Comparative value of erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) testing in combination versus individually for the diagnosis of undifferentiated patients with suspected inflammatory disease or serious infection: a systematic review and economic analysis

DATE: 25 March , 2015

Report authors:

Nazila Assasi MD PhD[1,2]
Gord Blackhouse MSc[1,2]
Kaitryn Campbell MLIS[1,2]
Robert Hopkins[1,2]

[1]Programs for Assessment of Technology in Health (PATH) Research Institute, St. Joseph's Healthcare Hamilton, Ontario, Canada

[2]Department of Clinical Epidemiology & Biostatistics, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada

Contents

1 Introduction

1.1 Context and Policy Issues

Patients who have suspected or are being monitored for inflammatory disease or serious infection may undergo a diagnostic workup that involves multiple laboratory tests. Two such laboratory tests are erythrocyte sedimentation rate (ESR) and c-reactive protein (CRP). These tests are nonspecific blood tests that are used widely to aid the diagnosis of numerous conditions, and ESR and CRP are both well-established tests that are frequently ordered together. The simultaneous and widespread use of both of these tests has raised concerns about the potential overutilization of these tests, particularly if they are providing little valuable information regarding the patient management and outcomes.

The aim of this report is to address the issue of when, if ever, it is appropriate to concurrently test ESR and CRP (as opposed to testing only ESR or CRP) to help diagnose inflammatory disease or serious infection. Accordingly, research questions (see Section 2) were developed to elucidate the added value associated with performing both tests rather than one test.

1.2 Background Information

CRP and ESR are among the most widely used diagnostic tests in detecting and monitoring inflammatory conditions that may be caused by infection, autoimmune disorders, malignancies, or tissue necrosis.[1] CRP is a test that measures the level of a plasma protein (C-reactive protein) that is produced by liver cells in response to acute inflammation or infection. Unlike CRP which is a direct measure of the inflammation, ESR indirectly measures the level of inflammation in the body. This test measures the rate at which red blood cells settle in a specially marked tube of anti-coagulated blood.

Both CRP and ESR are usually increased in acute inflammatory conditions. However, patterns of response are different for each test. CRP rises within hours of onset of an infection of inflammatory condition and returns back to normal, within 3-7 days, if the acute process is resolved. ESR, on the other hand, increases in a slower manner and remains elevated for a longer period of time.[2] In addition, ESR is a non-specific measure that can be affected by other factors than inflammation such as the size, shape and number of red blood cells, levels of serum fibrinogen and immunoglobulins, renal function, age and sex, pregnancy, or use of medications.[2,3] Due to the above-mentioned differences, CRP is generally preferred in the literature over ESR in the assessment of early inflammation.[2,4-6] However, there is no consensus on the preferential use of one of these tests;[3,7] and as a result, physicians often request both tests. Based on the statistics provided by Alberta Health Services, of the 650,000 requests of ESR and CRP tests in Alberta (Canada) in 2013, 45% ordered both ESR and CRP.[8]

There is scarce data to support the simultaneous use of CRP and ESR. In addition, it has been shown that, when conducted simultaneously, CRP and ESR are likely to yield concordant results in 67 to- 81% of the time.[9-11] Data from a recent study in an academic tertiary care children's hospital in the United States (US) suggested that elimination of concurrent CRP and ESR testing in both pediatric and adult practice could result in cost savings of approximately $250 000 to $400 000 per year, to the hospital[10]

2 Research Questions

What is the comparative diagnostic performance of ESR+CRP versus ESR or CRP (quantified as, sensitivity, specificity, positive and negative predictive values, and overall accuracy) in different conditions?
What is the comparative cost for ESR+CRP versus ESR or CRP (differential cost versus differential outcome) for different conditions?

3 Methods

3.1 Literature Search Strategy

The literature search was performed by an information specialist using a peer-reviewed search strategy.

Published literature was identified by searching the following bibliographic databases: PubMed; and The Cochrane Library (2014, Issue 6) via Wiley. The search strategy was comprised of both controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords. The main search concepts were erythrocyte sedimentation rate and C-reactive protein.

No methodological filters were applied. Where possible, retrieval was limited to the human population. Retrieval was limited to items published between January 2004 and June 11, 2014 and the English language. Editorials, comments, letters, and newspaper articles were excluded from the search results. See Appendix 1 for the detailed search strategy.

The initial search was completed on June 11, 2014. Regular alerts were established to update the search until the publication of the final report. Regular search updates were performed on databases that do not provide alert services.

In order to complement our initial search, a second search of the same databases was performed on January 15, 2015. Retrieval was limited to items published between January 1980 and January 15, 2015 and the English language.

Grey literature (literature that is not commercially published) was identified by searching the Grey Matters checklist (http://www.cadth.ca/resources/grey-matters), which includes the websites of regulatory agencies, health technology assessment agencies, clinical guideline repositories, and professional associations. Google and other Internet search engines were used to search for additional web-based materials. These searches were supplemented by reviewing the bibliographies of key papers and through contacts with appropriate experts and industry. See Appendix 1 for more information on the grey literature search strategy.

3.2 Selection Criteria and Methods

Studies suitable for inclusion were selected from those identified through the literature search, using the criteria listed below.

Population	Any undifferentiated population not being tested to monitor an existing condition

Intervention	ESR and CRP used in combination
Comparators	ESR or CRP alone
Outcomes	-Diagnostic Test Performance: Sensitivity Specificity Positive likelihood ratio Negative likelihood ratio Area Under the Receiver Operating Characteristic (ROC) Curve (AUC) Positive predictive value Negative predictive value Rates of false positive tests Rates of false negative tests Overall diagnostic accuracy -Cost per outcome unit -Cost per quality-adjusted life year (QALY) -Incremental cost-effectiveness ratio (ICER)
Study types	RCTs, prospective or retrospective observational (non-randomized) studies (cross-sectional diagnostic studies, cohort, case-control), economic evaluations

Two reviewers independently screened the titles and abstracts for relevance using a predefined checklist (Appendix 2). Any discrepancies between reviewers were resolved by consensus. Full texts of relevant titles and abstracts were retrieved, and assessed by two independent reviewers to make inclusion and exclusion decisions, using explicit pre-determined criteria (Appendix 3). Discrepancies between the reviewers were resolved by consensus, consulting a third reviewer when necessary.

3.3 Exclusion Criteria

Studies were excluded if they: did not include a defined population/condition; included tests that use the Wintrobe method for ESR; performed tests outside a central laboratory (e.g., in a physician's office); or utilized point-of-care methodology for performing the tests of interest.

3.4 Critical Appraisal of Individual Studies

3.4.1 Clinical Review

The methodological quality of the included diagnostic studies was assessed using the Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2; Appendix 4).[12] The QUADAS-2 is a tool that evaluates the risk of bias in a selection of patients, index test, reference standard, and flow and timing of the study. The tool also addresses concerns about the applicability of tests and provides signaling questions to help identify potential biases.

