically appraised each study for their methodological quality. This was based on ten signaling questions from the revised Quality Assessment of Diagnostic Studies (QUADRAS-2) critical appraisal tool for diagnostic test accuracy.[88] The four domains used to assess risk of bias included patient selection (questions 1-3), index tests (questions 4-5), reference standard (questions 6-7) and flow and timing (questions 8-10). Disagreements between the reviewers was resolved through discussion. The combined reviewer results from the critical appraisal process were exported directly from JBI SUMARI in tabulated format with a yes, no or unclear response to the questions from each study. Concerns regarding applicability of the included studies was also assessed. Judgements with respect to patient selection (severity, demographics, comorbidities, setting), index test (variation in technology, execution, interpretation) and target condition (matching review question population) were reported. A summary critical appraisal table of bias and applicability for each study was then constructed.

### 2.5.4 Data Extraction

The data related to the diagnostic test accuracy of the index tests being investigated were extracted using a standardised tool.[82] Sensitivity and specificity were chosen as principle discriminative measures of diagnostic accuracy for both index tests in this review. Data were extracted from all included studies either directly where 2 x 2 tables were provided or indirectly, using the raw numbers to construct the contingency table. Index test threshold levels were also extracted from each included study and where common thresholds were identified, the outcomes were evaluated using meta-analysis. Additional data extracted included the country where the study was performed, year of publication, basic patient demographic characteristics (e.g. age range), target and comparator conditions, clinical setting, study design and funding sources.

### 2.5.5 Data Synthesis

A statistician was consulted to assist with the statistical synthesis and analysis component of this systematic review. All statistical data analyses were performed using STATA version 15.0 (Stata

Corp, College Station, Texas, USA). This included presentation of sensitivity and specificity data, likelihood ratios (LR), odds ratio (OR) and predictive values with their respective 95% confidence interval (CI). This statistical data was calculated and tabulated for all included studies.

### 2.5.6 Meta-analysis

A random effects meta-analysis was performed in STATA where more than three studies for each index test reported on the same outcome. This was to provide sufficiently diverse data for statistical convergence in STATA modelling. Where meta-analysis was performed, sensitivity and specificity results were graphically displayed. For those diagnostic tests sharing a common threshold level, a Forest plot was generated and where thresholds varied a hierarchical summary receiver operating (HSROC) curve was displayed. Results of all meta-analyses were presented with 95% CIs.

The magnitude of heterogeneity was reported qualitatively as depicted graphically from the Forest plots and by the degree of closeness of the scattered observed study results from the HSROC curve. Where meta-analysis was not possible, a synthesis of findings was presented in a narrative format.

## 2.6 Conclusion

The protocol for this systematic review was published and registered in PROSPERO prior to knowledge of all of the available studies. This was important to minimise author bias at each step of the process, facilitate transparency in the methodology and provide peer review of the planned approach.

# 3 Results

## 3.1 Introduction

Chapter three provides the results from the systematic review and meta-analysis. The first section describes the outcome of the search with a description of the studies found. The middle section evaluates the methodological quality of the included studies and the final section provides the findings of the review including statistical analysis.

## 3.2 Description of Studies

### 3.2.1 Search and study selection process

A total of 3,086 articles were identified through the systematic search strategy conducted for each of the information sources documented (see Appendix 1). No other records of information were sourced. After removing 913 duplicates, a total of 2,173 articles remained. From screening of title and abstract only, a total of 2,072 records were excluded leaving 101 articles to be retrieved in full text. Following a detailed examination of full text, a further 93 records were excluded leaving eight published studies meeting all the inclusion and exclusion criteria. A flow chart of the search and study selection process is presented in Figure 1. The reason for exclusion of full text records have been categorized and summarised in Table 1. Refer to Appendix 2 for full citations of the 93 articles with individual reasons for their exclusion.

Table 1. Reasons for exclusion of records.

| Reasons for exclusion | Number of articles |
|---|---|
| Summary, review or guideline article | 14 |
| Insufficient data provided for accuracy test | 11 |
| Language other than English | 11 |
| Not a diagnostic test study | 10 |
| PCR or PCT not included | 8 |
| Sensitivity only provided | 8 |
| Included neonates and data not analysed separately | 5 |
| Included adults and data not analysed separately | 8 |
| Not limited to bone and joint infection | 7 |
| Reference standard not meeting criteria | 5 |
| Letter or case report only | 6 |

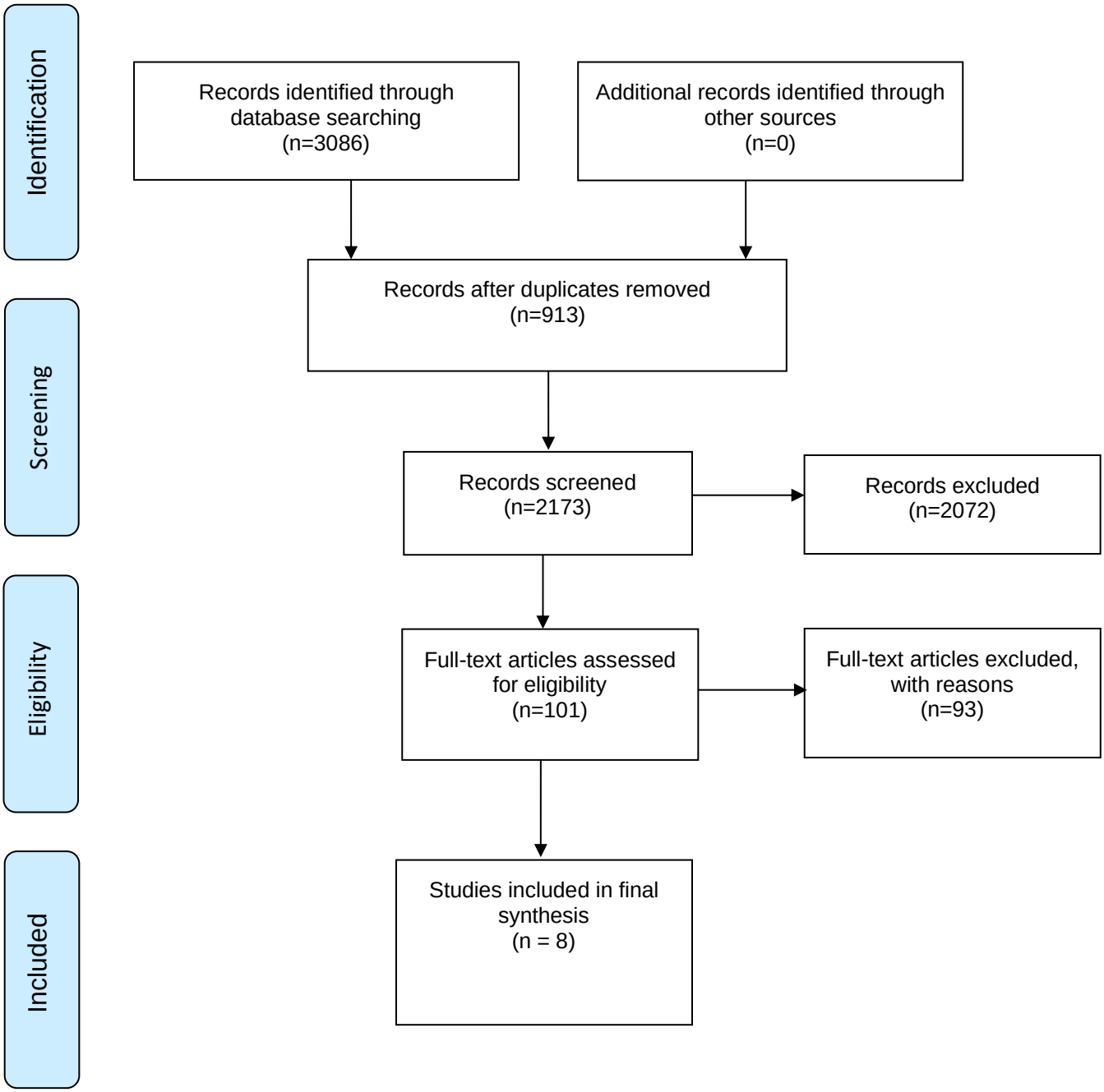


Figure 1. PRISMA flow diagram of search and study selection process.[89]

### 3.2.2 Characteristics of included studies

All of the eight studies in this review were published in peer reviewed medical journals.[31, 46, 90-95] The sample sizes varied from 33 to 313 with a total of 1,520 participants. As anticipated from the reference test criteria of bone and joint infection documented in the protocol, differences were observed in the reference test across the eight studies. Furthermore, there was no uniform consistency with the index test thresholds applied in these studies. Only two studies used a paired study design comparing both index tests (CRP and PCT) directly with the reference test in an individual patient.[31, 93] The key characteristics of each included study are summarised in the included studies table
(Table 2).

#### *3.2.2.1 Studies including CRP and reference test*

Five studies included an evaluation of the measurement of serum CRP in septic arthritis of the hip in children and adolescents.[46, 90-92, 94] These five studies were relatively uniform with respect to the population, clinical context, hospital setting and the index test. The test threshold for CRP was set at 20mg/L for CRP for four of these studies.[46, 91, 92, 96] In the remaining study the cut off was 10mg/L.[94] Variability between and within these studies with respect to the reference test was noted as not all patients received the same investigations. All studies were conducted in specialised tertiary hospitals with paediatric expertise. Only one article of the five was a prospective study.[46] The studies were all consecutive cohort selected cross sectional design, that is patient selection was based on suspicion of joint infection in these cases. Tables providing raw data for calculation of diagnostic test accuracy outcomes were available in four studies.[46, 91, 94 96] The remaining study required extraction of data from the text in the results section.[92] Participant numbers varied across the studies 33 to 311. Studies in this group were published between 2003 and 2011. The study flow is shown in Figure 2 below.

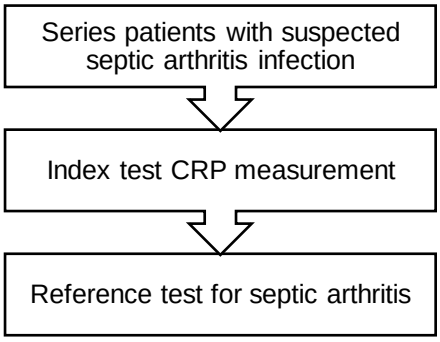


Figure 2. Representation of study flow for CRP

Table 2. Key study characteristics

| First author, year publication | Country | Participants N % female age range | Context | Clinical setting | Study design | Target condition | Index test | Reference test | Funding/COI |
|---|---|---|---|---|---|---|---|---|---|
| Caird MS, 2006.[46] | UK | N=48<br>52.0 %<br>7 months – 16.0 years | Presenting to clinic or emergency department with history suspicious for septic arthritis having an ultrasound and hip aspiration. | Tertiary children's hospital | Prospective consecutive cohort selected cross sectional study | Septic arthritis hip | Serum CRP>20 mg/L | WBC > 50.0 x $10^9$/L from hip aspirate ± culture positive from hip or blood OR bacteria seen on Gram stain from hip aspirate.<br>All 48 underwent hip aspiration | No grants or outside funding received |
| Cui C, 2017.[95] | China | N=267<br>44.4%<br>Average age 12.4 years | Presenting to hospital with suspected cases of osteomyelitis (187) and health volunteers (80). In cases suspected acute osteomyelitis. Based on mean weight and SD of groups, >99.7% probability of no neonates included in this study. | Hospital | Prospective case control selected cross sectional study | Acute osteomyelitis | Serum PCT ≥ 3.56 ng/mL | Bone biopsy culture | Nil funding recorded.<br>No COI |
| Eich GF, 1999.[92] | Switzerland | N=89<br>41.5%<br>1 month – 12.3 years | Presenting to children's hospital with acute hip pain. | University Children's hospital | Retrospective consecutive cohort selected cross sectional study | Septic arthritis hip | Serum CRP≥ 20mg/L | Pus aspirated from the hip joint and/or growth of pathogenic bacteria from the aspirate. (18 patients had joint aspiration) | Nil recorded |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Faesch S, 2009.[31] | France | N=339<br><br>1 month-14.0 years | Children presenting with non-traumatic decreased motion of a skeletal segment. | Academic tertiary care hospital | Prospective consecutive cohort selected cross sectional study | Osteomyelitis and septic arthritis | Serum PCT using >0.5ng/mL threshold | Positive bacteriological culture from blood, bone aspiration or joint fluid | No competing interests |
| Greeff E, 2012.[93] | South Africa | N=33<br><br>1 month-14.0 years | Children less than 14 years age presenting to hospital with signs and symptoms of osteomyelitis or septic arthritis. No child less than 30 days (personal correspondence) | Hospital | Prospective consecutive cohort selected cross sectional study | Osteomyelitis and septic arthritis | Serum Procalcitonin >0.2 ng/mL threshold data extracted only and serum CRP >10mg/L threshold | Pus at arthrotomy or from bone and microscopy and culture of tissue | No competing interests |
| Jung ST, 2003.[94] | Korea | N=124<br><br>1 month-15.0 years | Children with acute hip pain | University hospital | Retrospective consecutive cohort selected cross sectional cohort study | Septic arthritis | Serum CRP >10mg/L | Positive culture of hip fluid from arthrotomy. | Nil recorded |
| Singhal R, 2011.[91] | UK | N=311<br>30.2%<br>2.4 months-15.1 years | All children and adolescents with acute new onset atraumatic limp or hip pain. | Children's hospital | Retrospective consecutive cohort selected cross sectional study | Septic arthritis | Serum CRP >20mg/L | Positive culture hip aspirate or abundance WBC on microscopy (total 42 patients underwent arthrotomy hip) | Nil recorded |
| Sultan J, 2010.[96] | UK | N=96<br>35.4%<br>1.0-12.0 years | All children and adolescents admitted to hospital with irritable hip.<br>91(94.8%) Transient synovitis, 5 (5.2%) septic arthritis | Teaching hospital | Retrospective consecutive cohort selected cross sectional study | Septic arthritis | Raised serum CRP $\geq$ 20mg/L | Positive culture from hip joint/ blood culture and numerous WBC on hip microscopy | Nil recorded |

COI: conflict of interest; CRP: C-reactive protein; PCT: procalcitonin

#### 3.2.2.2 *Studies including both CRP and PCT and reference test*

Two studies in this review evaluated both CRP and PCT directly with the reference test.[31, 93] Both studies were prospective using a consecutive cohort selected cross-sectional study design.
In the smaller study from South Africa a total 33 children and adolescents with suspected bone and or joint infection were enrolled. This study used cut off of 10mg/L for CRP and 0.2 ng/mL for PCT threshold. [93] The hospital setting was similar to those studies above.
In the larger of these two studies from France a total of 339 patients presenting to a university teaching hospital with presumed osteoarticular infection were included.[31] In this study, a PCT serum measurement of 0.5 ng/mL was chosen as the threshold level. While CRP was also measured, it was also included to identify patients categorized as "suspected infection". For this reason CRP could not be evaluated as a diagnostic marker in this study. Four cases that included using CRP in their reference standard to classify patients as suspected infection were therefore excluded from data analysis. The paired index study flow is shown below.