Data from all included studies were extracted into pre-defined data extraction forms (Appendix 5). Relevant data were directly extracted from the text or tables. The data extraction forms were piloted by the reviewers, a priori, and a calibration exercise using data from 25% of studies was

undertaken to ensure consistency between the reviewers. The data extraction was performed by two independent reviewers. Any disagreements in data extraction were discussed and resolved by consensus.

3.4.2 Economic Review

The Drummond checklist was used to assess the methodologic quality of economic reviews.

3.5 Data Analyses and Synthesis

3.5.1 Clinical Review

Outcomes

Statistical outcomes that provided tests of differences in diagnostic test performance between types of tests included sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, AUC, positive predictive value, and negative predictive value. A definition of each outcome measure is provided below. More details on how each of these methods was derived are provided in Appendix 6.

The two-by-two contingency table comparing an indext test with a reference standard

	Reference test	
Index test	Positive	Negative
Positive	TP	FP
Negative	FN	TN

TP= True positives; when the positive index test agrees with the positive reference standard;
FP= False positives; when the positive index test disagrees with the negative reference standard;
FN= False negatives; when the negative index test disagrees with the positive reference standard;
TN= True negatives; when the negative index test agrees with the negative reference standard

Table 1- Definitions of diagnostic accuracy measures

Measure	Definition	Formula
Sensitivity	The proportion of persons with the disease who are correctly identified by a test	TP/ (TP/FN)
Specificity	The proportion of persons without a disease who are correctly identified by a test	TN/(TN+FP)
Positive Predictive Value	the proportion of patients with positive test results who are correctly diagnosed	TP/(TP+FP)
Negative Predictive Value	proportion of patients with negative test results who are correctly diagnosed	TN/(TN+FN):
Positive Likelihood Ratio	Indicates how much more likely it is to get a positive test in the diseased as opposed to the non-diseased group	sensitivity/(1-specificity)
Negative Likelihood Ratio	Indicates how much more likely it is to get a negative test in the non-diseased as opposed to the diseased group	(1-sensitivity)/specificity
AUC	The probability that a randomly chosen diseased subject is correctly diagnosed with greater suspicion than a randomly chosen non-diseased subject.	A plot of the sensitivity versus (1-specificity), where different points on the curve correspond to different test thresholds
Overall Diagnostic Accuracy	Proportion of correctly classified subjects among all subjects	(TP/TN)/(TP+FP+TN+FN)

AUC= Area under the Receiver Operating Characteristic (ROC) Curve

In this report we use measures of sensitivity, specificity and the global measure of overall diagnostic accuracy, for the purpose of reporting differences in diagnostic performance of the combined ESR and CRP testing versus either of the tests alone.

Comparisons

Each of the diagnostic performance measures were estimated for the comparison between ESR, CRP, and a combination of these two tests. The focus of the comparisons was:

- Combined ESR and CRP testing versus ESR alone
- Combined ESR and CRP testing versus CRP alone

Since different definitions are used for the combination of ESR and CRP tests, we considered two different situations when a combined test was defined as positive: (1) if both ESR and CRP tests were positive (referred to as ESR+CRP from hereon); and (2) if either ESR or CRP were positive (referred to as ESR/CRP from hereon).

Direct and Indirect Comparisons

The analysis of the diagnostic test performance involves two steps. In the first step, the direct comparison between each of the four comparators, i.e. ESR+CRP versus reference standard, ESR/CRP versus reference standard, ESR alone versus reference standard, and CRP alone versus reference standard was estimated for each study. When there was more than one study with the same disease condition, the results of multiple studies of the same test for a particular condition (e.g. ESR for the diagnosis of periprosthetic infection) were pooled to create one pooled estimate. A random-effect meta-analysis was not possible, since this method requires more than four different studies reporting the same outcome to provide sufficiently diverse data that is needed for a statistical convergence in STATA. Since no more than four similar studies (in terms of tests, condition, and outcome) were identified by the review, we used a fixed effects analysis with a simple sum of the elements in the two-by-two tables.[13]

In order to provide the relative performance of combined ESR and CRP testing i.e., ESR+CRP or ESR/CRP, versus individual ESR or CRP testing, indirect comparisons were conducted to provide a comparative estimate between the two tests. The indirect comparison was estimated with the publically available indirect treatment comparison software (http://www.cadth.ca/index.php/en/itc-user-guide) developed for CADTH by Wells et al. (2009)[14] and based on the Bucher method of indirect comparisons.[15] The Bucher method allows indirect comparisons between relative test performances if the patient populations are similar. Therefore, we provided an indirect comparison across studies with the same disease condition. All of the included studies used all study tests, i.e., ESR and CRP individually and in combination, on a common set of patients and the results were dependent on patients selected.

One caveat with this analysis of the pair-wise estimates such as sensitivity was that the estimation was conducted under the assumption of normality which creates confidence intervals not bounded by one, which was observed in the data. In particular, the confidence intervals for sensitivity or specificity within every study are always less than or equal to the upper bound of 1.0 (or 100%). In particular, sensitivity and specificity can never be higher than 100%, and the confidence intervals can never be higher than 100%, because of the use of binomial confidence intervals. However, when we conducted the indirect comparison, the relative sensitivity or relative specificity can be higher or lower than 1.0 (or 100%) and the relative diagnostic test performance should be interpreted as a ratio of estimates (relative sensitivity, relative specificity, relative overall diagnostic accuracy (ODA), etc.).[16] For example, if the first test had 90% sensitivity and the second test had 80% sensitivity, then the relative sensitivity is 0.90/0.80= 1.125.

Missing Data

For studies that did not report all of the statistical parameters and confidence intervals, wherever possible the missing parameters and confidence intervals were derived with available information. Specifically, not all studies reported the elements of the two-by-two contingency table, i.e., number of true positive, false positive, true negative and false negative. Unfortunately, these latter values are required for meta-analysis of diagnostic accuracy studies. To derive the missing information, we relied on the assumption that we can recreate the elements of the two-by-two table by using available information, described below.

In most cases, the sensitivity and specificity were provided without confidence intervals. When the confidence intervals were not provided, we can, based on the total study sample size, iteratively estimate the unique two-by-two table that would create the study's sensitivity and specificity. Occasionally other outcome measures such as predictive values were provided, again without confidence intervals, and we used this information to provide verification of the unique two-by-two table. Specifically, for a given study size, there is one unique set of 2x2 contingency table values that will create the sensitivity and specificity. The unique 2x2 contingency table will also create other diagnostic test performance estimates such as predictive values or likelihood ratios. If these latter values were reported in the published papers, we could confirm that the unique 2x2 contingency table crated these latter values.