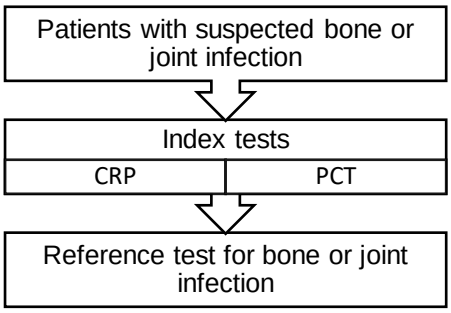


Figure 3. Representation of study flow for both CRP and PCT

#### 3.2.2.3 *Studies including PCT and reference test*

One study evaluated the index test PCT only with the reference test. [95] This study from China included children with acute and non-acute osteomyelitis.[31] This was a prospective trial, involving a total of 187 patients. Patients with acute osteomyelitis were classified on layered bone structure radiologically with or without bone biopsy as the reference test. A cut off level of 3.56 ng/mL for PCT was not predetermined. Inconsistent use of the SI units for the PCT test was noted in this paper. A single index test study flow design for PCT is shown below.

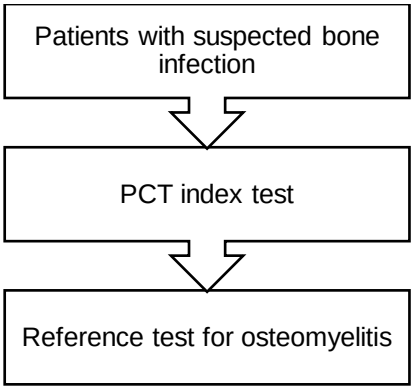


Figure 4. Representation of study flow for PCT

## 3.3 Methodological Quality of Studies

All studies were evaluated individually for bias across the four domains below. With the exception of the flow and timing domain, applicability was also included as part of the quality assessment.

### 3.3.1 Patient selection

In one study the enrolment was unclear.[94] All remaining studies appeared to enrol patients consecutively. Two studies used a case-control design.[91, 95]One study excluded patients after being enrolled. [91]All studies were of low concern regarding to the applicability of patient selection and those targeted by the review question.

### 3.3.2 Index test

Information regarding blinding from test results was provided in only one study and this was considered appropriate [93] Potential bias from blinding was considered unclear in the remaining studies. Three studies did not include a prespecified threshold for the index test being evaluated, potentially overestimating of the test performance outcomes in these cases.[93-95] Low concern with the applicability of the index tests used in the studies was raised.

### 3.3.3 Reference test

While some variation with the reference standard was noted, they were all assessed by the two reviewers to correctly classify the target condition of interest. No study included information as to the reference standard being interpreted without knowledge of the index test result. The target condition defined in the reference question was consistent with that defined by the individual studies and low concerns regarding applicability of the reference test was therefore recorded.

### 3.3.4 Flow and timing

One study reported on the time interval between index test and reference standard and this was noted to be appropriate.[31] The time interval was unclear in the other seven studies. Three studies reported using the same reference standard for all patients therefore minimising risk from verification bias.[46, 90, 93] With the exception of one study,[31] all patients were included in the final analysis.

To help summarise the overall findings from the QUADAS-2 assessment of bias, the following interpretations were first applied to the outcomes as recommended.[88] A study where a “yes” answer was recorded, was judged as low risk of bias, a “no” response was judged as high risk of bias and “unclear” judged as unclear risk of bias. For applicability, concerns were simply rated as either “high”, “low” or “unclear” risk.

If an individual study was judged as “low” on all the domains relating to bias or applicability, then the overall judgement was of low risk of bias or low concern related to applicability for that study. If a study however was judged as high risk or unclear risk in only one or more domains then it was judged as “at risk of bias” or having “concerns regarding applicability”.[88]

The results from critical appraisal of bias and applicability for each study are presented in Tables 3a and 3b respectively and then summarised in table 4. All studies were considered “at risk of bias” using the QUADAS-2 tool. Given the conservative approach for classifying “at risk of bias” using this methodology we chose not to exclude any of the eight studies from further statistical analysis but to acknowledge any potential limitations we identified in our reporting of outcomes.

Table 3a. Critical appraisal for bias

| Citation Fist author, date publication | **Q1** Was a consecutive or random sample of patients enrolled? | **Q2** Was a case-control design avoided ? | **Q3** Did the study avoid inappropriate exclusions? | **Q4** Were the index test results interpreted without knowledge of the results of the reference standard? | **Q5** If a threshold was used, was it pre-specified? | **Q6** Is the reference standard likely to correctly classify the target condition? | **Q7** Were reference standard results interpreted without knowledge of results of the index test? | **Q8** Was there an appropriate interval between index test and reference standard? | **Q9** Did all patients receive the same reference standard? | **Q10** Were all patients included in the analysis ? |
|---|---|---|---|---|---|---|---|---|---|---|
| Caird MS, 2006. | Y | Y | Y | U | Y | Y | U | U | Y | Y |
| Cui C, 2017. | Y | N | Y | U | N | Y | U | U | N | Y |
| Eich GF, 1999. | Y | Y | Y | U | Y | Y | U | U | N | Y |
| Faesch S, 2009. | Y | Y | Y | U | Y | Y | U | Y | N | N |
| Greeff E, 2012. | Y | Y | Y | Y | N | Y | U | U | Y | Y |
| Jung ST, 2003. | U | Y | Y | U | N | Y | U | U | N | Y |
| Singhal R, 2011. | Y | Y | U | U | Y | Y | U | U | N | Y |
| Sultan J, 2010. | Y | Y | Y | U | Y | Y | U | U | Y | Y |

Y=yes, N=no, U=unclear

Table 3b. Critical appraisal applicability

| Citation Frist author, date publication | Concerns regarding severity of target condition | Concerns regarding demographic features | Concerns regarding comorbid conditions | Concerns regarding setting of study | Concerns regarding variation in index test technology , execution or interpretation of | Concerns regarding differences in target condition specified and target condition in study |
|---|---|---|---|---|---|---|
| Caird MS, 2006 | Low | Low | Low | Low | Low | Low |
| Cui C, 2017 | Low | Low | Low | Low | Low | Low |
| Eich GF, 1999 | Low | Low | Low | Low | Low | Low |
| Faesch S, 2009 | Low | Low | Low | Low | Low | Low |
| Greef E, 2012 | Low | Low | Low | Low | Low | Low |
| Jung ST, 2003 | Low | Low | Low | Low | Low | Low |
| Singhal R, 2011 | Low | Low | Low | Low | Low | Low |
| Sultan J 2010 | Low | Low | Low | Low | Low | Low |

Low=low concerns regarding applicability

Table 4. Summary of risk of bias and applicability for QADAS-2 domains

| | | Risk of Bias | | | | Applicability concerns | |
|---|---|---|---|---|---|---|---|
| First author and year | Patient selection | Index test | Reference standard | Flow and timing | Patient selection | Index test | Reference standard |
| Caird MS, 2006 | | | | | | | |
| Cui C, 2017 | | | | | | | |
| Eich GF, 1999 | | | | | | | |
| Faesch S, 2009 | | | | | | | |
| Greef E, 2012 | | | | | | | |
| Jung ST, 2003 | | | | | | | |
| Singhal R, 2011 | | | | | | | |
| Sultan J 2010 | | | | | | | |

Low risk of bias

At risk of bias

## 3.4 Review Findings

### 3.4.1 Diagnostic test measure outcomes from individual studies

The sensitivity and specificity of the index tests for each individual study are included in Table 5. Variation in diagnostic accuracy was observed between the two different index tests and within index tests across different thresholds. For CRP, the highest sensitivity and specificity recorded from an individual study was 100.0 % and 85.9% respectively at a threshold of 20mg/L.[92] In comparison for PCT, the individual study with the highest sensitivity and specificity recorded was 91.7% and 81.0%, at a threshold of 0.2ng/mL.[93] The other calculated measures of diagnostic accuracy are also included in Table 5 for comparison between index tests. These include positive and negative LR, OR and positive and negative predictive values.

Table 5. Summary of diagnostic test measures from included studies.

| Frist author, year publication | Index test Cut off | TP | FN | FP | TN | Sensitivity (95% CI) | Specificity (95%CI) | LR+ (95%CI) | LR- (95%CI) | DOR (95%CI) | PPV (95%CI) | NPV (95%CI) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Caird MS, 2006. | CRP>20 mg/L | 29 | 5 | 4 | 10 | 85.3 (68.9-95.0) | 71.4 (41.9-91.6) | 3.0 (1.3-6.9) | 0.2 (0.0-0.5) | 14.5 (3.4-62.5) | 87.9 (71.8-96.6) | 66.7 (38.4-88.2) |
| Cui C, 2017. | PCT ≥ 3.56 ng/mL | 71 | 21 | 53 | 122 | 88.3 (68.9-95.0) | 71.4 (41.9-91.6) | 2.9 (1.3-6.9) | 0.2 (.08-0.5) | 14.5 (3.4-62.5) | 87.9 (71.8-96.6) | 66.7 (38.4-88.2) |
| Eich GF, 1999. | CRP≥ 20mg/L | 8 | 0 | 9 | 55 | 100 (63.1-100.0) | 85.9 (75.0-93.4) | 6.5 (3.5-11.9) | 0.0 (0.0-0.9) | 99.3 (5.3-1867) | 47.1 (23.0-72.2) | 100 (93.5-100.0) |
| Faesch S, 2009. | PCT >0.5ng/mL | 2 | 6 | 9 | 282 | 25.0 (3.19-65.1) | 96.9 (94.2-98.6) | 8.08 (2.0-31.5) | 0.77 (0.5-1.7) | 10.5 (0-53.1) | 18.2 (2.3-51.8) | 97.7 (95.5-99.2) |
| Greeff E, 2012. | PCT >0.2 ng/mL | 11 | 1 | 4 | 17 | 91.7 (61.5-99.8) | 81.0 (58.1-94.6) | 4.8 (1.9-11.8) | 0.1 (.01-0.7) | 46.8 (5.51-475) | 73.3 (44.9-92.2) | 94.4 (72.7-99.9) |
| | CRP >10mg/L | 12 | 0 | 14 | 5 | 100 (73.5-100) | 26.3 (9.2-51.2) | 1.3 (0.9-1.8) | 0.1 (0.0-2.3) | 9.5 (0.5-189) | 46.2 (26.6-66.6) | 100 (47.8-100) |
| Jung ST, 2003. | CRP >10mg/L | 24 | 3 | 7 | 90 | 88.5 (69.8-97.6) | 92.8 (85.7-97.0) | 12.3 (5.9-25.4) | 0.1 (0.0-0.3) | 98.6 (24.6-387) | 76.7 (57.7-90.1) | 96.8 (90.9-99.3) |
| Singhal R, 2011. | CRP >20mg/L | 25 | 4 | 20 | 262 | 86.2 (68.3-96.1) | 92.9 (89.3-95.6) | 12.2 (7.8-19.0) | 0.1 (0.0-0.4) | 81.9 (26.9-247) | 55.6 (40.0-70.4) | 98.5- (96.2-99.6) |
| Sultan J, 2010. | CRP ≥ 20mg/L | 3 | 2 | 9 | 82 | 60.0 (14.7-94.7) | 90.1 (82.1-95.4) | 6.0 (2.3-15.6) | 0.4 (0.2-1.3) | 13.7 (2.4-78.2) | 25.0 (5.5-57.2) | 97.6 (91.7-99.7) |

CRP=C-reactive protein, PCT=procalcitonin, TP=true positive, FP=false positive, TN=true negative, FN=false negative, LR+=positive likelihood ratio, LR-= negative likelihood ratio, DOR=diagnostic odds ratio, PPV= positive predictive value, NPP=negative predictive value.

### 3.4.2 Meta-analysis of CRP using index tests with common threshold

Using a bivariate model, meta-analysis was performed for CRP studies with a common threshold of 20mg/L. The summary point for pooled sensitivity with 95% CI was 0.86 (0.68-0.96) and for specificity 0.90 (0.83-0.94) (see Figure 5). This corresponds to a positive LR of 8.6 (4.75-15.5) and negative LR of 0.15 (0.09-0.25). In all four studies [90-92, 97] however the target condition was septic arthritis. No cases of osteomyelitis were included at this threshold level.

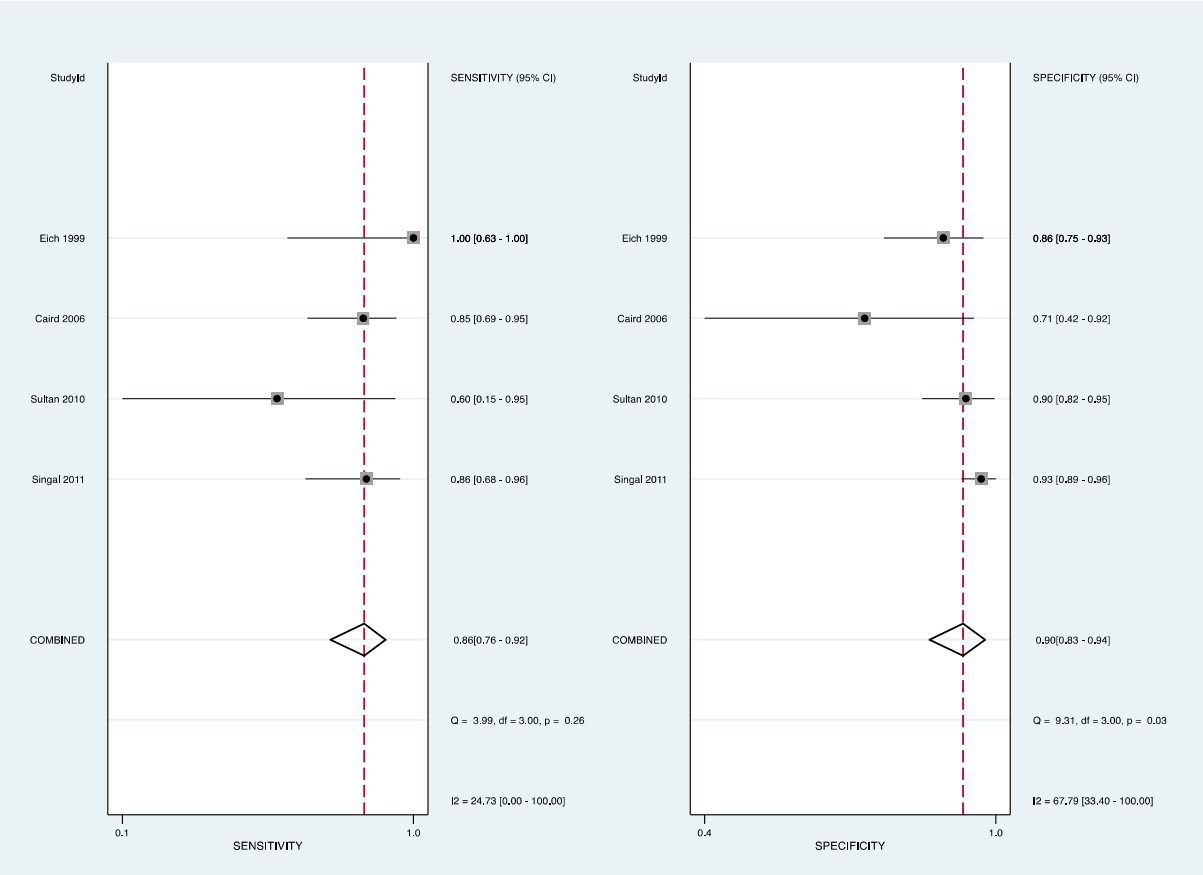


Figure 5. Forest plots of sensitivity and specificity from studies using a CRP threshold of 20mg/L in adolescents and children with septic arthritis.

### 3.4.3 HSROC curve for CRP combining different positivity thresholds

Hierarchical logistic regression models were used to display estimates of HSROC parameters and simple summary measures. The HSROC curve plot best describes how sensitivity and specificity trade off at different positivity thresholds values. This was achieved by using only one pair of sensitivity and specificity values from each of the four CRP studies at threshold of 20mg/L together with one pair each from the two additional studies using CRP threshold of 10 mg/L (see Figure 6).

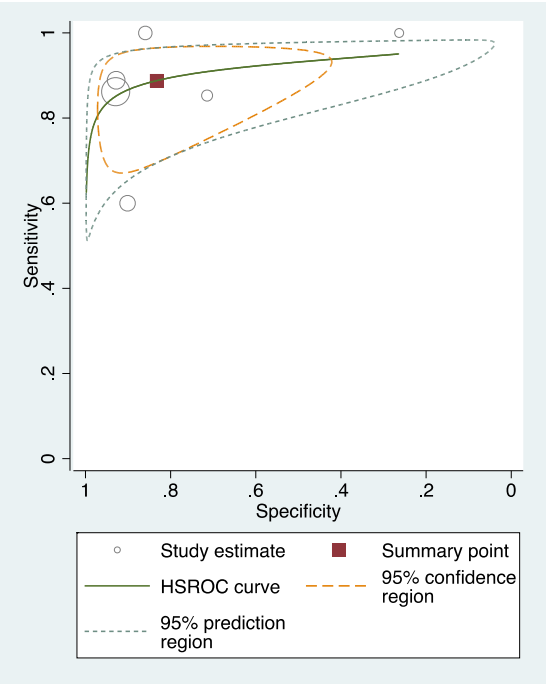


Figure 6. Hierarchical Summary Receiver Operating Characteristic curve for CRP.

### 3.4.4 Other CRP diagnostic accuracy measures using combined positivity thresholds

Other summary data measures of diagnostic accuracy for CRP included a diagnostic OR (DOR) of 39.4 (14.85-104.9), positive LR (LR+) 5.3 (2.3-11.9) and negative LR (LR-) 0.1 (0.07-0.2). A pre-test versus post-test probability curve for osteoarticular infection using CRP is shown in Figure 7.

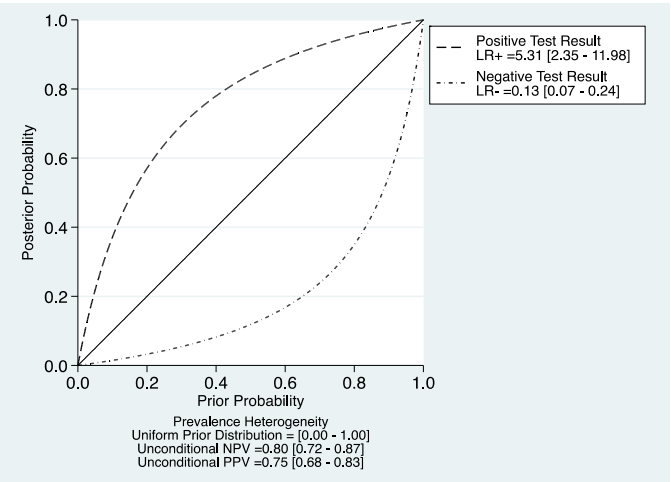


Figure 7. Line graph of post-test probabilities versus prior probability of osteoarticular infection based on positive or negative CRP test result.

### 3.4.5 Diagnostic accuracy outcomes for PCT

Since only three studies were included for evaluating the diagnostic test accuracy of PCT for osteoarticular infection in children and adolescents, it was not possible to a perform meta-analysis in STATA to construct either a Forest plot or HSROC curve for PCT. Hence it was also not possible to

directly compare the diagnostic accuracy of measurement of CRP with PCT from pooled data using these techniques.

In the only paediatric osteoarticular infection study directly comparing the sensitivity and specificity of PCT at a threshold of 0.2ng/mL with CRP at a threshold of 10mg/mL[93], PCT was shown to be less sensitive (91.7% [61.5-99.8] versus 100% [73.5-100.0]) for CPR but more specific (81% for PCT [58.1-94.6]) versus 26.3% (9.2-51.2) for CRP.[93] From the calculated DOR as a global measure of diagnostic accuracy in this study, PCT outperformed CRP, 46.8 (5.51-475.0) to 9.5 (0.5-189.0) respectively.[93]

In the second PCT study that also included both septic arthritis and osteomyelitis cases,[31] a test threshold of 0.5ng/mL resulted in a sensitivity of 25% (3.2-65.1) and increased specificity of 96.9% (94.2-98.6).[31] The DOR from this study however, was lower at 10.5 (0.0-53.1).

In the final PCT study using the highest reported threshold for PCT at 3.56 ng/mL in children with osteomyelitis only, the specificity was still high at 88.3% (68.9-95.0) and sensitivity 71.4% (41.9-91.6) with a DOR of 14.5 (3.4-62.5).[95]

### 3.4.6 Heterogeneity

Visual inspection of the Forest plots for CRP (see Figure 2) demonstrated overlapping CIs for both sensitivity and specificity and while this does not exclude heterogeneity, it does infer this may not be excessive. The distribution of individual studies from the HSROC curve, together with the number of studies lying outside the 95% prediction area plus a prediction region that is larger than the 95% CI region, all confirm a degree of heterogeneity exists (see Figure 6).

### 3.4.7 Publication bias

Deeks' funnel plot for examining publication bias for meta-analysis of diagnostic test accuracy studies was performed and did not demonstrate statistically significant (p=0.11) asymmetry (see Figure 8).

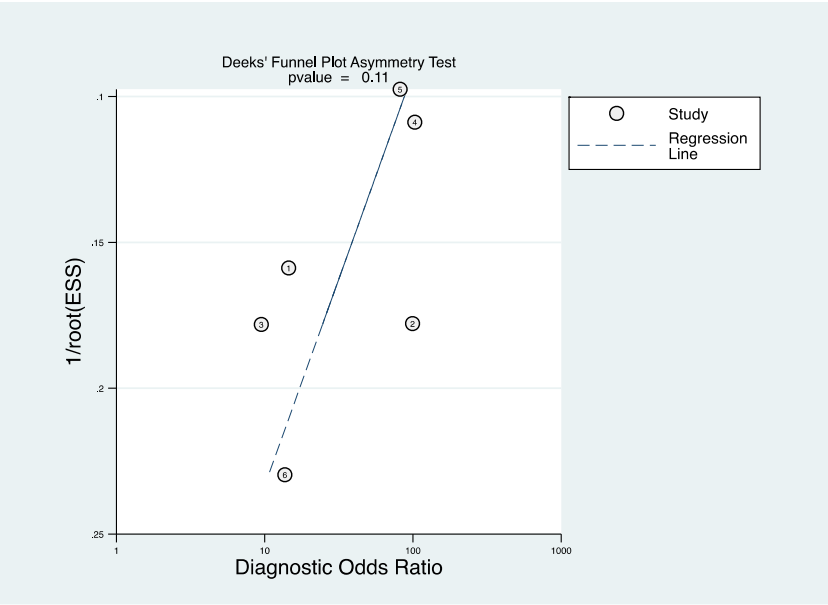


Figure 8. Deeks' Funnel Plot for publication bias for CRP studies

## 3.5 Conclusion

This systematic review has identified eight studies from a total of 3,086 that include the measurement of serum CRP and or PCT in children and adolescents presenting to hospital with suspected bone and joint infection. The results from these analyses confirm the summary point pooled estimates of both sensitivity and specificity when using a CRP threshold of 20mg/L as a diagnostic test. We have also provided additional statistical precision of test accuracy when combining test thresholds for both 10mg/L and 20mg/L for CRP using HSROC modelling. This included a diagnostic OR (DOR) of 39.4 (14.85-104.9), positive LR (LR+) 5.3 (2.3-11.9) and negative LR (LR-) 0.1 (0.07-0.2). At the time of performing this review, there was insufficient data to perform a meta-analysis for PCT. Formal critical appraisal highlighted the potential for "at risk of bias" in those studies included in statistical analysis. This review presents new quantitative data on the diagnostic accuracy of CRP for both bone and joint infection in children and adolescents.

# 4 Discussion and conclusion

## 4.1 Introduction

This chapter explores and explains the major findings from this review in the context of the existing body of knowledge in this area. It also considers the limitations of this review and discusses the implications of the results for future clinical practice and research.

## 4.2 Research objective

The objective of this systematic review was to identify, evaluate, extract and synthesise the best available scientific evidence for the diagnostic test accuracy of serum measurement of PCT compared to CRP in children and adolescents aged between one month and 18 years of age, admitted to hospital with suspected osteoarticular infection.

The findings from this review support the use of CRP as a preferred biomarker over PCT in this setting. Based on the currently available data, there is insufficient evidence to recommend that serum measurement of PCT can be used as either an additional or replacement test for the routine evaluation of suspected bone and joint infection in children and adolescents.

## 4.3 Systematic review findings in the context of existing results

### 4.3.1 The diagnostic test performance of CRP using a common test threshold of 20mg/L in children and adolescents with suspected septic arthritis

This review has provided the first summary operating point estimates for CRP exclusively in children and adolescents with suspected septic arthritis. Using a bivariate model, the pooled sensitivity with 95%CI for CRP at a common diagnostic threshold of 20mg/L was shown to be 0.86 (0.68-0.96) and specificity 0.9 (0.83-0.94). From the systematic literature search performed in this review, no other meta-analysis was identified when confined to the paediatric age group with this condition.

In a recent meta-analysis where researchers did combine both adult and paediatric patients with septic arthritis, to evaluate both PCT and CRP as diagnostic markers, the pooled sensitivity and specificity reported for CRP were both significantly lower at 0.45 (0.35-0.55) and 0.07 (0.02-0.25) respectively. [12] The analysis for the CRP index test included five adult and one paediatric study only and included CRP thresholds ranging from 18-50 mg/L. Estimating sensitivity and specificity by pooling studies which mix thresholds however will produce an estimate that relates to some notional unspecified average of thresholds and this may be clinically unhelpful.[98]

The positive and negative LRs provided in this systematic review were calculated from the summary point estimates for both sensitivity and specificity using the common 20mg/L threshold. Likelihood ratios provide health care practitioners with more clinically meaningful information at an individual

patient level. When evaluating the post-test probability for joint infection in the paediatric setting alone, the positive LR of 8.6 as reported indicates that a patient with a positive test result has more than 8.6 times the odds of having septic arthritis. The negative LR of 0.15 indicates a patient with a negative test result has 15 times less the odds of having the disease.

When comparing the LRs calculated from this review in children and adolescents with the pooled adult and paediatric data combined for CRP for septic arthritis only, the LRs are surprisingly inconsistent. The combined study resulted in an LR+ of 0.4 and LR- of 6.79.[12] In practical terms an LR + less than 1.0 means that a positive test result for CRP is less likely to occur in patients with septic arthritis than without the infection while an LR- greater than 1.0 means that a negative CRP test is more likely to occur in people with the septic arthritis than without. Clearly with such poor test accuracy performance, measuring serum CRP in this setting would offer no advantage to either patient or clinician.

The cause for such a wide discrepancy in the diagnostic accuracy performance is likely to be multi-factorial. While the sensitivity and specificity of a diagnostic test varies as the positivity thresholds change, other important factors may contribute to the variation observed. [99] As diagnostic test accuracy outcome measures are not fixed properties of the test, variables including patient groups, spectrum of disease, clinical setting and test interpreters can significantly affect sensitivity and specificity outcomes.[100]

Assumptions that children and adults can be combined in a meta-analysis to determine summary point estimates of sensitivity and specificity may be inappropriate from a biological plausibility perspective. Evidence from another earlier published systematic review evaluating the diagnostic accuracy of CRP in adults with septic arthritis clearly demonstrated consistently lower specificity values when compared to the paediatric literature.[101] Differences in the disease presentation, pathogens responsible for infection, pathophysiology, immune response to infection, and kinetics of CRP include just some of the reasons why combining these two populations into a single summary measurement could be problematic. The published systematic review for septic arthritis did not perform subgroup analysis for CRP by population type to confirm this as a significant source of heterogeneity.[12] The methodological concerns raised with this systematic review may have also contributed towards the variation in test accuracy outcomes observed.[84]

In summary the diagnostic accuracy of CRP was observed to be higher in the paediatric setting for acute septic arthritis in this review compared with the published data that combined both adults and children. The size of this effect we observed from this review and meta-analysis appears to be significant without any overlap in the 95% CIs and while “risk of bias” existed in the quality of the evidence used, there were low concerns regarding the applicability of this evidence.

### 4.3.2 The diagnostic accuracy performance of CRP across positive thresholds in children and adolescents with either bone and/or joint infection

To determine the summary test performance of serum measurement of CRP across all osteoarticular infections in children and adolescents it was necessary to combine the data from different test thresholds and include both bone and joint infections as the target condition. Using a hierarchical logistic regression model analysis it was possible to generate an HSROC curve for CRP (to show the trade-off of sensitivity and specificity), estimate global measures of diagnostic accuracy for CRP such as DOR and provide overall summary estimates of LRs to improve clinical interpretation.

From this meta-analysis the DOR for the overall test performance of CRP in children and adolescents with osteoarticular infection was 39.4 (14.8-104.9). When comparing these data to the only other systematic review and meta-analysis using combined adult and children and including both bone and joint infection, the DOR for CRP was reported to be lower at 3.5 (1.3-9.7).[11] While a wide confidence interval was observed in the paediatric only data, there was no overlap in CIs between these two cohorts and the greater than ten-fold difference in the global test accuracy of CRP was noted.

When evaluating the likelihood ratios from this review for overall test performance of CRP for both bone and joint infection, the LR+ was 5.3 (2.3-11.9) and LR- 0.1 (0.07-0.2). This compares favorably with the published finding from the combined adult and paediatric CRP data with a reported LR+ of 1.39 (1.17-1.65) and LR- of 0.4 (0.12-1.36).[11]

Again, differences in the diagnostic test outcome measures between the population cohorts were observed. The same rationale for the differences observed in sensitivity and specificity of CRP between paediatric only versus adult and paediatric data for septic arthritis as the target condition also applies to the variations observed using the DOR and LRs when combining both bone and joint infections together as the target condition.

The results from this systematic review and meta-analysis suggest measurement of CRP may perform better when suspecting osteoarticular infections in children and adolescent patients compared with adults. Recommendations for the use of CRP in children and adolescents based on outcomes from combining adult and paediatric data may therefore be inappropriate and misleading. Future evaluations of test accuracy outcome measures for CRP in osteoarticular infections should always include separate analysis of children and adolescents from adults.

In summary by combining only paediatric data for CRP, we observed an improvement in the diagnostic test accuracy for this biomarker in both bone and joint infection in children and adolescents compared to the test performance outcomes when combining all age groups. This effect was most obvious with the LR test outcomes.

### 4.3.3 The diagnostic test accuracy performance of PCT in children and adolescents in bone and joint infections.

The diagnostic test outcome measures for PCT in this setting are based on a limited number of studies. This review performed a broader search strategy than both of the previous published systematic reviews combining adults and children.[11, 12] While identifying more PCT diagnostic test accuracy studies using the search strategy described in this review, in total only three paediatric studies satisfied all the inclusion and methodological criteria for subsequent evaluation and analysis. [31, 93 95] There was only one study that was included in both the previously published systematic reviews and this review. [31]

A comparative evaluation of the test performance of PCT in this setting was the primary objective of this review. A meta-analysis using the PCT studies however was not feasible due to the low number of studies. As per protocol, this was replaced by narrative summary of the diagnostic test outcomes for PCT.