When confidence intervals were provided, we assumed that the confidence intervals were derived with binomial approximation methods, which is the most common statistical distribution. From the available estimate and confidence interval, we iteratively estimated the unique number of true positives and false negatives to recreate the confidence interval. After a similar exercise for the non-disease cases, the numbers of true negatives and false positives were derived.[17] With the derived two-by-two table estimates, the estimates and confidence intervals were recreated to ensure approximate consistency, as well as being verified with other estimates such as positive predictive values. Studies that reported estimates with insufficient decimal places, e.g., sensitivity =0 94 instead of 94.4%, led to the creation of a range of possible values in the two-by-two table, which can lead to inconsistencies between the derived 2x2 table and that in the original study. To minimize such discrepancies, the mean estimated values of the two-by-two estimates were used.

3.5.2 Economic Review

Only one economic study was identified in the search, therefore no pooling of results was completed. A narrative review of this economic evaluation was conducted.

4 Summary of Clinical Evidence

4.1 Quantity of Research Available

A total of 3,236 potential citations were identified by the clinical search, with 3,184 citations being excluded during the title and abstract review based on irrelevance to the questions of interest. The full text documents of the remaining 52 articles were retrieved. Of those 52 articles, 41 did not meet the eligibility criteria and were excluded and one reported cost-effectiveness data, leaving 10 articles for the clinical review.[18-25] Our review identified one relevant systematic review that studied the value of ESR and CRP tests in diagnosis of inflammatory bowel diseases.[26] Of the 24 studies included in the aforementioned review, only one[24] evaluated the combination of ESR and CRP tests and compared the results to those of ESR and CRP alone. This study had already been identified and included in our review. A PRISMA diagram demonstrating the study selection process is presented in Appendix 7. A list of included and excluded studies is provided in Appendix 8.

4.2 Summary of Study Characteristics

A summary of individual study characteristics is presented in Appendix 9. As the appendix shows, four of the included articles reported on the diagnostic performance of the study tests in diagnosis of periprosthetic infections after total hip or knee arthroplasty procedures in adults.[18,21,23,25] Two studies reported on the diagnostic performance of ESR and CRP, individually and in combination, in pediatric orthopedic infections;[20,22] one study in pediatric bronchiolitis,[27] one study in inflammatory bowel diseases;[24] and two studies in giant cell arteritis.[19,28] The included studies originated from the following countries: the US (6 studies),[18,19,21,23,25,28] the United Kingdom (2 studies),[20,24] Poland (one study),[27] and Finland (one study).[22]

Eight of the included studies were observational diagnostic accuracy studies; of which, five used a prospective data collection approach,[18,21-23,27] while three reviewed medical charts retrospectively.[19,20,25] Two studies used a case–control design. In one of these two studies a group of undifferentiated patients (suspected inflammatory bowel disorder) who underwent both ESR and CRP tests along with the study's reference standard (barium flow) were compared with disease positive and healthy control groups.[24] However, it was not clear how the data from the control groups were used in the analysis. Healthy controls in this study seemed not to undergo ESR and CRP testing. The second study compared ESR and CRP test results from the patients with a positive reference standard test (temporal artery biopsy) with those of the healthy control group.[28] One of the prospective diagnostic studies, which included children with suspected bronchiolitis, classified study participants to two subgroups of children with symptoms of viral and those with symptoms of bacterial respiratory infections, in order to perform an internal case-control type analysis.[27] Sample sizes varied across the studies, ranging from 63[24] to 764[19] participants, mostly derived from individual academic hospitals. One study enrolled patients from multiple referral centers;[22] while the remaining seven studies were conducted in a single medical center.

In the selected studies, different definitions were used for the combination of ESR and CRP tests. Three studies[18,21,25] reported on both ESR+CRP and ES/CRP test combinations, while four other studies considered only ESR+CRP as the test combination option.[19,20,24,28] The definition of a combined test was not clear in the remaining three studies.[22,23,27] However, a closer examination of the results of the two of these studies revealed that what had been

considered as a combined test in both studies had a higher sensitivity and lower specificity than either of the individual tests.[23 22] This likely reflects a positive results if either ESR or CRP is positive, i.e., the ESR/CRP combination. Accordingly, the data on the combined test from these two studies were assumed to reflect the ESR/CRP and were analyzed as such. The definition of the combined test in the third study remained unclear.[27]

4.3 Summary of Critical Appraisal

4.3.1 Risk of bias

Appendix 10 summarizes the results of the QUADAS-2 assessments. As shown in this appendix, three of the studies were rated as being in a high risk of selection bias because they used a case-control design,[24,28] or applied extensive exclusion criteria.[22,24] In seven studies[18-21,23,25,27] it was unclear whether their recruitment approach or exclusion criteria could have introduced any selection bias; five of which included patients only if they had both ESR and CRP test results;[18-21,23] five studies excluded patients who were at a higher[22] or lower[18] risk of having the study outcome, or those with concurrent conditions that might have affected their ESR or CRP levels;[23,25,27] three used a retrospective data collection approach,[19,20,25] and one recruited hospitalized patients under the age of 24-months.[27]The possibility of low generalizability of findings due to the study's sampling or site selection approach was discussed by the authors in two studies.[19,22]

None of the studies reported blinding of the results of ESR and CRP at the time of determining the final diagnosis, or provided specific information to assess whether the knowledge of the reference standard might have influenced interpretation of index test results. . Only one of the studies[19] mentioned the time interval between ESR and CRP tests and the reference standard. One study[20] was considered to be at a high risk of bias for index test and reference standard domains, due to not using pre-specified cut-off-points for index tests (i.e. ESR and CRP) and using an imperfect reference standard (i.e. history of relevant clinical presentation, referral to an orthopedic team and antibiotic-therapy) for verification of final diagnosis. It was unclear if the reference standard used in one of the studies could correctly classify the target condition.[19] In this study, positive giant cell artery biopsy results were used for confirmation of diagnosis of giant cell arteritis. The authors noted in their discussion that the possibility of misclassification of biopsy-negative giant cell arteritis patients as disease-free could not be precluded in their study.