From the summary results (Table 5), variability across all key test performance outcomes for PCT, was found. The range observed for sensitivity was wide, (25.0% to 91.7%), but narrower for specificity (71.4% to 96.9%). The positive LRs were also widely spread from 2.9 to 8.0 with LR- tighter in range (0.1 to 0.2). Finally, the DOR variability observed across all three PCT studies was from 10.5 to 46.8. The most important explanation for this observed variability across studies is the different test threshold levels employed for PCT. The threshold levels were also widespread including values of 0.2ng/L, 0.5ng/L and 3.56ng/L. The study reporting the highest sensitivity (91.7%), specificity (81.0%) and DOR (46.8) occurred at the lower test cut off for PCT at 0.2ng/L.[93]

This review excluded a paediatric study that was included in both of the other systematic reviews comparing PCT and CRP in children and adults with osteoarticular infections. [30] The reason for exclusion from this review was that neonates and adolescents up to 19-years of age were part of the analysis and this was outside the age range in the published protocol. With this limitation in mind, the sensitivity and specificity reported from this study using a cut off of 0.50ng/L for PCT was 43.5% and 100% respectively and most consistent with the included study in this review also using the same cut-off.[30, 31] Again this demonstrates the importance of the threshold effect on the diagnostic test accuracy outcome measures.

Interestingly, published meta-analyses performed to evaluate the diagnostic test accuracy of PCT in other disease settings such as sepsis more often differentiate adults from the paediatric age and within paediatrics further separate neonates from other cohorts.[102 103, 104] Clearly other investigators recognise the potential for age to be a significant potential confounder in the assessment of PCT test outcomes. This further supports the premise that using combined age population data for inference across all age cohorts may has limitations.

In summary there appears to be variability across the test performance measurements for PCT in the paediatric setting when used as a diagnostic test for suspecting osteoarticular infection.

### 4.3.4 Comparing the diagnostic test accuracy performance of PCT and CRP

The test performance results where diagnostic index tests are compared indirectly between patients may differ to those in which all patients receive all tests and the reference test. The direct comparative design however is the preferred methodology to help guide clinicians in their test selection and this approach should be routinely undertaken whenever possible.[105]

In the only direct comparison study included in this review, PCT (at a threshold of 0.2 ng/mL) was reported to be more specific, (0.81 [0.58-0.95]) than CRP at a threshold of 10 mg/L, (0.26 [0.92-0.51] for children presenting to hospital with suspected acute septic arthritis and osteomyelitis.[93] Sensitivity however was higher for CRP at 1.00 (0.73-1.00) versus 0.92 (0.61-0.99) for PCT.[93] The small number of enrolled cases (n=33) was a limitation of this study.[93] In the only other direct comparative published study identified in children and adolescents with acute osteoarticular infection, PCT again demonstrated higher specificity (100%) compared to CRP (19%) but again with lower sensitivity 43.4 versus 78.2 respectively.[30] This study was excluded as it included both neonates and adults up to 19 years of age.

In both of the combined adult and paediatric meta-analyses performed, the specificity of PCT was also noted to be higher than for CRP, but the data using direct comparisons were incomplete for both these studies.[11, 12] In the meta-analysis examining septic arthritis cases only, a total of six of the 10 studies analysed included both index tests with the reference test in the same patient.[12] In the second review of both bone and joint infections, only three of the included seven studies incorporated both index tests and reference tests for each patient.[11] Even with this constraint, there remains consistent evidence across all studies to suggest higher a specificity can be expected when measuring serum PCT compared to serum CRP as a diagnostic test for children and adolescents with presumed osteoarticular infection.

The published combined adult and paediatric meta-analyses for PCT and CRP also performed subgroup analysis after removal of one paediatric study in the first study and two in the second study.[11, 12] In both cases the re-analyses did not appear to significantly change any of the diagnostic accuracy parameters examined given significant overlap was still observed with the wide confidence intervals.

In summary, given the limited direct and indirect comparative data between CRP and PCT from this systematic review and the wide variability in results observed with serum measurement of PCT, it was not possible to provide an accurate assessment of the diagnostic test performance of PCT in children and adolescents presenting to hospital with osteoarticular infection.

## 4.4 Unexpected findings

The most unexpected finding from this review was the number of diagnostic test accuracy studies that did not directly compare the two index tests (CRP and PCT) to the reference test in the same patient. This was also observed in many of the studies not meeting the inclusion criteria for this review. This finding however is not limited to the paediatric setting and is shown to be consistent across the published diagnostic test accuracy literature.[106] Improved guidance on both methodology and reporting will assist to improve the quality of these studies.

## 4.5 Strengths of this review

The value of any systematic review depends upon the quality quantity, homogeneity of the included studies but equally important the rigor of the methodology adopted in both of the systematic review and included studies.[73] This review followed the preferred reporting items for systematic reviews and meta-analysis of diagnostic test accuracy studies using the PRISMA-DTA, 27 step check list first published in 2009.[107] Significantly this review was conducted without any deviation from the published protocol.[1]

The review question was of high clinical importance. Given the uncertainty surrounding the diagnostic accuracy of PCT in children and adolescents with suspected osteoarticular infection and the financial burden to our health system in performing a more expensive alternative test, more evidence was needed to guide clinicians. Importantly the inclusion criteria in this review were carefully defined to focus on the paediatric setting not specifically addressed in previous systematic reviews. [11, 12]

While evidence is emerging to support more efficient search strategies for systematic reviews of diagnostic test accuracy by limiting searches to data bases such as MEDLINE (PubMed) and Embase (Ovid) a more comprehensive search was used for this review as described in our methods.[108] To allow replication of our methodology, full search strategies are included in Appendix 1. In the previously published systematic reviews no more than three data bases were searched for each review, resulting in 68 citations in one review and 205 in another before undertaking screening.[11, 12] In comparison this review included four data bases, unpublished and grey literature and started with over 900 citations after duplicates were removed. This review also avoided the use of methodological filters shown to miss relevant diagnostic test accuracy studies.[109] The overall aim of this more extensive search was to identify as many relevant paediatric studies as feasible and to help mitigate publication bias.

Another strength of this review was the use of two independent reviewers to critically appraise each study for methodological quality by applying the QUADAS-2 critical appraisal tool for diagnostic test accuracy.[88] This tool aims to improve transparency in the assessment of bias.

It was not always clear under what circumstances the variations seen with the different methodological designs may change the estimates of diagnostic accuracy performance. This uncertainty particularly applies to diagnostic case-control studies, where it has been shown that bias may lead to a two or three fold higher estimate of diagnostic accuracy compared to cohort selection studies.[110] In this review, seven of the eight included studies used a cohort study design further minimising this risk of bias.

## 4.6 Limitations of this review

Several limitations were identified with this systematic review and meta-analysis. Only studies in English language were included in this review. A systematic review of English language restriction on systematic review based meta-analyses however found overall no evidence of bias from language restriction in conventional medicine.[111, 112]

While one of the strengths of this review was the inclusion of studies using a consecutive cohort-cross sectional design, the critical appraisal identified some potential risks of bias. The most common and consistent potential source of bias identified included interpretation related bias. Insufficient information was a systematic and significant finding across many of the studies and contributed to having to document an unclear answer to many of the appraisal questions. Less common sources of bias included failure to pre-specify the test threshold in half of the studies and a degree of inconsistency from not all patients not receiving the same reference test. The extent from the impact of this on the overall statistical findings however is considered to be low given that the index tests are usually performed and reported independently and that the reference test includes objective criteria measures.

In the absence of any sample size calculations of individual studies it is difficult to comment on whether any of the included studies were powered adequately.

It was not feasible to perform subgroup analyses to statistically identify differences in the diagnostic test performance of either CRP or PCT based on the primary site of infection, either bone or joint. While this would be interestingly academically from a clinical perspective it is less important given differentiating the two conditions is not always possible and both conditions can occur simultaneously.

As with the other published meta-analyses on this topic, the meta-analyses in this review were also performed using a small number of studies (n= 4-6) with sample sizes ranging from 33 to 339. It is well recognised that small numbers of both may contribute to sampling and statistical error resulting in both low precision and low confidence in the estimated effects observed.[113]

While heterogeneity is anticipated in diagnostic test accuracy reviews, using methods such as Cochrane Q or $I^2$ statistics may not be informative as they do not account for heterogeneity due to threshold effects.[99] Instead this review relied on graphical visualisation from the CIs with Forest plots

and prediction regions from the HSROC curve. Using this approach, it was not possible to accurately quantify the degree of heterogeneity where it exists.

In this review it was not possible to evaluate the test accuracy performance of the PCT in children and adolescents using statistical methods. A commentary and overview were therefore provided on the three studies included. This summary is subjective with the same limitations of a narrative review.

The risk from publication bias in this meta-analysis of diagnostic test accuracy was evaluated as recommended using Deeks' funnel plot for asymmetry.[114] While this test was not significant, it is underpowered when using a small number of studies (as was the case here) and may have been falsely reassuring.[115]

The GRADE guidelines use a structured framework for evaluating certainty across a body of evidence and provide readers with a level of confidence in the effect estimates observed.[116] This review included a combination of both meta-analysis and narrative summary to evaluate the test accuracy performance which proved challenging to incorporate into the current GRADE appraisal format. It has been acknowledged however that further development is still required before it can be widely applied to all systematic reviews of test accuracy.[117]

## 4.7 Implications for clinical practice

The results from this study provide clinicians with increased statistical precision regarding the diagnostic test performance of CRP in children and adolescents with suspected acute bone and joint infection. It is recommended that CRP remain the preferred biomarker to help define the post-test probability of suspected acute osteoarticular infection (Grade A, JBI grade of recommendation: a strong recommendation for a certain health management strategy where (1) it is clear that desirable effects outweigh undesirable effects of the strategy; (2) where there is evidence of adequate quality supporting its use; (3) there is a benefit or on impact on resource use, and (4) values, preferences and the patient experience have been taken into account.).

Due to the limited evidence available, no recommendation can be made regarding the use of PCT as either and adjunct or replacement test in the paediatric setting; further research is required (Grade A, JBI Grade of Recommendation). Clinicians should be cautious when requesting and or recommending PCT routinely in isolation for children and adolescents with suspected acute bone and joint infection, until it is possible to make a better informed decision regarding the diagnostic accuracy of this test in this population.

While the results from predominately adult-based systematic reviews may be illuminating for clinicians working in paediatrics, this information alone must not form the basis on which decisions should be made regarding children and adolescents. Based on the findings and recommendations from this review it is suggested that local, national and international clinical guidelines continue to recommend

measurement of serum CRP in their diagnostic evaluation of bone and joint infection in children and clinicians should await further evidence before consideration is given to the role of PCT. These recommendations are consistent with the core aims of the Choosing Wisely® campaign of promoting only those investigations that are evidence based.

The publication of the systematic review protocol and the systematic review should help provide clinicians and researchers with access to new scientific information specifically for the paediatric setting. Hopefully this will generate ongoing clinical debate and further recommendations regarding the existing use and/or future use of PCT in children and adolescents presenting to hospital with bone and joint infection in their own health care setting.

## 4.8 Implications for research

The research process for this systematic review and the results obtained from the meta-analysis have identified several important gaps in the knowledge base that may have important implications for the direction of future research in this area. These are discussed below.

### 4.8.1 The next step

The most expedient and least resource demanding response to address the research questions left unanswered would be to first extend the original age inclusion criteria in the published protocol so that 19-year-old patients and neonates would be eligible. This step would increase the number of included studies potentially making meta-analyses for PCT studies more feasible, but with this approach comes additional limitations (see below section 4.8.2). The added advantage of this approach is that it may also allow subgroup analysis by different age groups. If this step was unsuccessful however, it would be reasonable to wait for more publications before updating this systematic review. Although this approach may capture newer studies it would also delay any further formal recommendations and may not improve the methodologically quality of included test diagnostic accuracy studies. Improvements in the quality of studies would require a direct comparative consecutive cohort cross sectional study design including both index tests simultaneously with a reference test. Ideally this would be prospective and across a defined age group. A sample size calculation would be essential to ensure adequate power to detect the anticipated effect was reached. In reality all of these approaches may be required to make the best evidence-based recommendations regarding the diagnostic accuracy of PCT in children and adolescents with acute osteoarticular disease. This review provides the foundation on which to build these additional layers.

### 4.8.2 PRISMA-C, PRISMA-PC and PRISMA-DTA

Given the large degree of physiological, pharmacological, and psychological differences among different age groups (e.g. neonates, children, adolescents, adults, elderly people), it is important that research questions be age-specific and that evidence be reported in such a way that the resulting recommendations are appropriate for each age group.[118] Furthermore child health systematic reviews

and meta-analyses exhibit incomplete reporting of child-centric topics. Reports and protocols with a mixed children/adult population are more prone to incomplete reporting than child-only populations.[118] To overcome these limitations and address some of the more child specific features, PRISMA-C for children and the PRISMA-PC protocol for children are being developed.[119] These will be welcomed for paediatric researchers but careful consideration and clear guidance will be needed to harmonise the paediatric specific methodology with the newly updated PRISMA-DTA guidelines released in August 2020.[120] As the PRISMA framework forms the preferred methodology for systematic reviews and meta-analyses, there is a greater sense of urgency for completion of the paediatric guidelines and to provide the necessary direction and support for researchers working in the paediatric diagnostic test accuracy field.

### 4.8.3 Sample sizes and power analysis

None of the individual diagnostic accuracy studies included in this review performed a sample size calculation. This is an important consideration before conducting a diagnostic study to justify the results, conclusions and recommendations. This helps ensure that the results obtained from the subsequent analysis will provide the test with a desired minimum value for both sensitivity and specificity together with a sufficient level of power to detect and effect and low level of type 1 error.[121] In a published study conducted to evaluate the impact of study size on meta-analyses, the authors found that when at least two adequately powered studies were included, the underpowered studies often contributed little additional information. However they found that underpowered studies made up the entirety of the evidence in most Cochrane reviews.[122] Researches wishing to perform rapid reviews however may benefit from this knowledge when seeking to reduce costs and time. Sample size assessment should therefore be considered in future methodological critical appraisal tools of diagnostic test accuracy studies.

One of the most powerful and persuasive reasons for performing meta-analyses is the desire to increase the statistical power to detect an effect when one exists. In meta-analyses which lack power, precision in estimation is also lacking, and therefore the summary effect is imprecisely estimated.[123] If meta-analyses are performed too early, before enough studies are available, there is potential risk from erroneous summary statistics and incorrect conclusions. More dialogue with statisticians would be welcomed in exploring the application of power analysis for systematic reviews with meta-analysis in diagnostic test accuracy studies into the future.

### 4.8.4 Hypothesis testing and study design

Most randomised controlled trials pre-specify a test hypothesis that the intervention under evaluation improves outcomes which is statistically tested against a null hypothesis that has no effect. The sample size is calculated and a test statistic with corresponding p-value determined.[124] The situation for individual diagnostic test accuracy studies however is very different. According to one study only, 12% reported any statistical test of a hypothesis related to the study objectives and no more than 11%

reported a sample size justification.[125] Having a clear and pre-specified hypothesis: 1) forces researchers to express minimally acceptable criteria for accuracy values (e.g. sensitivity and specificity) that would make the test clinically fit for purpose before initiating the study, 2) enables informed judgement of the appropriateness of the study design, sample size, statistical analysis and conclusions, and 3) may minimise the risk from over-interpretation of findings.[124] The same hypothesis driven approach should also be applied and documented in protocols for meta-analyses of diagnostic test accuracy.