Three studies were assessed as being at unclear risk of bias for the flow and timing domain.[22,24,27] The study by Paakkonen et al[22] seemed to combine microbiological culture result from joint and bone aspirations with different imaging modalities, such as ultrasound, radiographs or magnetic resonance imaging, to make the final diagnosis of various osteoarticular infections. The schedules of measuring ESR and CRP were reported to be slightly different in this study. Grzesk et al[27] used chest x-ray findings along with white blood cell count, bacteriologic test and clinical symptoms as the reference standard. However, their description of study methodology implied that chest x-ray and bacteriologic examination had not been performed for all study participants. The study by Hayreh et al[28] was considered to be at a high risk of bias for the flow and timing domain. In this study a subset of patients who were recruited prior to 1985 did not receive CRP test. Therefore, diagnostic performance of ESR was estimated using data from all patients; whereas diagnostic performance measures for CRP and combined ESR and CRP testing were estimated in a smaller set of study participants. The timing of the ESR and CRP tests was not clearly indicated in the case-control diagnostic study by Dolwani et al.[24]

4.3.2 Applicability

Overall, applicability concerns related to the patient selection, index test and reference standard domains were low for the majority of the included studies. In order to be rated as applicable to the research question, the studies evaluated using QUADAS2 tool should have included the same patient population, same index tests (ESR and CRP), and same reference standard as were defined in our study questions. Since the objective of our study was to compare a combination of ESR and CRP with either ESR or CRP for diagnosis of severe infectious and inflammatory disorders, we decided that any reference standard which could provide a basis for accurate diagnosis would be applicable. One study,[20] which used retrospectively collected data on a combination of clinical symptoms, referral history and antibiotic-therapy to verify the diagnosis of pediatric bone and joint infections was classified as raising some levels of applicability concern in terms of reference standard test. It was also unclear if the exclusion of culture (reference-standard) negative patients in the study by Paakkonen et al,[22] and limiting inclusion criteria to the age group of 0-24 months in the study by Grzesk et al[27] could raise any applicability concerns in terms of patient selection. The possibility of low generalizability of findings due to the study's sampling or site selection approach was noted by the authors in two studies.[19,22] In addition, the generalizability of study results could be limited due unusually high prevalence of disease, in studies by Costa et al[18] and Johnson et al,[21] which can be ascribed to the referral nature of the study settings.

4.4 Summary of Findings

4.4.1 Diagnostic performance of ESR and CRP, when performed individually and in combination, for diagnosis of different conditions

An overview of the diagnostic performance of ESR, CRP, ESR+CRP and ESR/CRP testing from the available studies is presented in Appendices 11 and 12. Appendix 11 provides the details of the study index tests (ESR and/or CRP), the reference standard used for the confirmation of diagnosis, the relative frequency of disease-positive individuals diagnosed by the reference standard, and the number of truly or falsely diagnosed patients by the index tests of interest. Appendix 12 shows the diagnostic performance measures for the tests on interest when compared with the reference standard employed in each study (direct comparisons). The included studies are categorized based on the studies' target conditions into four categories: periprosthetic infections; pediatric orthopedic infections; inflammatory bowel diseases; and giant cell arteritis.

4.4.1.1 Periprosthetic infections

Four studies used ESR (at the threshold of >30 millimeter per hour (mm/h)) and CRP (at the threshold of >10 milligram per liter (mg/L)), individually and in combination, for the diagnosis of periprosthetic infections.[18,21,23,25] Two of these studies used a single criterion of a positive synovial fluid bacterial culture as the reference standard,[18,23] while the other two employed similar multiple criteria diagnostic tools to confirm the diagnosis (see Appendix 11 for details). The prevalence of periprosthetic infection was considerably higher in the studies by Costa et al (89.61%)[18] and Johnson (84.07%)[21] than the two other studies (39.19%[23] and 29.50%[25]). As it is shown in Appendix 12, among the four studies, the sensitivity and specificity of ESR ranged from 0.89[18] to 0.94[25] and from 0.33[21] to 0.72,[23] respectively. The sensitivity of CRP varied between 0.91[25] and 0.95[21] and its specificity between 0.20[21] to 0.77.[25] Data on performance of ESR+CRP in periprosthetic infections were available from three studies,[18,21,25] which reported sensitivity values for ESR+CRP ranging from 0.84[18] to 0.89[21] and specificity values ranging from

0.29[21] to 0.75.[18] The sensitivity and specificity of the ESR/CRP for periprosthetic infections ranged from 0.96[23] to 0.98,[25] and from 0.20[18] to 0.59[25], respectively.

The results of direct pooled analyses of the tests of interest versus the reference standard employed in each study are provided in Table 2. Among the four types of tests, the highest sensitivity for the diagnosis of periprosthetic infections was found for ESR/CRP (0.96 (95% CI 0.94, 0.98)) and the highest specificity was found for ESR+CRP (0.85 (95% CI 0.81, 0.89)). The pooled analysis also showed that, when compared to the reference standard, the ODA of ESR+CRP (0.86 (95% CI 0.83, 0.89)) was slightly higher than that of ESR/CRP (0.82 (95% CI 0.80, 0.85)) and CRP (0.82 (95% CI 0.79, 0.84)); and ESR had the lowest overall accuracy (0.79 (95% CI 0.76, 0.81)) among the all four types of tests, in diagnosis of periprosthetic infections.

Table 2- pooled analysis of data for diagnostic performance of esr and crp individually and in combination for the diagnosis of periprosthetic infection

Test	#Studies (ref #)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	LR(+) (95% CI)	LR(-) (95% CI)	AUC (95% CI)	ODA (95% CI)
ESR	4 ([18,21,23,25])	0.92 (0.89, 0.94)	0.70 (0.66, 0.73)	0.69 (0.65, 0.73)	0.92 (0.89, 0.95)	3.01 (2.17, 4.17)	0.12 (0.11, 0.14)	0.806 (0.751, 0.860)	0.79 (0.76, 0.81)
CRP	4 ([18,21,23,25])	0.93 (0.91, 0.96)	0.73 (0.70, 0.77)	0.72 (0.68, 0.76)	0.94 (0.91, 0.96)	3.51 (2.45, 5.04)	0.09 (0.08, 0.11)	0.833 (0.775, 0.891)	0.82 (0.79, 0.84)
ESR+ CRP	3 ([18,21,23,25])	0.87 (0.83, 0.91)	0.85 (0.81, 0.89)	0.82 (0.77, 0.86)	0.90 (0.87, 0.93)	5.79 (4.27, 7.85)	0.15 (0.12, 0.19)	0.861 (0.784, 0.939)	0.86 (0.83, 0.89)
ESR/ CRP	4 ([18,21,25])	0.96 (0.94, 0.98)	0.57 (0.53, 0.61)	0.74 (0.70, 0.78)	0.95 (0.92, 0.97)	3.18 (2.02, 5.02)	0.06 (0.05, 0.07)	0.828 (0.762, 0.894)	0.82 (0.80, 0.85)

Abbreviations: AUC= area under the curve; CRP= C-reactive protein; CI= confidence interval; ESR= erythrocyte sedimentation rate; LR(+)= positive likelihood ratio; LR(-)= negative likelihood ratio; NPV= negative predictive value; ODA=overall diagnostic accuracy; PPV= positive predictive value.

4.4.1.2 Pediatric orthopedic infections

Data on the diagnostic performance of the study tests in pediatric orthopedic infections were available from two studies.[20,22] However, due to the paucity of data and significant differences between the two studies, in terms of study population, index test cut-off-points and the reference standard, pooled analysis was not conducted for this subgroup of the included studies.

Robinson et al[20] retrospectively included children under 13 years of age who presented with atraumatic limb pain at the study hospital's emergency room. The authors presented the diagnostic performance results for ESR (at the threshold of >12 mm/h), CRP (at the threshold of >7 mg/L) and ESR+CRP against the reference standard that was defined as a documented history of referral to an orthopedic team, diagnosis of infection and antibiotic-therapy. In this study ESR was reported to have the highest sensitivity 0.88 (95% CI 0.73, 1.04) and ESR+CRP to have the highest specificity 0.90 (95% CI 0.86, 0.94) among the three study tests (Appendix 12). The test thresholds in this study were determined based on the ROC curve analysis.

Paakkonen et al[22] prospectively recruited non-immune deficient children between the age of 3 months and 15 years, who were referred to multiple tertiary care centers in Finland with a suspected acute joint or bone infection. The diagnostic threshold was considered to be greater than 20mm/h for ESR, and greater than 20 mg/L for CRP. A positive bone or joint bacterial culture was the main diagnostic criteria (reference standard) used in this study. However, patients were excluded if they had a negative culture, which allowed only sensitivity values to be

reported. This study demonstrated sensitivities of 0.94 (95% CI 0.90, 0.96), 0.95 (95% CI 0.90, 0.97), and 0.98 (95% CI 0.96, 0.99), for ESR, CRP and ESR/CRP respectively (Appendix 12).

4.4.1.3 Inflammatory bowel diseases

Dolwani et al[24] reported on the diagnostic performance of ESR, CRP and ESR+CRP for diagnosis of inflammatory bowel conditions. This study was conducted in a mixed in-patient/out-patient population consisting of 63 unclassified/undifferentiated patients presenting with symptoms of inflammatory bowel disease, 25 patients with known active Crohn's disease and 25 healthy controls. The study findings, however, seem to compare the diagnostic performance of ESR, CRP and ESR+CRP with that of small bowel barium follow-through imaging (reference standard) in 63 undifferentiated cases. As shown in Appendix 12, ESR had the highest sensitivity (0.79 (95% CI 0.60, 1.00)) and ESR+CRP had the highest specificity (0.84 (95% CI 0.73, 0.94)). However, the overall accuracy of the three types of tests demonstrated little variation (between 0.70 and 0.76) in diagnosing inflammatory bowel diseases.

4.4.1.4 Giant cell arteritis

Two studies evaluated the diagnostic value of ESR, CRP and ESR+CRP for diagnosis of giant cell arteritis.[19,28] However, data was not pooled due to clinical and methodological heterogeneity in the studies. The two studies used different study designs and different thresholds to define positive and negative test results. In one of the studies,[28] there was also intrastudy variability in terms of number and characteristics of patient population included in ESR and ESR+CRP comparison groups. Descriptions of the results from these two studies are provided below.

Kermani et al[19] used ESR at the threshold of greater than 22 mm/h for men and 29 mm/h for women; CRP at the threshold of greater than 8 mg/L; and temporal artery biopsy as the gold standard. This study found sensitivity values of 0.84 (95% CI 0.79, 0.90), 0.86 (95% CI 0.81, 0.92), and 0.81 (95% CI 0.75, 0.87) for ESR, CRP and ESR+CRP, respectively. ESR+CRP was reported to have a higher specificity (0.41 (95% CI 0.37, 0.45)) in detecting giant cell arteritis, as compared to ESR (0.30 (95% CI 0.26, 0.33)), or CRP (0.31 (95% CI 0.27, 0.37)) alone.

The diagnostic case-control study by Hayreh et al[28] used ESR at the threshold of greater than 10 mm/h for men and 20 mm/h for women; CRP at the threshold of greater than 5 mg/L; and temporal artery biopsy as the reference standard. The study included patients with suspected giant cell arteritis for whom temporal artery biopsy had been requested. All patients underwent ESR testing. However, CRP test that became available midway through the study was used for the study participants (cases and controls) who were recruited after this test became available (in 1985). The study found sensitivity values of 0.97 (95% CI 0.92, 0.99), 1.00 (95% CI 0.93, 1.00), and 0.98 (95% CI 0.88, 1.00) for ESR, CRP and ESR+CRP, respectively. Specificity values were 0.67 (95% CI 0.64, 0.72), 0.82 (95% CI 0.74, 0.88), and 0.92 (95% CI 0.86, 0.96) for ESR, CRP and ESR+CRP, respectively.

4.4.1.5 Pediatric bronchiolitis

Grzesk et al[27] studied the performance of ESR and CRP, when performed individually or in combination, in differentiating viral from bacterial bronchiolitis in children with clinical symptoms of bronchiolitis. They used ESR at the threshold of greater than 15 mm/h; CRP at the threshold of greater than 15 mg/L; and chest x-ray findings in combination with WBC count of greater than 12 moles per liter (M/L) and clinical symptoms as the gold standard. This study reported only AUC values, as the diagnostic test performance measure. Lack of data reported on proportions

of truly of falsely diagnosed cases did not allow for any calculations of sensitivity, specify and other diagnostic performance metrics. The authors reported AUCs of 0.71 (95% CI 0.60, 0.83) for ESR, 0.63 (95% CI 0.51, 0.75) for CRP and 0.74 (95% CI 0.60, 0.88) for the combined ESR and CRP testing.

4.4.2 Comparative diagnostic performance of ESR and CRP, when performed individually and in combination for different conditions

The relative diagnostic performance of ESR, CRP, ESR+CRP, and ESR/CRP (indirect comparisons) is determined using the sensitivity and specificity, and ODA values from the included studies as demonstrated in Tables 3-6.

4.4.2.1 Periprosthetic infections

The relative performance measures calculated using pooled data from four studies[18,21,23,25] (Table 3) suggest that, for in detection of periprosthetic infections, ESR+CRP can have statistically higher specificity than both ESR (relative specificity, RSpec)= 1.21 (95% CI 1.13, 1.30) and CRP (RSpec= 1.16 (95% CI 1.09, 1.25)), and a statistically lower sensitivity than CRP (relative sensitivity (RSens= 0.94 (95% CI 0.89, 0.99)), but not ESR (RSens= 0.95 (95% CI 0.90, 1.00)). The sensitivity of ESR/CRP was statistically higher than that of ESR (RSens= 0.94 (95% CI 1.01, 1.08)) alone and could be marginally higher than that of CRP (RSens= 1.03 (95% CI 1.00, 1.07)). The results of our indirect comparisons suggest that, in terms of overall accuracy, ESR+CRP is more accurate than ESR alone (relative ODA= 1.09 (95% CI 1.04, 1.14)) and can be equal to, if not better than, CRP (relative ODA= 1.05 (95% CI 1.00, 1.10)). However, no statistically significant differences were found between ESR/CRP combination and individual results of either of ESR or CRP (Table 3).