Critical appraisal of the methodological quality of diagnostic test accuracy studies includes assessment and evaluation of the study design.[88] Clear and consistent distinction of the different study designs is important for diagnostic test accuracy studies to ascertain the risk of bias.[126] Unfortunately test accuracy studies are often unclearly labelled in the medical literature or not defined at all, making assessment more difficult and leaving the reader to make a conclusion.[126] Fortunately an algorithm for the classification of test accuracy studies in systematic reviews has been proposed to improve consistent classification of study designs and this could be incorporated into the handbooks and guidelines for systematic review researchers in this field.[126] Making explicit statements regarding the study design in individual diagnostic test accuracy studies as recommended in the Standards for Reporting of Diagnostic Accuracy Studies (STARD) should be a mandatory requirement.[127] Efforts are currently underway to extend the QUADAS-2 tool for assessment of methodological quality of diagnostic tests with incorporation of the new comparative QUADAS-C tool.[128]

Where new diagnostic tests are being evaluated as either a replacement or additional test, direct comparison of the new index test with the existing best practice or routine test should be the preferred methodology, as this approach provides the most robust comparison with less risk of bias.[129, 130] In this systematic review only one study included direct comparison of CRP with PCT. Where this direct approach is not feasible then the use of network meta-analysis may provide an alternative option, but this requires researchers including statisticians to become familiar with this more complex test methodology. Few published articles were identified using network meta-analysis specifically for diagnostic test accuracy studies when exploring this option for this systematic review. Clearly this is an emerging field, but more detailed guidance for researchers with unambiguous examples would be welcomed.

### 4.8.5 Rapid reviews for diagnostic test accuracy studies

Considerable time and resources were directed to following the comprehensive search strategy methodology for this systematic review. Streamlining this procedure by using a rapid review process for searching and synthesis of information offers researchers an attractive and pragmatic option.[131] This could also improve upon the turnaround time when more rapid laboratory information is required. Half of Cochrane reviews were published later than the stated aim of two years.[132] Using limited or accelerated methods that have been validated for diagnostic test accuracy studies to expediate the

time and to obtain a comparable outcome would appear to be a significant advantage. This work is currently under way and researchers need to be aware of such advances.[131]

### 4.8.6 Assessing heterogeneity in diagnostic test accuracy studies

The assessment, measurement, interpretation and presentation of heterogeneity in meta-analysis of diagnostic test accuracy remains complex and challenging. Relying on subjective visual assessment to identify statistical heterogeneity is not uncommon but does not allow the extent of heterogeneity to be accurately assessed. There is general acknowledgement of the level of difficulty in identifying and measuring the degree of heterogeneity and the ways this is handled.[133] While these reviews should be carried out with a statistician familiar with this field, substantial education and training to assist researchers in performing these tasks is needed.

### 4.8.7 Closer research collaboration

Closer collaboration with paediatric clinicians, medical laboratory scientists and researchers for diagnostic test accuracy studies could improve the quality of published systematic reviews and meta-analyses. Topic content expertise is critical to ensuring clinical relevance within an appropriate setting, but this must be carefully balanced with systematic review expertise to provide the best available scientific evidence. Meta-analyses for diagnostic tests accuracy should always have available a specialist statistician to guide researchers.

## 4.9 Conclusion

This systematic review with meta-analyses provides clinicians and researchers with new evidence related to the test accuracy performance of CRP in children and adolescents with suspected osteoarticular infection.

We identified an insufficient number of scientific studies from our inclusion criteria to perform a meta-analysis of the test accuracy of PCT and compare this statistically with CRP. While other authors have achieved this by combining adult, child and adolescent data, we have demonstrated the potential limitations of transferring predominately adult outcomes to paediatric patients. Clinicians should remain vigilant to the indiscriminate creep of PCT use into routine practice without a clear understanding of the test accuracy performance in the paediatric setting.

Ideally, further studies using direct comparisons of CRP and PCT with a single reference standard in children and adolescents will be conducted. Future diagnostic test accuracy studies and systematic reviews should include a pre-specified hypothesis that can be statistically tested against a null hypothesis using sample size calculations when formally evaluating these index tests.

This review will permit researchers to update the meta-analysis for CRP and PCT in this setting by either adjusting inclusion criteria and or awaiting new publications. This research will hopefully

facilitate the next step using meta-analyses to compare the diagnostic test accuracy of CRP and PCT in children and adolescents with suspected osteoarticular infection.

# 5 Appendices

## 5.1 Appendix 1

5.1.1 The full search strategy for EMBASE (Ovid®) and Cochrane centre of controlled trials (Ovid®) is included below. All final searches were performed on the 30th July 2019.

#1 exp Osteomyelitis
#2 Arthritis, Infectious/
#3 Osteomyelitis.ti,ab,kw.
#4 ((osteoarticular or bone* or joint*) adj4 infect*).ti,ab,kw.
#5 (arthritis adj2 (septic or sepsis or infectious or bacterial)).ti,ab,kw.
#6 or/1-5
#7 Procalcitonin/
#8 (pro?calcitonin or PCT).ti,ab,kw.
#9 C-reactive protein/
#10 (C-reactive protein or CRP).ti,ab,kw.
#11 or/7-10
#12 Infant/ or exp Child/ or Adolescent/ or Paediatrics/
#13 (infant* or child* or adolescent* or juvenile*or p?ediatric* or girl* or boy*).af.
#14 12 or 13
#15 6 and 11 and 14
#16 15 not Animals/ not (Animals/ and Human/)

5.1.2 The full search strategy for MEDLINE (PubMed) is included below:

((((((((((osteomyelitis[MeSH Terms]) OR (osteomyelitis AND infection[MeSH Terms])) OR arthritis, infectious[MeSH Terms]) OR osteomyelitis[Title/Abstract]) OR (osteoarticular infection*[Title/Abstract] OR bone infection*[Title/Abstract] OR joint infection*[Title/Abstract])) OR (septic arthritis[Title/Abstract] OR infectious arthritis[Title/Abstract] OR bacterial arthritis[Title/Abstract]))) AND ((((procalcitonin[MeSH Terms]) OR (procalcitonin[Title/Abstract] OR PCT[Title/Abstract])) OR C-reactive protein[MeSH Terms]) OR (C-reactive protein[Title/Abstract] OR CRP[Title/Abstract]))) AND (((Infant or child or adolescent or pediatrics[MeSH Terms])) OR (infant*[Title/Abstract] OR child*[Title/Abstract] OR adolescent*[Title/Abstract] OR pediatric*[Title/Abstract] OR paediatric*[Title/Abstract]))))) NOT ((animals[MeSH Terms]) NOT humans[MeSH Terms])))

5.1.3 The full search strategy for Web of Science is included below:

TS=(osteomyelitis OR "septic arthritis" OR "infectious arthritis" OR "bacterial arthritis" OR "bone infection*" OR "joint infection*"OR "osteoarticular infection*") AND TS=(“c-reactive protein” OR CRP OR procalcitonin OR PCT) AND TS=(Infant* OR child* OR adolescent* OR p*ediatric*)

5.1.4 The full search strategy for Mednar :

(osteomyelitis OR “septic arthritis” OR “osteoarticular infection” OR “bone infection” OR “joint infection” OR “infectious arthritis” OR “bacterial arthritis” ) AND (“C-reactive protein” OR CRP OR procalcitonin OR PCT) AND (Infant* OR child* OR adolescent* OR p?ediatric)

5.1.5 The full search strategy for Open Grey:
osteomyelitis OR “septic arthritis” OR “osteoarticular infection” OR “bone infection” OR “joint infection” OR “infectious arthritis” OR “bacterial arthritis” AND “C-reactive protein” OR CRP OR procalcitonin OR PCT AND Infant* OR child* OR adolescent* OR paediatric* OR pediatric*

5.1.6 The search strategy for ProQuest dissertations and theses:

AB,TI(osteomyelitis OR “septic arthritis” OR “osteoarticular infection” OR “infectious arthritis” OR “bone infection” OR “joint infection” OR “bacterial arthritis”) AND AB,TI(infant* OR child* OR p?diatric* OR adolescent*)

5.1.7 The search strategy for Google scholar:

osteomyelitis OR “septic arthritis” OR “osteoarticular infection” OR “bone infection” OR “joint infection” OR “infectious arthritis” OR “bacterial arthritis” AND “C-reactive protein” OR CRP OR procalcitonin OR PCT AND Infant* OR child* OR adolescent* OR paediatric* OR pediatric

## 5.2 Appendix 2

Bibliography of excluded references and reasons:

1. Aigner RM, Fueger GF and Vejda M. Follow-up of osteomyelitis of infants with systemic serum parameters and bone scintigraphy. Nuklearmedizin 1996; 35(4): 116-121.
Reason for exclusion: No CRP cut off in article and no raw data included in analysis.

2. Akinkugbe O, Stewart C and McKenna C. Presentation and Investigation of Pediatric Bone and Joint Infections in the Pediatric Emergency Department. Pediatr Emerg Care 2018.
Reason for exclusion: No criteria included for the diagnosis of osteoarticular disease. No defined case definition.

3. Arnold JC, Cannavino CR, Ross MK, Westley B, Miller TC, Riffenburgh RH, et al. Acute bacterial osteoarticular infections: eight-year analysis of C-reactive protein for oral step-down therapy. Pediatrics 2012; 130(4): e821-828.
Reason for exclusion: CRP for oral stepdown and not a diagnostic paper.

4. Aupiais C, Basmaci R, Ilharreborde B, Blachier A, Desmarest M, Job-Deslandre C, et al. Arthritis in children: comparison of clinical and biological characteristics of septic arthritis and juvenile idiopathic arthritis. Arch Dis Child 2017; 102(4): 316-322.
Reason for exclusion: This study does not allow for extraction or calculation of any of the parameters to determine diagnostic test accuracy.

5. Bacon MA, Bailie HC and Stohr KK. Kocher's Criteria: Is it always useful? Archives of Disease in Childhood 2016; 101 (Supplement 1): A36-A37.
Reason for exclusion: No CRP raw data included and no definition of acute osteoaricular infection included.

6. Baker AD and Macnicol MF. Haematogenous osteomyelitis in children: epidemiology, classification, aetiology and treatment. Paediatrics and Child Health 2008; 18(2): 75-84.
Reason for exclusion: Review article only, no raw data included in this paper.

7. Barriocanal MB, Jimenez MR, Amador JTR, Insuga VS, Sanchez AB and Jareno MLL. Acute osteomyelitis: epidemiology, clinical manifestations, diagnosis and treatment. Anales De Pediatria 2013; 78(6): 367-373.
Reason for exclusion: Article written in Spanish.

8. Basmaci R, Ilharreborde B, Bonacorsi S, Kahil M, Mallet C, Aupiais C, et al. Septic arthritis in children with normal initial C-reactive protein: Clinical and biological features. Archives De Pediatrie 2014; 21(11): 1195-1199.
Reason for exclusion: Article written in French.

9. Benvenuti MA, An TJ, Mignemi ME, Martus JE, Mencio GA, Lovejoy SA, et al. A Clinical Prediction Algorithm to Stratify Pediatric Musculoskeletal Infection by Severity. Journal of Pediatric Orthopaedics 2019; 39(3): 153-157.
Reason for exclusion: This paper includes soft tissue infections and does not discriminate between bone and joint and soft tissue infections.

10. Bonhoeffer J, Haeberle B, Schaad UB and Heininger U. Diagnosis of acute haematogenous osteomyelitis and septic arthritis: 20 years experience at the University Children's Hospital Basel. Swiss medical weekly 2001; 131(39-40): 575-581.
Reason for exclusion: Criteria for reference test are insufficient.

11. Brown MD. Test characteristics of C-reactive protein (CRP) for pediatric septic arthritis. Journal of pediatric orthopedics 2004; 24(3): 344.
Reason for exclusion: Letter to editor.

12. Bueno Barriocanal M, Ruiz Jimenez M, Ramos Amador JT, Soto Insuga V, Bueno Sanchez A and Lorente Jareno ML. Acute osteomyelitis: Epidemiology, clinical manifestations, diagnosis and treatment. Anales de Pediatria 2013; 78(6): 367-373.
Reason for exclusion: Article written in Spanish.

13. Butbul-Aviel Y, Koren A, Halevy R and Sakran W. Procalcitonin as a diagnostic aid in osteomyelitis and septic arthritis. Pediatr Emerg Care 2005; 21(12): 828-832.
Reason for exclusion: Age group does not meet inclusion criteria.

14. Calvo C, Nunez E, Camacho M, Clemente D, Fernandez-Cooke E, Alcobendas R, et al. Epidemiology and management of acute, uncomplicated septic arthritis and osteomyelitis spanish multicenter study. Pediatric Infectious Disease Journal 2016; 35(12): 1288-1293.
Reason for exclusion: Sensitivity outcomes provided without corresponding specificity or other outcomes of diagnositic test accuracy.

15. Ceroni D, Regusci M, Pazos J, Dayer R and Kaelin A. Acute bone and joint infections in children: how much attention should be paid to persistent fever during intravenous antibiotic therapy? Revue De Chirurgie Orthopedique Et Reparatrice De L Appareil Moteur 2003; 89(3): 250-256.
Reason for exclusion: Article written in French, with only abstract in English.

16. Dartnell J, Ramachandran M and Katchburian M. Haematogenous acute and subacute paediatric osteomyelitis: A systematic review of the literature. Journal of Bone and Joint Surgery-British Volume 2012; 94B(5): 584-595.
Reason for exclusion: Systematic review without raw data for diagnosis test accuracy evaluation.

17. De Boeck H. Osteomyelitis and septic arthritis in children. Acta orthopaedica belgica 2005; 71(5): 505.
Reason for exclusion: Summary paper only without raw data for CRP diagnostic accuracy testing.

18. Delaney RA, Lenehan B, O'Sullivan L, McGuinness AJ and Street JT. The limping child: an algorithm to outrule musculoskeletal sepsis. Ir J Med Sci 2007; 176(3): 181-187.
Reason for exclusion: No raw data provided and no case definitions of OA/SA included.

19. Digman JM, Roddy R, Kaczenski A, Godbold S and Abramo T. Evaluation of procalcitonin as a negative predictor for serious bacterial infection in pediatrics. Annals of Emergency Medicine 2016; 68 (4 Supplement 1): S93.
Reason for exclusion: This article included all serious infections and not possible to extract osteoarticular infections. Abstract only format.

20. Dubois-FerriSre V, Belaieff W, Lascombes P, de Coulon G and Ceroni D. Transient synovitis of the hip: which investigations are truly useful? Swiss medical weekly 2015; 145.
Reason for exclusion: This paper does not include septic arthritis or osteomyelitis.

21. Ernst AA, Weiss SJ, Tracy LA and Weiss NR. Usefulness of CRP and ESR in predicting septic joints. South Med J 2010; 103(6): 522-526.
Reason for exclusion: Adult data only provided.

22. Gage MJ, Twomey KD, Sala DA, Maguire KJ, Hanstein R, Hennrikus WL, et al. Identifying Predictive Factors of Pediatric Septic Arthritis of the Knee in a Lyme Endemic Area. Bulletin of the Hospital for Joint Diseases 2018; 76(3): 161-164.
Reason for exclusion: Study type does not allow extraction of either sensitivity or specificity data for CRP.