Table 3- Relative diagnostic performance of ESR and CRP individually and in combination for diagnosis of periprosthetic infection (4 studies)[18,21,23,25]

Comparison	Relative Sensitivity (95% CI)	P-value	Relative Specificity (95% CI)	P-value	Relative ODA (95% CI)	P-value
ESR+CRP vs ESR	0.95 (0.90, 1.00)	0.063	1.21 (1.13, 1.30)	<0.001	1.09 (1.04, 1.14)	<0.001
ESR+CRP vs CRP	0.94 (0.89, 0.99)	0.027	1.16 (1.09, 1.25)	<0.001	1.05 (1.00, 1.10)	0.050
ESR/CRP vs ESR	1.04 (1.01, 1.08)	0.009	1.00 (0.93, 1.08)	1.000	1.04 (0.99, 1.09)	0.119
ESR/CRP vs CRP	1.03 (1.00, 1.07)	0.050	0.96 (0.89, 1.03)	0.291	1.00 (0.96, 1.04)	1.000

Abbreviations: CI= confidence interval; CRP= C-reactive protein; ESR= erythrocyte sedimentation rate; NA= not available; ODA= overall diagnostic accuracy

4.4.2.2 Pediatric orthopedic infections

The limited evidence from two studies[20,22] in this category showed that, in diagnosis of pediatric orthopedic infections, ESR+CRP was statistically more specific than ESR (RSpec= 1.36 (95% CI 1.23, 1.51)), but not CRP (RSpec=.1.03 (95% CI 0.98, 1.03)). The sensitivity of ESR/CRP combination was statistically higher than that of ESR (RSens= 1.04 (95% CI 1.01, 1.08)) and marginally higher than that of CRP (RSens= 1.03 (95% CI 1.00, 1.07)). Data on specificity and overall accuracy of ESR/CRP was not available. ESR+CRP was shown to have a statistically higher overall accuracy than ESR (relative ODA= 1.31 (95% CI 1.19, 1.45)), but not CRP (relative ODA= 1.02 (95% CI 0.95, 1.10)) (Table 4).

Table 4- Relative diagnostic performance of ESR and CRP individually and in combination for diagnosis of pediatric orthopedic infections (2 study)[20,22]

Comparison	Relative Sensitivity (95% CI)	P-value	Relative Specificity (95% CI)	P-value	Relative ODA (95% CI)	P-value
ESR+CRP vs ESR*	0.74 (0.50, 1.12)	0.132	1.36 (1.23, 1.51)	<0.001	1.31 (1.19, 1.45)	<0.001
ESR+CRP vs CRP*	0.92 (0.55, 1.53)	0.751	1.03 (0.98, 1.09)	0.244	1.02 (0.95, 1.10)	0.585
ESR/CRP vs ESR**	1.04 (1.01, 1.08)	0.009	NA	NA	NA	NA
ESR/CRP vs CRP**	1.03 (1.00, 1.07)	0.050	NA	NA	NA	NA

Abbreviations: CI= confidence interval; CRP= C-reactive protein; ESR= erythrocyte sedimentation rate; NA= not available; ODA= overall diagnostic accuracy
*Data for this comparison was available from the study by Robinson et al (one study).[20]
**Data for this comparison was available from the study by Paakkonen et al (one study).[22]

4.4.2.3 Inflammatory bowel diseases

The relative measures of sensitivity, specificity or ODA for ESR+CRP versus ESR or CRP alone, for detecting inflammatory bowel conditions, were estimated indirectly using the data from a single identified study.[24] Data on diagnostic performance of ESR/CRP was not available from this study. As shown in Table 5, no statistically significant differences were found between ESR+CRP and either of the ESR or CRP tests, in terms of sensitivity, specificity or ODA.

Table 5- Relative diagnostic performance of ESR and CRP individually and in combination for diagnosis of Inflammatory Bowel Disease (1 study)[24]

Comparison	Relative Sensitivity (95% CI)	P-value	Relative Specificity (95% CI)	P-value	Relative ODA (95% CI)	P-value
ESR+CRP vs ESR	0.63 (0.36, 1.13)	0.106	1.25 (0.99, 1.60)	0.061	1.07 (0.88, 1.34)	0.498
ESR+CRP vs CRP	0.65 (0.34, 1.16)	0.193	1.20 (0.96, 1.50)	0.109	1.04 (0.85, 1.28)	0.703
ESR/CRP vs ESR	NA	NA	NA	NA	NA	NA
ESR/CRP vs CRP	NA	NA	NA	NA	NA	NA

Abbreviations: CI= confidence interval; CRP= C-reactive protein; ESR= erythrocyte sedimentation rate; NA= not available; ODA= overall diagnostic accuracy

4.4.2.4 Giant cell arteritis

The limited evidence from the two identified studies was used to indirectly determine relative diagnostic performance of ESR+CRP versus ESR or CRP tests alone. Data on diagnostic performance of ESR/CRP was not available from neither of the studies reporting on giant cell arthritis. In both studies ESR+CRP was statistically more specific than either ESR or CRP alone; and the sensitivity of the combined test was comparable to both individually interpreted ESR and CRP tests. However, due to heterogeneity in terms of study design and the diagnostic thresholds used for a positive test, the relative diagnostic measures will be described separately for these two studies (see tables 6 and 7 for details).

Based on our estimates, in the study by Kermani et al,[19] ESR+CRP was shown to have a statistically higher specificity than both ESR (RSpec= 1.36 (95% CI 1.17, 1.60)) and CRP (RSpec= 1.32 (95% CI 1.14, 1.54)), in diagnosis of giant cell arteritis. The sensitivity of the combined ESR+CRP test was comparable to both individually interpreted ESR and CRP test results (Table 6). In terms of overall diagnostic accuracy, ESR+CRP was found to be statistically more accurate than both ESR (relative ODA= 1.19 (95% CI 1.07, 1.33)) and CRP (relative ODA= 1.16 (95% CI 1.05, 1.29)).

Table 6- Relative diagnostic performance of ESR (>22 mm/h) and CRP(>8 mg/L) individually and in combination for diagnosis of Giant Cell Arteritis (1 study)[19]

Comparison	Relative Sensitivity (95% CI)	P-value	Relative Specificity (95% CI)	P-value	Relative ODA (95% CI)	P-value
ESR+CRP vs ESR	0.96 (0.87, 1.06)	0.416	1.36 (1.17, 1.60)	<0.001	1.19 (1.07, 1.33)	0.001
ESR+CRP vs CRP	0.94 (0.85, 1.04)	0.228	1.32 (1.14, 1.54)	<0.001	1.16 (1.05, 1.29)	0.004
ESR/CRP vs ESR	NA	NA	NA	NA	NA	NA
ESR/CRP vs CRP	NA	NA	NA	NA	NA	NA

Abbreviations: CI= confidence interval; CRP= C-reactive protein; ESR= erythrocyte sedimentation rate; NA= not available; ODA= overall diagnostic accuracy

In the study by Hayreh et al[28] ESR+CRP was shown to have a statistically higher specificity than both ESR (RSpec= 1.37 (95% CI 1.27, 1.49)) and CRP (RSpec= 1.12 (95% CI 1.01, 1.24)), in diagnosis of giant cell arteritis. The sensitivity of the combined ESR+CRP test was comparable to ESR alone (RSens= 1.01 (95% CI 0.94, 1.09)) and to CRP alone (RSens= 0.98 (95% CI 0.91 1.05)) results. In terms of overall diagnostic accuracy, ESR+CRP was found to be statistically more accurate than both ESR (relative ODA= 1.31 (95% CI 1.24, 1.39)) and CRP (relative ODA= 1.08 (95% CI 1.01, 1.26)) (Table 7).

Table 7- Relative diagnostic performance of ESR (>10 mm/h for men and >20 mm/h for women) and CRP (>5 mg/L) individually and in combination for diagnosis of Giant Cell Arteritis (1 study)[28]

Comparison	Relative Sensitivity (95% CI)	P-value	Relative Specificity (95% CI)	P-value	Relative ODA (95% CI)	P-value
ESR+CRP vs ESR	1.01 (0.94, 1.09)	0.395	1.37 (1.27, 1.49)	<0.001	1.31 (1.24, 1.39)	<0.001
ESR+CRP vs CRP	0.98 (0.91, 1.05)	0.588	1.12 (1.01, 1.24)	0.014	1.08 (1.01, 1.26)	0.013
ESR/CRP vs ESR	NA	NA	NA	NA	NA	NA
ESR/CRP vs CRP	NA	NA	NA	NA	NA	NA

Abbreviations: CI= confidence interval; CRP= C-reactive protein; ESR= erythrocyte sedimentation rate; NA= not available; ODA= overall diagnostic accuracy

4.4.2.5 Pediatric bronchiolitis

The study by Grzesk et al[27] compared AUCs for the combined ESR and CRP testing versus either ESR or CRP alone used for differentiation of viral from bacterial bronchiolitis in pediatric population. The study reported p-values for the comparison of combined ESR and CRP testing versus ESR alone and combined testing versus CRP alone to be to be 0.742, and 0.231, respectively. The authors concluded that using the result of a single test, either ESR or CRP, in diagnosis of the type of bronchiolitis and decisions about antibiotic-therapy may be as helpful as using the results of combined ESR and CRP testing.

5 Summary of Economic Evidence

5.1 Quantity of Research Available

One economic evaluation was identified that compared combined ESR and CRP testing to testing of either test alone was identified. This evaluation was a Japanese based[29] study that evaluated the costs and number of "useful results" for a number of laboratory tests.

5.2 Summary of Critical Appraisal

The Drummond checklist[30] was used to critically appraise the one economic evaluation identified. Based on the checklist[30] there were some deficiencies in the study design as neither the perspective of the analysis nor justification of the form of economic evaluation were explicitly stated. As for data collection in the study, the authors do not provide a clear definition of their clinical outcome being evaluated "useful result". The authors did not undertake any sensitivity analysis as part of analysis of results. Additionally, no conclusions on whether combined ESR and CRP testing is cost-effective compared to either test alone was provided.

5.3 Summary of Findings

In a Japanese based study on 177 outpatients, Takemura et al[29] estimated the cost-effectiveness of a number of laboratory tests undertaken for suspected infection and inflammatory conditions. Two of the tests that evaluated were ESR and CRP. The costs for each test were presented in Japanese Yen. The unit of effectiveness for the study was the proportion of tests that produced a "useful result". The authors indicated that a "useful result" was defined a test result that contributed to a change in diagnosis or decision making once the test result was taken into consideration. The determination of a "useful result" was made based on a chart review by a group of three physicians, one of whom was a participant in the initial clinical practice of the patients, which may increase the potential for bias. The cost per ESR and CRP test was reported to be ¥102.8 and ¥340.21 respectively. Costs were reported in 2002 Japanese Yen. The proportion of test results that were found to be "useful" was 0.58 for combined ESR and CRP testing, 0.41 for ESR testing alone and 0.53 for CRP testing alone. The authors report the incremental cost-effectiveness of combined ESR and CRP testing compared to CRP testing alone be ¥1737 per useful result produced. Though not presented in the paper, the incremental cost-effectiveness of combined ESR and CRP testing compared to ESR testing alone can be calculated as ¥2001 per useful result produced. The cost per useful event can be converted from 2002 Japanese Yen to 2015 Canadian dollars by applying the historical 2002 exchange rate[31] (1¥ =$Can 0.012554) and inflating this to 2015 $CAN using the Canadian Consumer Price Index health care component[32] (122.3; January 2015 vs. 2002) Based on this conversion, the incremental cost of combined ESR and CRP testing per useful result compared to CRP alone and to ESR alone was estimated to be $CAN26.67 and $CAN30.72 respectively. The authors did not provide any conclusions around the cost-effectiveness of combined ESR and CRP testing compared to ESR or CRP alone.

Because the only identified published economic study was conducted in Japan its generalizability to the Canadian setting is questionable. Furthermore the outcome of "useful result" was not an outcome included in our clinical review, and its generalizability is questionable given the way in which the data were determined. Therefore a primary economic evaluation was conducted to evaluate the cost effectiveness of combined testing with ESR and CRP compared to either test alone using the diagnostic accuracy data collected in our clinical review.

6 Primary Economic Evaluation

6.1 Methods

Type of evaluation
A cost-effectiveness analysis was conducted to compare combined ESR and CRP testing to either ESR or CRP alone. The cost-effectiveness outcome evaluated depended on whether a positive combined test was based on 1) **both** ESR and CRP being positive (ESR+CRP) or 2) **either** ESR or CRP being positive (ESR/CRP).

If a combined positive test was based on both tests being positive (ESR+CRP), it was assumed that the benefit of a combined test compared to a single test would be to increase specificity and reduce the number of false positive results. Therefore the cost-effectiveness outcome for this definition was the incremental cost per false positive avoided.

If a combined positive test was based on either test being positive (ESR/CRP), it was assumed that the benefit of a combined test compared to a single test would be to increase sensitivity and reduce the number of false negative results. Therefore the cost-effectiveness outcome for this definition was the incremental cost per false negative avoided.

Target population
There were a number of several different target patient populations for this economic analysis. These were based upon the patient populations in which comparative diagnostic accuracy data was were found in the clinical review of this report. Specifically the four target populations of the analyses are:

1) Patients suspected of periprosthetic infections;
2) Children suspected of having orthopedic infections;
3) Patients suspected of having inflammatory bowel diseases; and
4) Patients suspected of having giant cell arteritis.