23. Gandini D. Acute septic arthritis of the hip in children in northern Australia. ANZ J Surg 2003; 73(3): 136-139.
Reason for exclusion: Sensitivity data only is provided.

24. García-Arias M, Balsa A and Mola EM. Septic arthritis. Best practice & research Clinical rheumatology 2011; 25(3): 407-421.
Reason for exclusion: Summary paper, no diagnostic test accuracy data available.

25. Gibian JT, Daryoush JR, Wollenman CC, Johnson SR, Henry A, Koehler RJ, et al. The Heterogeneity of Pediatric Knee Infections: A Retrospective Analysis. Journal of pediatric orthopedics 2019.
Reason for exclusion: Calculation of sensitivity only therefore unable to determine diagnostic accuracy

26. Giordano M, Aulisa AG, Guzzanti V, Careri S, Krzysztofiak A and Toniolo RM. Managing of musculoskeletal infections in children. Eur Rev Med Pharmacol Sci 2019; 23(2 Suppl): 179-186.
Reason for exclusion: No raw data provided in this article.

27. Grummert E and Michael K. Pediatric hip pain: Transient synovitis versus septic arthritis. Journal of Diagnostic Medical Sonography 2006; 22(3): 185-188.
Reason for exclusion: Case presentation only

28. Hariharan P and Kabrhel C. Sensitivity of erythrocyte sedimentation rate and C-reactive protein for the exclusion of septic arthritis in emergency department patients. J Emerg Med 2011; 40(4): 428-431.
Reason for exclusion: Adult patients only presented.

29. Harris JC, Caesar DH, Davison C, Phibbs R and Than MP. How useful are laboratory investigations in the emergency department evaluation of possible osteomyelitis? Emerg Med Australas 2011; 23(3): 317-330.
Reason for exclusion: Summary information, no raw data provided.

30. Highton E, Perez MG, Cedillo Villamagua C, Sormani MI, Mussini MS, Isasmendi A, et al. Osteoarticular infections in a tertiary care children's hospital: Epidemiology and clinical characteristics in association with bacteremia. Arch Argent Pediatr 2018; 116(2): e204-e209.
Reason for exclusion: Article witten in Spanish.

31. Holloway E, Ruffles T and Godden C. Diagnosis and management in the child with suspected septic arthritis a 14 year retrospective study and review of the evidence. Archives of Disease in Childhood 2011; 96: A62.
Reason for exclusion: Abstract without raw data and case definitions not provided.

32. Howard A and Wilson M. Easily missed? Septic arthritis in children. BMJ (Online) 2010; 341(7776): 776-777.
Reason for exclusion: Letter to the editor.

33. Htun S. Phogenic osteomyelitis of long bones in children. Ann Arbor: The University of Liverpool (United Kingdom); 1975. p. 1.
Reason for exclusion: No test accuracy data.

34. Hugle T, Schuetz P, Mueller B, Laifer G, Tyndall A, Regenass S, et al. Serum procalcitonin for discrimination between septic and non-septic arthritis. Clinical and Experimental Rheumatology 2008; 26(3): 453-456.
Reason for exclusion: Unable to separate patients 16-18 years from general adult cohort.

35. Inusa BPD, Oyewo A, Brokke F, Santhikumaran G and Jogeesvaran KH. Dilemma in Differentiating between Acute Osteomyelitis and Bone Infarction in Children with Sickle Cell Disease: The Role of Ultrasound. PLoS ONE 2013; 8(6).
Reason for exclusion: Variable case definitions included without being clearly defined.

36. Jain S, Tittal P, Rohilla N, Sud A, Yadav C, Kanojia RK, et al. Acute septic arthritis revisited: a prospective study in 93 patients correlating C-reactive protein levels with duration of intravenous antibiotic therapy, clinical and radiological outcomes. European Journal of Orthopaedic Surgery & Traumatology 2009; 19(7): 447-455.
Reason for exclusion: Case definitions of patients imprecise.

37. Jana FN, Ortega CS and Goiano EO. Epidemiological study of osteoarticular infections in children. Acta ortopedica brasileira 2018; 26(3): 201-205.
Reason for exclusion: Includes neonates.

38. Journeau P, Wein F, Popkov D, Philippe R, Haumont T and Lascombes P. Hip septic arthritis in children: assessment of treatment using needle aspiration/irrigation. Orthop Traumatol Surg Res 2011; 97(3): 308-313.
Reason for exclusion: Not a diagnostic testing assay article.

39. Kallio MJ, Unkila-Kallio L, Aalto K and Peltola H. Serum C-reactive protein, erythrocyte sedimentation rate and white blood cell count in septic arthritis of children. Pediatr Infect Dis J 1997; 16(4): 411-413.
Reason for exclusion: Sensitivity only provided and unable to calculate diagnostic test accuracy.

40. Kao HC, Huang YC, Chiu C-H, Chang LY, Lee ZL, Chung PW, et al. Acute hematogenous osteomyelitis and septic arthritis in children. Journal of microbiology immunology and infection 2003; 36(4): 260-265.
Reason for exclusion: Summary data only included in this review.

41. Karambin MM and Hashemian H. Childhood arthritis: Rate of different types. Acta Medica Iranica 2009; 47(1): 31-34.
Reason for exclusion: Examined only arthritis, not a diagnostic test paper to evaluate CRP.

42. Kocher MS, Lee B, Dolan M, Weinberg J and Shulman ST. Pediatric orthopedic infections: Early detection and treatment. Pediatric Annals 2006; 35(2): 112-122.
Reason for exclusion: Review article with only summary information.

43. Kocher MS, Mandiga R, Zurakowski D, Barnewolt C and Kasser JR. Validation of a clinical prediction rule for the differentiation between septic arthritis and transient synovitis of the hip in children. JBJS 2004; 86(8): 1629-1635.
Reason for exclusion: CRP not included in the analysis.

44. Kopchak O and Kostik M. Differential diagnosis of non-bacterial and acute hemotogenous osteomyelitis. Pediatric Rheumatology 2017; 15 (Supplement 2): 102.
Reason for exclusion: Not diagnostic test evaluation, summary data only included.

45. Krzysztofiak A, Bozzola E, Lancella L, Marchesi A and Villani A. Osteomyelitis in children. Italian Journal of Pediatrics Conference: 72th Congress of the Italian Society of Pediatrics Italy 2017; 43(1 Supplement).
Reason for exclusion: Summary abstract only, no CRP or procacitonin included.

46. Kunnamo I, Kallio P, Pelkonen P and Hovi T. Clinical signs and laboratory tests in the differential diagnosis of arthritis in children. Am J Dis Child 1987; 141(1): 34-40.
Reason for exclusion: Could not exclude neonates in this review.

47. Kuong EE, To M, Chow W, Yuen MH, Choi AKY and Fong CM. Pitfalls in diagnosing septic arthritis in hong kong children: Ten years' experience. Hong Kong Medical Journal 2012; 18(6): 482-487.
Reason for exclusion: Only included sensitivity data therefore unable to calculate diagnostic accuracy parameters.

48. Lalanda M and Alonso JA. Improving the management of the child with an unexplained limp. Clinical Governance 2006; 11(4): 308-315.
Reason for exclusion: Testing algorithim only provided.

49. Lavy CBD. The clinical features and surgical treatment of acute septic arthritis in Malawian children. Ann Arbor: University of London, University College London (United Kingdom); 2006. p. 154.
Reason for exclusion: No diagnostic accuracy test outcomes.

50. Lazzarini L, Mader JT and Calhoun JH. Osteomyelitis in long bones. JBJS 2004; 86(10): 2305-2318.
Reason for exclusion: Treatment article in summary format, not diagnostic test accuracy paper.

51. Le Saux N. The diagnosis and treatment of acute osteoarticular infections in children. Paediatrics & child health 2018; 23(5): 344-352.
Reason for exclusion: Article written in French.

52. Lee S, Kim HW, Cho HK, Yun YH, Ryu KH and Kim KH. Clinical presentations and causative organisms in children and adolescents with osteoarticular infections: A retrospective study. [Korean]. Pediatric Infection and Vaccine 2015; 22(3): 154-163.
Reason for exclusion: Article written in Korean.

53. Levine MJ, McGuire KJ, McGowan KL and Flynn JM. Assessment of the test characteristics of C-reactive protein for septic arthritis in children. Journal of pediatric orthopedics 2003; 23(3): 373-377.
Reason for exclusion: Unable to determine upper or lower age limits of patients included.

54. Lew DP and Waldvogel FA. Osteomyelitis. The Lancet 2004; 364(9431): 369-379.
Reason for exclusion: Summary review article, no diagnostic test accuracy data.

55. Li SF, Cassidy C, Chang C, Gharib S and Torres J. Diagnostic utility of laboratory tests in septic arthritis. Emergency Medicine Journal 2007; 24(2): 75-77.
Reason for exclusion: No CRP or procalcitonin included in this article.

56. Lin Z, Vasudevan A and Tambyah PA. Use of erythrocyte sedimentation rate and C-reactive protein to predict osteomyelitis recurrence. J Orthop Surg (Hong Kong) 2016; 24(1): 77-83.
Reason for exclusion: No cases of acute osteomyelitis.

57. Lorrot M, Fitoussi F, Faye A, Mariani P, Job-Deslandre C, Penneçot G, et al. Laboratory studies in pediatric bone and joint infections. Archives de pediatrie: organe officiel de la Societe francaise de pediatrie 2007; 14: S86-90.
Reason for exclusion: Article written in French.

58. Macnicol MF. Patterns of musculoskeletal infection in childhood. Journal of Bone and Joint Surgery - Series B 2001; 83(1): 1-2.
Reason for exclusion: Editorial only.

59. Macnicol MF and Watts AC. Haematogenous osteomyelitis. Surgery (Oxford) 2005; 23(1): 25-30.
Reason for exclusion: Summary article.

60. Madruga Dias J, Costa MM, Pereira da Silva JA and Viana de Queiroz M. Septic arthritis: patients with or without isolated infectious agents have similar characteristics. Infection 2014; 42(2): 385-391.
Reason for exclusion: Adult cohort of patients.

61. Maharajan K, Patro DK, Menon J, Hariharan AP, Parija SC, Poduval M, et al. Serum Procalcitonin is a sensitive and specific marker in the diagnosis of septic arthritis and acute osteomyelitis. J Orthop Surg Res 2013; 8: 19.
Reason for exclusion: Included adults.

62. Mahmoudi S, Pourakbari B, Borhani K, Khodabandeh M, Valian SK, Aziz-Ahari A, et al. Acute osteomyelitis and septic arthritis in children: A referral hospital-based study in Iran. Wiener Medizinische Wochenschrift 2017; 167(11-12): 259-263. S
Reason for exclusion: Sensitivity only provided and unable to extract or calculate diagnostic accuracy.

63. Mamishi S, Kalantari N and Pourakbari B. Clinical feature and etiology of septic arthritis and osteomyelitis in children in a 10 year period. Acta Medica Iranica 2007; 45(1): 58-62.
64. Matan AJ and Smith JT. Pediatric septic arthritis. Orthopedics 1997; 20(7): 630-635.
Reason for exclusion: No CRP or procalcitonin included.

65. McMahon A-M. Bone and joint infections. Paediatrics and Child Health 2011; 21(12): 546-551.
Reason for exclusion:Summary article only.

66. Mistry RM, Lennon D, Boyle MJ, Chivers K, Frampton C, Nicholson R, et al. Septic arthritis and acute rheumatic fever in children: the diagnostic value of serological inflammatory markers. Journal of pediatric orthopedics 2015; 35(3): 318-322.
Reason for exclusion: Not a diagnostic test accuracy study, recommendations only.

67. Mitchell PD, Viswanath A, Obi N, Littlewood A and Latimer M. A prospective study of screening for musculoskeletal pathology in the child with a limp or pseudoparalysis using erythrocyte sedimentation rate, C-reactive protein and MRI. J Child Orthop 2018; 12(4): 398-405.
Reason for exclusion: Included all infections and unable to isolate osteoarticular infections.

68. Monachello Araujo D, Cordero Garcia JM, Garcia Zogby L, Guzman Cruz AJ, Rodado Marina S, Coronado Poggio M, et al. Osteomyelitis in children: Association between analytical parameters and three-phase bone scan findings. European Journal of Nuclear Medicine and Molecular Imaging 2018; 45 (Supplement 1): S350.
Reason for exclusion:Abstract without raw data for analysis.

69. Moro-Lago I, Talavera G, Moraleda L and Gonzalez-Moran G. Clinical presentation and treatment of septic arthritis in children. Revista espanola de cirugia ortopedica y traumatologia 2017; 61(3): 170-175.
Reason for exclusion: Article written in Spanish text.

70. Nannery R and Heinz P. Approach to joint pain in children. Paediatrics and Child Health (United Kingdom) 2018; 28(2): 43-49.
Reason for exclusion: Summary article.

71. Paakkonen M, Kallio MJ, Kallio PE and Peltola H. Sensitivity of erythrocyte sedimentation rate and C-reactive protein in childhood bone and joint infections. Clin Orthop Relat Res 2010; 468(3): 861-866.
Reason for exclusion: Sensitivity data only provided and unable to calculate diagnostic accuracy test performance.

72. Paakkonen M, Kallio MJ, Kallio PE and Peltola H. C-reactive protein versus erythrocyte sedimentation rate, white blood cell count and alkaline phosphatase in diagnosing bacteraemia in bone and joint infections. J Paediatr Child Health 2013; 49(3): E189-192.
Reason for exclusion: Focus on children with bacteraemia and testing accuracy and not primary osteoarticular infection.

73. Patel L, Michael J, Schroeder L, Sherman AK, Berglund L and Newland JG. Can a Septic Hip Decision Rule Aid in the Evaluation of Suspected Pediatric Musculoskeletal Infections? J Emerg Med 2019; 56(3): 241-247.
Reason for exclusion: Included all musculoskeletal infections.

74. Peltola H, Vahvanen V and Aalto K. Fever, C-reactive protein, and erythrocyte sedimentation rate in monitoring recovery from septic arthritis: a preliminary study. Journal of pediatric orthopedics 1984; 4(2): 170-174.
Reason for exclusion: Markers used to monitor the effectiveness of treatment only.

75. Perry M. Erythrocyte sedimentation rate and C reactive protein in the assessment of suspected bone infection--are they reliable indices? J R Coll Surg Edinb 1996; 41(2): 116-118.
Reason for exclusion: Single case report.

76. Reed L, Baskett A and Watkins N. Managing children with acute non-traumatic limp: The utility of clinical findings, laboratory inflammatory markers and X-rays: Paediatric Emergency Medicine. EMA - Emergency Medicine Australasia 2009; 21(2): 136-142.
Reason for exclusion: Included neonates, no defined reference test description and no case definitions.

77. Reitzenstein JE, Yamamoto LG and Mavoori H. Similar erythrocyte sedimentation rate and C-reactive protein sensitivities at the onset of septic arthritis, osteomyelitis, acute rheumatic fever. Pediatric reports 2010; 2(1): e10.
Reason for exclusion: No criteria provided to determine the reference test standard.

78. Riise ØR, Kirkhus E, Handeland KS, Flatø B, Reiseter T, Cvancarova M, et al. Childhood osteomyelitis-incidence and differentiation from other acute onset musculoskeletal features in a population-based study. BMC pediatrics 2008; 8.
Reason for exclusion: Included all infections and not examining diagnostic test performance.

79. Robinson S and Leonard P. C reactive protein, erythrocyte sedimentation rate, or both, in the diagnosis of atraumatic paediatric limb pain? Emergency Medicine Journal 2012; 29(12): 969-971.
Reason for exclusion: Gold standard reference test opinion by orthopaedic team. Not objective.

80. Rutz E and Spoerri M. Septic arthritis of the paediatric hip - A review of current diagnostic approaches and therapeutic concepts. Acta Orthop Belg 2013; 79(2): 123-134.
Reason for exclusion:Diagnostic algorithm

81. Saavedra-Lozano J, Falup-Pecurariu O, Faust SN, Girschick H, Hartwig N, Kaplan S, et al. Bone and Joint Infections. Pediatric Infectious Disease Journal 2017; 36(8): 788-799.

Reason for exclusion: Guideline

82. Sanabria-Baez G, Delgado V and Arbo A. Clinical and microbiological features of septic arthritis in children: Experience of one Latin American country. Journal of Infection in Developing Countries 2017; 11(12): 971-973.
Reason for exclusion: CRP evaluation not included in this paper.

83. Sanchez-Moreno P, Ardanuy-Pizarro AV, Navarro L, Melon M, Falcon-Neyra MD and Camacho-Lovillo M. Acute osteoarticular infections in children in a tertiary hospital: Our experience across 5 years. Annals of the Rheumatic Diseases 2015; 74: 1233. abstract without details for use of CRP
84. Simon L, Gauvin F, Amre DK, Saint-Louis P and Lacroix J. Serum procalcitonin and C-reactive protein levels as markers of bacterial infection: a systematic review and meta-analysis. Clinical infectious diseases 2004; 39(2): 206-217.
Reason for exclusion: Meta-analysis for all bacterial infections in general and not osteoarticular infections.

85. Smith NSM, Wilson N and Janet GM. Differentiating between jia and septic arthritis: Identification of important features in history, examination and laboratory investigation. Rheumatology (United Kingdom) 2012; 51: viii11.
Reason for exclusion: Abstract, no CRP raw data that can be extracted provided.

86. Talebi-Taher M, Shirani F, Nikanjam N and Shekarabi M. Septic versus inflammatory arthritis: Discriminating the ability of serum inflammatory markers. Rheumatology International 2013; 33(2): 319-324.
Reason for exclusion:Adult population only.

87. Unkila-Kallio L, Kallio M and Peltola H. Acute haematogenous osteomyelitis in children in Finland. Finnish Study Group. Annals of medicine 1993; 25(6): 545-549.
Reason for exclusion: Only sensitivity data provided without corresponding specificity.Unable to calculate diagnostic accuracy

88. Unkila-Kallio L, Kallio M and Peltola H. The usefulness of C-reactive protein levels in the identification of concurrent septic arthritis in children who have acute hematogenous osteomyelitis. A comparison with the usefulness of the erythrocyte sedimentation rate and the white blood-cell count. The Journal of bone and joint surgery American volume 1994; 76(6): 848-853.
Reason for exclusion: Comparing osteomyeltitis and septic arthritis in neonates only.

89. Unkila-Kallio L, Kallio MJ, Eskola J and Peltola H. Serum C-reactive protein, erythrocyte sedimentation rate, and white blood cell count in acute hematogenous osteomyelitis of children. Pediatrics 1994; 93(1): 59-62.

Reason for exclusion:Only sensitivity data provided and unable to calculate the diagnostic accuracy.

90. Wasz-Hockert O, Stenman UH and Kostia J. Monarthritis in children. An etiological, clinical and differential diagnostical study on 264 children. Ann Paediatr Fenn 1965; 11(3): 119-133.
Reason for exclusion: no CRP included

91. Welling BD, Haruno LS and Rosenfeld SB. Validating an Algorithm to Predict Adjacent Musculoskeletal Infections in Pediatric Patients With Septic Arthritis. Clin Orthop Relat Res 2018; 476(1): 153-159.
Reason for exclusion: CRP raw data or diagnostic test accuracy parameters not provided.

92. Yeo A and Ramachandran M. Acute haematogenous osteomyelitis in children. Bmj-British Medical Journal 2014; 348.
Reason for exclusion: Review article.

93. Zamani A, Raeeskarami SR, Akbrai Asbagh P, Oloomi Yazdi Z, Matloob R, Zamani N, et al. Pediatric septic arthritis: A 10- year epidemiologic study in Imam Khomeini Hospital Complex. Tehran University Medical Journal 2010; 67(10): 736-742.
Reason for exclusion: Article written in Arabic.

# 6 Bibliography


1. Ritchie B, Porritt K, Marin T, Williams N. Diagnostic test accuracy of serum measurement of procalcitonin and C-reactive protein for bone and joint infection in children and adolescents: a systematic review protocol. JBI Evid Synth. 2020;18(3):564-70.
2. Gillen G, Hunter EG, Lieberman D, Stutzbach M. AOTA's Top 5 Choosing Wisely((R)) Recommendations. Am J Occup Ther. 2019;73(2):7302420010p1-p9.
3. Lindner RA. Choosing Wisely Australia: changing behaviour in health care. Med J Aust. 2018;208(3):105-6.
4. Weidner W, Wagenlehner FM. Re: Procalcitonin as a diagnostic marker for sepsis: a systematic review and meta-analysis. Eur Urol. 2014;66(1):178.
5. Tan M, Lu Y, Jiang H, Zhang L. The diagnostic accuracy of procalcitonin and C-reactive protein for sepsis: A systematic review and meta-analysis. J Cell Biochem. 2019;120(4):5852-9.
6. Ruan L, Chen GY, Liu Z, Zhao Y, Xu GY, Li SF, et al. The combination of procalcitonin and C-reactive protein or presepsin alone improves the accuracy of diagnosis of neonatal sepsis: a meta-analysis and systematic review. Crit Care. 2018;22(1):316.
7. Simon L, Gauvin F, Amre DK, Saint-Louis P, Lacroix J. Serum procalcitonin and C-reactive protein levels as markers of bacterial infection: a systematic review and meta-analysis. Clinical infectious diseases. 2004;39(2):206-17.
8. Yo CH, Hsieh PS, Lee SH, Wu JY, Chang SS, Tasi KC, et al. Comparison of the test characteristics of procalcitonin to C-reactive protein and leukocytosis for the detection of serious bacterial infections in children presenting with fever without source: a systematic review and meta-analysis. Ann Emerg Med. 2012;60(5):591-600.
9. Trippella G, Galli L, De Martino M, Lisi C, Chiappini E. Procalcitonin performance in detecting serious and invasive bacterial infections in children with fever without apparent source: a systematic review and meta-analysis. Expert Rev Anti Infect Ther. 2017;15(11):1041-57.
10. Saavedra-Lozano J, Falup-Pecurariu O, Faust SN, Girschick H, Hartwig N, Kaplan S, et al. Bone and Joint Infections. Pediatr Infect Dis J. 2017;36(8):788-99.
11. Shen CJ, Wu MS, Lin KH, Lin WL, Chen HC, Wu JY, et al. The use of procalcitonin in the diagnosis of bone and joint infection: a systemic review and meta-analysis. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology. 2013;32(6):807-14.
12. Zhao J, Zhang S, Zhang L, Dong X, Li J, Wang Y, et al. Serum procalcitonin levels as a diagnostic marker for septic arthritis: A meta-analysis. Am J Emerg Med. 2017;35(8):1166-71.
13. Gutierrez K. Bone and joint infections in children. Pediatr Clin North Am. 2005;52(3):779-94, vi.
14. Dodwell ER. Osteomyelitis and septic arthritis in children: current concepts. Curr Opin Pediatr. 2013;25(1):58-63.
15. Brischetto A, Leung G, Marshall CS, Bowen AC. A retrospective case-series of children with bone and joint infection from northern Australia. Medicine (Baltimore). 2016;95(8):e2885.
16. Paakkonen M, Peltola H. Acute osteomyelitis in children. N Engl J Med. 2014;370(14):1365-6.
17. Caksen H, Ozturk MK, Uzum K, Yuksel S, Ustunbas HB, Per H. Septic arthritis in childhood. Pediatr Int. 2000;42(5):534-40.
18. Jagodzinski NA, Kanwar R, Graham K, Bache CE. Prospective evaluation of a shortened regimen of treatment for acute osteomyelitis and septic arthritis in children. JPediatr Orthop. 2009;29(5):518-25.
19. Le Saux N. Diagnosis and management of acute osteoarticular infections in children. Paediatrics & child health. 2018;23(5):336-43.
20. Williams N, Cooper C, Cundy P. Kingella kingae septic arthritis in children: recognising an elusive pathogen. J Child Orthop. 2014;8(1):91-5.
21. Lew DP, Waldvogel FA. Osteomyelitis. Lancet. 2004;364(9431):369-79.
22. Mooney JF, 3rd, Murphy RF. Septic arthritis of the pediatric hip: update on diagnosis and treatment. Curr Opin Pediatr. 2019;31(1):79-85.
23. Dartnell J, Ramachandran M, Katchburian M. Haematogenous acute and subacute paediatric osteomyelitis: a systematic review of the literature. J Bone Joint Surg Br. 2012;94(5):584-95.
24. Iliadis AD, Ramachandran M. Paediatric bone and joint infection. EFORT Open Rev. 2017;2(1):7-12.
25. Thevenin-Lemoine C, Vial J, Labbe JL, Lepage B, Ilharreborde B, Accadbled F, et al. MRI of acute osteomyelitis in long bones of children: Pathophysiology study. Orthop Traumatol Surg Res. 2016;102(7):831-7.

26. Riise OR, Kirkhus E, Handeland KS, Flato B, Reiseter T, Cvancarova M, et al. Childhood osteomyelitis-incidence and differentiation from other acute onset musculoskeletal features in a population-based study. BMC Pediatr. 2008;8:45.
27. Levine MJ, McGuire KJ, McGowan KL, Flynn JM. Assessment of the test characteristics of C-reactive protein for septic arthritis in children. Journal of pediatric orthopedics. 2003;23(3):373-7.
28. Arnold JC, Cannavino CR, Ross MK, Westley B, Miller TC, Riffenburgh RH, et al. Acute bacterial osteoarticular infections: eight-year analysis of C-reactive protein for oral step-down therapy. Pediatrics. 2012;130(4):e821-8.
29. Wu JS, Gorbachova T, Morrison WB, Haims AH. Imaging-guided bone biopsy for osteomyelitis: are there factors associated with positive or negative cultures? AJR Am J Roentgenol. 2007;188(6):1529-34.
30. Butbul-Aviel Y, Koren A, Halevy R, Sakran W. Procalcitonin as a diagnostic aid in osteomyelitis and septic arthritis. Pediatr Emerg Care. 2005;21(12):828-32.
31. Faesch S, Cojocaru B, Hennequin C, Pannier S, Glorion C, Lacour B, et al. Can procalcitonin measurement help the diagnosis of osteomyelitis and septic arthritis? A prospective trial. Ital J Pediatr. 2009;35(1):33.
32. Manz N, Krieg AH, Heininger U, Ritz N. Evaluation of the current use of imaging modalities and pathogen detection in children with acute osteomyelitis and septic arthritis. Eur J Pediatr. 2018;177(7):1071-80.
33. Monsalve J, Kan JH, Schallert EK, Bisset GS, Zhang W, Rosenfeld SB. Septic arthritis in children: frequency of coexisting unsuspected osteomyelitis and implications on imaging work-up and management. AJR Am J Roentgenol. 2015;204(6):1289-95.
34. Russell CD, Ramaesh R, Kalima P, Murray A, Gaston MS. Microbiological characteristics of acute osteoarticular infections in children. J Med Microbiol. 2015;64(Pt 4):446-53.
35. Dierig A, Berger C, Agyeman PKA, Bernhard-Stirnemann S, Giannoni E, Stocker M, et al. Time-to-Positivity of Blood Cultures in Children With Sepsis. Front Pediatr. 2018;6:222.
36. Paakkonen M, Kallio MJ, Kallio PE, Peltola H. Sensitivity of erythrocyte sedimentation rate and C-reactive protein in childhood bone and joint infections. Clin Orthop Relat Res. 2010;468(3):861-6.
37. Nabulsi M, Hani A, Karam M. Impact of C-reactive protein test results on evidence-based decision-making in cases of bacterial infection. BMC Pediatr. 2012;12:140.
38. Aguiar FJ, Ferreira-Junior M, Sales MM, Cruz-Neto LM, Fonseca LA, Sumita NM, et al. C-reactive protein: clinical applications and proposals for a rational use. Rev Assoc Med Bras (1992). 2013;59(1):85-92.
39. Pepys MB, Hirschfield GM. C-reactive protein: a critical update. J Clin Invest. 2003;111(12):1805-12.
40. Salazar J, Martinez MS, Chavez-Castillo M, Nunez V, Anez R, Torres Y, et al. C-Reactive Protein: An In-Depth Look into Structure, Function, and Regulation. Int Sch Res Notices. 2014;2014:653045.
41. Schlenz H, Intemann T, Wolters M, Gonzalez-Gil EM, Nappo A, Fraterman A, et al. C-reactive protein reference percentiles among pre-adolescent children in Europe based on the IDEFICS study population. Int J Obes (Lond). 2014;38 Suppl 2:S26-31.
42. Rodoo P, Ridefelt P, Aldrimer M, Niklasson F, Gustafsson J, Hellberg D. Population-based pediatric reference intervals for HbA1c, bilirubin, albumin, CRP, myoglobin and serum enzymes. Scand J Clin Lab Invest. 2013;73(5):361-7.
43. Nora D, Salluh J, Martin-Loeches I, Povoa P. Biomarker-guided antibiotic therapy-strengths and limitations. Ann Transl Med. 2017;5(10):208.
44. Sproston NR, Ashworth JJ. Role of C-Reactive Protein at Sites of Inflammation and Infection. Front Immunol. 2018;9:754.
45. Landry A, Docherty P, Ouellette S, Cartier LJ. Causes and outcomes of markedly elevated C-reactive protein levels. Can Fam Physician. 2017;63(6):e316-e23.
46. Caird MS, Flynn JM, Leung YL, Millman JE, D'Italia JG, Dormans JP. Factors distinguishing septic arthritis from transient synovitis of the hip in children. A prospective study. J Bone Joint Surg Am. 2006;88(6):1251-7.
47. Lammin KAT, J. ; Zenios ,M. The usefulness of CRP and ESR in diagnosing long bone osteomyelitis in children-are we being falsely reassured? Orthopaedic Proceedings. 2012;94-B:156.
48. Samsudin I, Vasikaran SD. Clinical utility and measurement of procalcitonin. Clin Biochem Rev. 2017;38(2):59-68.
49. Becker KL, Snider R, Nylen ES. Procalcitonin in sepsis and systemic inflammation: a harmful biomarker and a therapeutic target. Br J Pharmacol. 2010;159(2):253-64.

50. Rowland T, Hilliard H, Barlow G. Procalcitonin: potential role in diagnosis and management of sepsis. Adv Clin Chem. 2015;68:71-86.
51. Matera G, Quirino A, Giancotti A, Pulicari MC, Rametti L, Rodriguez ML, et al. Procalcitonin neutralizes bacterial LPS and reduces LPS-induced cytokine release in human peripheral blood mononuclear cells. BMC Microbiol. 2012;12:68.
52. Meisner M. Update on procalcitonin measurements. Ann Lab Med. 2014;34(4):263-73.
53. Gilbert DN. Use of plasma procalcitonin levels as an adjunct to clinical microbiology. J Clin Microbiol. 2010;48(7):2325-9.
54. Taylor R, Jones A, Kelly S, Simpson M, Mabey J. A Review of the Value of Procalcitonin as a Marker of Infection. Cureus. 2017;9(4):e1148.
55. McCluskey SM, Schuetz P, Abers MS, Bearnot B, Morales ME, Hoffman D, et al. Serial Procalcitonin as a Predictor of Bacteremia and Need for Intensive Care Unit Care in Adults With Pneumonia, Including Those With Highest Severity: A Prospective Cohort Study. Open Forum Infect Dis. 2017;4(1):ofw238.
56. Schuetz P, Birkhahn R, Sherwin R, Jones AE, Singer A, Kline JA, et al. Serial Procalcitonin Predicts Mortality in Severe Sepsis Patients: Results From the Multicenter Procalcitonin MOnitoring SEpsis (MOSES) Study. Crit Care Med. 2017;45(5):781-9.
57. Maruna P, Nedelnikova K, Gurlich R. Physiology and genetics of procalcitonin. Physiol Res. 2000;49 Suppl 1:S57-61.
58. Bobillo-Perez S, Rodriguez-Fanjul J, Jordan Garcia I. Is Procalcitonin Useful in Pediatric Critical Care Patients? Biomark Insights. 2018;13:1177271918792244.
59. Maharajan K, Patro DK, Menon J, Hariharan AP, Parija SC, Poduval M, et al. Serum Procalcitonin is a sensitive and specific marker in the diagnosis of septic arthritis and acute osteomyelitis. J Orthop Surg Res. 2013;8:19.
60. Kamat IS, Ramachandran V, Eswaran H, Guffey D, Musher DM. Procalcitonin to Distinguish Viral From Bacterial Pneumonia: A Systematic Review and Meta-analysis. Clin Infect Dis. 2020;70(3):538-42.
61. Sager R, Kutz A, Mueller B, Schuetz P. Procalcitonin-guided diagnosis and antibiotic stewardship revisited. BMC Med. 2017;15(1):15.
62. England JT, Del Vecchio MT, Aronoff SC. Use of serum procalcitonin in evaluation of febrile infants: a meta-analysis of 2317 patients. J Emerg Med. 2014;47(6):682-8.
63. Hu L, Shi Q, Shi M, Liu R, Wang C. Diagnostic value of PCT and CRP for detecting serious bacterial infections in patients with fever of unknown origin: a systematic review and meta-analysis. Appl Immunohistochem Mol Morphol. 2017;25(8):e61-e9.
64. Vouloumanou EK, Plessa E, Karageorgopoulos DE, Mantadakis E, Falagas ME. Serum procalcitonin as a diagnostic marker for neonatal sepsis: a systematic review and meta-analysis. Intensive Care Med. 2011;37(5):747-62.
65. Yu Z, Liu J, Sun Q, Qiu Y, Han S, Guo X. The accuracy of the procalcitonin test for the diagnosis of neonatal sepsis: a meta-analysis. Scand J Infect Dis. 2010;42(10):723-33.
66. Brown RB. Breakthrough Knowledge Synthesis in the Age of Google. Philosopies. 2020;5(4):1-13.
67. Garritty C, Stevens A, Hamel C, Golfam M, Hutton B, Wolfe D. Knowledge Synthesis in Evidence-Based Medicine. Semin Nucl Med. 2019;49(2):136-44.
68. Chalmers I, Hedges LV, Cooper H. A brief history of research synthesis. Eval Health Prof. 2002;25(1):12-37.
69. Tricco AC, Tetzlaff J, Moher D. The art and science of knowledge synthesis. J Clin Epidemiol. 2011;64(1):11-20.
70. Munn Z, Stern C, Aromataris E, Lockwood C, Jordan Z. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. BMC Med Res Methodol. 2018;18(1):5.
71. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, the P-DTAG, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA. 2018;319(4):388-96.
72. Wormald R, Evans J. What Makes Systematic Reviews Systematic and Why are They the Highest Level of Evidence? Ophthalmic Epidemiol. 2018;25(1):27-30.
73. Moller AM, Myles PS. What makes a good systematic review and meta-analysis? Br J Anaesth. 2016;117(4):428-30.
74. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015;350:g7647.

75. Leeflang MM. Systematic reviews and meta-analyses of diagnostic test accuracy. Clin Microbiol Infect. 2014;20(2):105-13.
76. Reitsma JB, Moons KG, Bossuyt PM, Linnet K. Systematic reviews of studies quantifying the accuracy of diagnostic tests and markers. Clin Chem. 2012;58(11):1534-45.
77. Gopalakrishnan S, Ganeshkumar P. Systematic Reviews and Meta-analysis: Understanding the Best Evidence in Primary Healthcare. J Family Med Prim Care. 2013;2(1):9-14.
78. Mulrow CD. Rationale for systematic reviews. BMJ. 1994;309(6954):597-9.
79. Vetter TR. Systematic Review and Meta-analysis: Sometimes Bigger Is Indeed Better. Anesth Analg. 2019;128(3):575-83.
80. Dickersin K, Berlin JA. Meta-analysis: state-of-the-science. Epidemiol Rev. 1992;14:154-76.
81. Cook DJ, Mulrow CD, Haynes RB. Systematic reviews: synthesis of best evidence for clinical decisions. Ann Intern Med. 1997;126(5):376-80.
82. Campbell JM, Klugar M, Ding S, Carmody DP, Hakonsen SJ, Jadotte YT, et al. Diagnostic test accuracy: methods for systematic review and meta-analysis. Int J Evid Based Healthc. 2015;13(3):154-62.
83. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open. 2016;6(11):e012799.
84. Wu S, Zhou X, Ye Y. Doubts on the meta-analysis of serum procalcitonin levels as a diagnostic marker for septic arthritis. Am J Emerg Med. 2018;36(11):2104-5.
85. Simeng W, Zhou X, Ye Y, Tang G, Wu F. A systematic review and meta-analysis of serum procalcitonin levels as a diagnostic marker of joint infection. CRD42018085549: PROSPERO; 2018.
86. Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. Value Health. 2011;14(4):417-28.
87. McGrath TA, Alabousi M, Skidmore B, Korevaar DA, Bossuyt PMM, Moher D, et al. Recommendations for reporting of systematic reviews and meta-analyses of diagnostic test accuracy: a systematic review. Syst Rev. 2017;6(1):194.
88. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36.
89. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151(4):264-9, W64.
90. Sultan J, Hughes PJ. Septic arthritis or transient synovitis of the hip in children THE VALUE OF CLINICAL PREDICTION ALGORITHMS. Journal of Bone and Joint Surgery-British Volume. 2010;92B(9):1289-93.
91. Singhal R, Perry DC, Khan FN, Cohen D, Stevenson HL, James LA, et al. The use of CRP within a clinical prediction algorithm for the differentiation of septic arthritis and transient synovitis in children. J Bone Joint Surg Br. 2011;93(11):1556-61.
92. Eich GF, Superti-Furga A, Umbricht FS, Willi UV. The painful hip: evaluation of criteria for clinical decision-making. Eur J Pediatr. 1999;158(11):923-8.
93. Greeff E. Is procalcitonin useful in diagnosing septic arthritis and osteomyelitis in children? SA Orthopaedic Journal. 2012;11(1):52-6.
94. Jung ST, Rowe SM, Moon ES, Song EK, Yoon TR, Seo HY. Significance of laboratory and radiologic findings for differentiating between septic arthritis and transient synovitis of the hip. Journal of pediatric orthopedics. 2003;23(3):368-72.
95. Cui C, Fu M, Gao B. Procalcitonin and Pancreatic Stone Protein Function as Biomarkers in Early Diagnosis of Pediatric Acute Osteomyelitis. Med Sci Monit. 2017;23:5211-7.
96. 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(9):1289-93.
97. Caird MS, Flynn JM, Leung YL, Millman JE, D'Italia JG, Dormans JP. Factors distinguishing septic arthritis from transient synovitis of the hip in children - A prospective study. Journal of Bone and Joint Surgery-American Volume. 2006;88A(6):1251-7.
98. Takwoingi Y, Riley RD, Deeks JJ. Meta-analysis of diagnostic accuracy studies in mental health. Evid Based Ment Health. 2015;18(4):103-9.
99. Lee J, Kim KW, Choi SH, Huh J, Park SH. Systematic Review and Meta-Analysis of Studies Evaluating Diagnostic Test Accuracy: A Practical Review for Clinical Researchers-Part II. Statistical Methods of Meta-Analysis. Korean J Radiol. 2015;16(6):1188-96.

100. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM, Cochrane Diagnostic Test Accuracy Working G. Systematic reviews of diagnostic test accuracy. Ann Intern Med. 2008;149(12):889-97.
101. Carpenter CR, Schuur JD, Everett WW, Pines JM. Evidence-based diagnostics: adult septic arthritis. Acad Emerg Med. 2011;18(8):781-96.
102. Fioretto JR, Borin FC, Bonatto RC, Ricchetti SM, Kurokawa CS, de Moraes M, et al. Procalcitonin in children with sepsis and septic shock. J Pediatr (Rio J). 2007;83(4):323-8.
103. Pontrelli G, De Crescenzo F, Buzzetti R, Jenkner A, Balduzzi S, Calo Carducci F, et al. Accuracy of serum procalcitonin for the diagnosis of sepsis in neonates and children with systemic inflammatory syndrome: a meta-analysis. BMC Infect Dis. 2017;17(1):302.
104. Liu D, Su L, Han G, Yan P, Xie L. Prognostic Value of Procalcitonin in Adult Patients with Sepsis: A Systematic Review and Meta-Analysis. PLoS One. 2015;10(6):e0129450.
105. Takwoingi Y, Leeflang MM, Deeks JJ. Empirical evidence of the importance of comparative studies of diagnostic test accuracy. Ann Intern Med. 2013;158(7):544-54.
106. Takwoingi Y, Partlett C, Riley RD, Hyde C, Deeks JJ. Methods and reporting of systematic reviews of comparative accuracy were deficient: a methodological survey and proposed guidance. J Clin Epidemiol. 2020;121:1-14.
107. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. Open Med. 2009;3(3):e123-30.
108. Preston L, Carroll C, Gardois P, Paisley S, Kaltenthaler E. Improving search efficiency for systematic reviews of diagnostic test accuracy: an exploratory study to assess the viability of limiting to MEDLINE, EMBASE and reference checking. Syst Rev. 2015;4:82.
109. Leeflang MM, Scholten RJ, Rutjes AW, Reitsma JB, Bossuyt PM. Use of methodological search filters to identify diagnostic accuracy studies can lead to the omission of relevant studies. J Clin Epidemiol. 2006;59(3):234-40.
110. Rutjes AW, Reitsma JB, Vandenbroucke JP, Glas AS, Bossuyt PM. Case-control and two-gate designs in diagnostic accuracy studies. Clin Chem. 2005;51(8):1335-41.
111. Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M, et al. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. Int J Technol Assess Health Care. 2012;28(2):138-44.
112. Di Bitetti MS, Ferreras JA. Publish (in English) or perish: The effect on citation rate of using languages other than English in scientific publications. Ambio. 2017;46(1):121-7.
113. Lin L. Bias caused by sampling error in meta-analysis with small sample sizes. PLoS One. 2018;13(9):e0204056.
114. 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(9):882-93.
115. Sotiriadis A, Papatheodorou SI, Martins WP. Synthesizing Evidence from Diagnostic Accuracy TEsts: the SEDATE guideline. Ultrasound Obstet Gynecol. 2016;47(3):386-95.
116. Hsu J, Brozek JL, Terracciano L, Kreis J, Compalati E, Stein AT, et al. Application of GRADE: making evidence-based recommendations about diagnostic tests in clinical practice guidelines. Implement Sci. 2011;6:62.
117. Gopalakrishna G, Mustafa RA, Davenport C, Scholten RJ, Hyde C, Brozek J, et al. Applying Grading of Recommendations Assessment, Development and Evaluation (GRADE) to diagnostic tests was challenging but doable. J Clin Epidemiol. 2014;67(7):760-8.
118. Farid-Kapadia M, Joachim KC, Balasingham C, Clyburne-Sherin A, Offringa M. Are child-centric aspects in newborn and child health systematic review and meta-analysis protocols and reports adequately reported?-two systematic reviews. Syst Rev. 2017;6(1):31.
119. Kapadia MZ, Askie L, Hartling L, Contopoulos-Ioannidis D, Bhutta ZA, Soll R, et al. PRISMA-Children (C) and PRISMA-Protocol for Children (P-C) Extensions: a study protocol for the development of guidelines for the conduct and reporting of systematic reviews and meta-analyses of newborn and child health research. BMJ Open. 2016;6(4):e010270.
120. Salameh JP, Bossuyt PM, McGrath TA, Thombs BD, Hyde CJ, Macaskill P, et al. Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA): explanation, elaboration, and checklist. BMJ. 2020;370:m2632.
121. Bujang MA, Adnan TH. Requirements for Minimum Sample Size for Sensitivity and Specificity Analysis. J Clin Diagn Res. 2016;10(10):YE01-YE6.
122. Turner RM, Bird SM, Higgins JP. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. PLoS One. 2013;8(3):e59202.
123. Jackson D, Turner R. Power analysis for random-effects meta-analysis. Res Synth Methods. 2017;8(3):290-302.

124. Korevaar DA, Gopalakrishna G, Cohen JF, Bossuyt PM. Targeted test evaluation: a framework for designing diagnostic accuracy studies with clear study hypotheses. Diagn Progn Res. 2019;3:22.
125. Ochodo EA, de Haan MC, Reitsma JB, Hooft L, Bossuyt PM, Leeflang MM. Overinterpretation and misreporting of diagnostic accuracy studies: evidence of "spin". Radiology. 2013;267(2):581-8.
126. Mathes T, Pieper D. An algorithm for the classification of study designs to assess diagnostic, prognostic and predictive test accuracy in systematic reviews. Syst Rev. 2019;8(1):226.
127. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. 2015;351:h5527.
128. Yang B, Vali Y, Dehmoobad Sharifabadi A, Harris IM, Beese S, Davenport C, et al. Risk of bias assessment of test comparisons was uncommon in comparative accuracy systematic reviews: an overview of reviews. J Clin Epidemiol. 2020.
129. Menten J, Lesaffre E. A general framework for comparative Bayesian meta-analysis of diagnostic studies. BMC Med Res Methodol. 2015;15:70.
130. Leeflang MMG, Reitsma JB. Systematic reviews and meta-analyses addressing comparative test accuracy questions. Diagn Progn Res. 2018;2:17.
131. Arevalo-Rodriguez I, Tricco AC, Steingart KR, Nussbaumer-Streit B, Kaunelis D, Alonso-Coello P, et al. Challenges of rapid reviews for diagnostic test accuracy questions: a protocol for an international survey and expert consultation. Diagn Progn Res. 2019;3:7.
132. Andersen MZ, Gulen S, Fonnes S, Andresen K, Rosenberg J. Half of Cochrane reviews were published more than 2 years after the protocol. J Clin Epidemiol. 2020;124:85-93.
133. Dinnes J, Deeks J, Kirby J, Roderick P. A methodological review of how heterogeneity has been examined in systematic reviews of diagnostic test accuracy. Health Technol Assess. 2005;9(12):1-113, iii.