As described in the clinical evaluation, because of the methodological differences in the two studies that evaluated giant cell arteritis,[19,28] pooling of diagnostic accuracy data was not completed. Therefore two cost-effectiveness analyses were conducted for the giant cell arteritis population. One used the diagnostic accuracy and prevalence data from Kermani et al.[19] The other used diagnostic accuracy and prevalence data from Hayreh et al.[28]

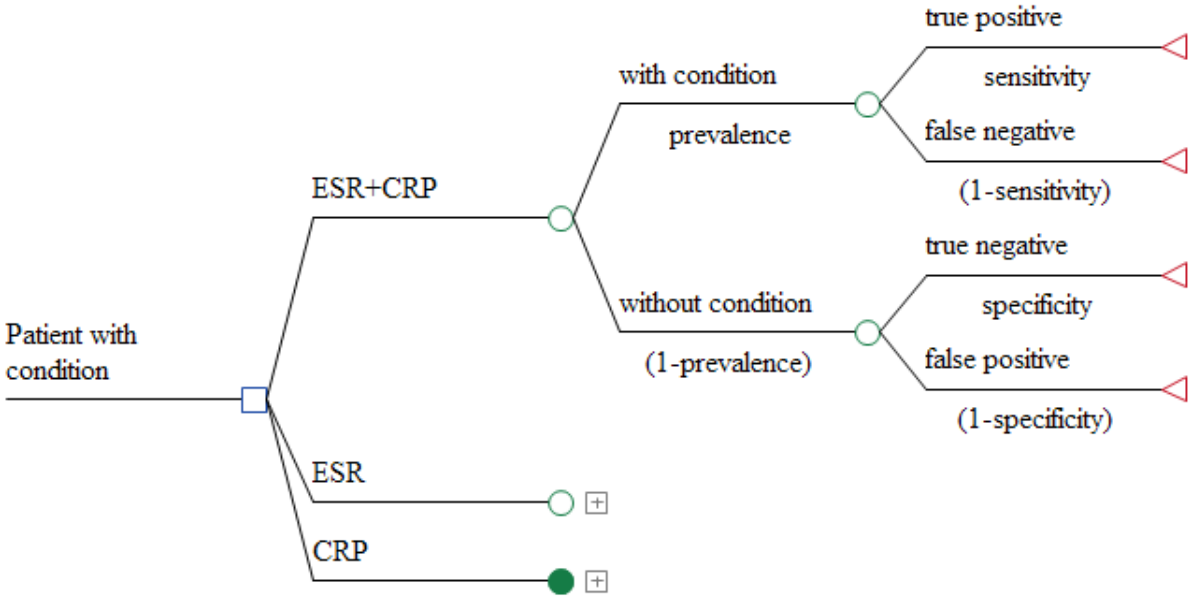
Comparators

The comparators in the evaluation are 1) Combined ESR and CRP testing;2) ESR testing alone; and 3) CRP testing alone. In this analysis, pairwise comparisons of ESR + CRP vs. ESR alone and ESR+CRP vs CRP alone were conducted. Because it was not intended to be part of this review, cost-effectiveness comparisons of ESR alone versus CRP alone were not reported.

Perspective
A third party payer perspective such as that of a provincial ministry of health was undertaken.

Model structure
A graphical representation of the structure of the model is provided in Figure 1. As shown the model starts out with patients being tested for one of the conditions of the four target populations. There is an underlying prevalence of disease that categorizes patients of either having the condition being tested for or not having the condition being tested for. Based on the test results patients having the condition can either be diagnosed correctly as a true positive or diagnosed incorrectly as a false negative. Similarly patients without the condition are either correctly diagnosed as a true negative or incorrectly as a false positive. The diagnostic status (i.e., true positive, false negative, true negative and false positive) of patients is dependent on the prevalence of the disease, along with the sensitivity and specificity of the diagnostic test. As indicated in the diagram, the proportion of false positives can be calculated as 1 minus condition prevalence multiplied by 1 minus test specificity. Similarly the proportion of false negative results is calculated as condition prevalence multiplied by 1 minus the sensitivity of the test.

Figure 1- Graphical representation of the model structure



Clinical model inputs
As shown in Figure 1, in order to estimate the diagnosis status of patients for each testing strategy, the prevalence of the condition along with the sensitivity and specificity of each testing

strategy are required. The prevalence rates used in the model for each of the four populations are provided in Table 8. The prevalence rates are based upon the disease frequencies found in the included studies (see Appendix 11). The prevalence for the periprosthetic infection population was based on the weighted average disease frequency reported in the four studies evaluating this population[18,21,23,25] . For the pediatric orthopedic infections and the inflammatory bowel disease populations, there was only a single study in which disease frequency could be derived. For giant cell arteritis, prevalence rates specific to the studies by Kermani[19] and Hayreh[28] were used in the analysis.

Table 8- Prevalence rates by population

Population	Prevalence
Periprosthetic infection	0.422[18,21,23,25]
Pediatric orthopedic infections	0.066[20]
Inflammatory bowel disease	0.238[24]
Giant Cell Arteritis	
Kermani	0.239[19]
Hayreh	0.292[28]

The sensitivity and specificity for each test strategy was based on the findings in the clinical review (see Appendix 12 and Table 2). Tables 9 and 10 present the sensitivity and specificity for each test by population. Table 2 shows data when a positive combined test is based on both ESR and CRP being positive. Table 3 shows diagnostic accuracy data when a combined test is defined as either ESR or CRP being positive. For the perioperative infection population, sensitivity and specificity are based on the pooled analysis in the clinical review (see Table 2). For the other populations, sensitivity and specificity were based on single studies (see Appendix 12). For giant cell arteritis, separate analyses were conducted based on data from by Kermani et al[19] and Hayreh et al.[28] Data were available for only two populations when combined test is based on either test being positive. For the pediatric infection population, Robinson et al[20] informed a diagnostic accuracy when the combined positive test was defined as both tests being positive, while Paakkonen et al[22] informed data when the combined positive test was defined as when either test was positive.

Table 9- Sensitivity and specificity by patient population when a positive combined test assumes both ESR and CRP are positive (ESR+CRP)

	Periprosthetic infection		Pediatric orthopedic infections		inflammatory bowel disease		Giant Cell Arteritis Kermani et al[19]		Giant Cell Arteritis Hayreh et al[28]	
Test	**sens**	**spec**	**sens**	**spec**	**sens**	**spec**	**sens**	**spec**	**sens**	**spec**
ESR	0.92	0.70	0.88	0.66	0.79	0.67	0.84	0.30	0.97	0.67
CRP	0.93	0.73	0.71	0.87	0.77	0.7	0.86	0.31	1.0	0.82
ESR+CRP	0.87	0.85	0.65	0.90	0.50	0.84	0.81	0.41	0.98	0.92

Abbreviations: CRP= C-reactive protein; ESR= erythrocyte sedimentation rate; sens= sensitivity; spec= specificity

Table 10- Sensitivity and specificity by patient population when a positive combined test assumes either ESR or CRP are positive (ESR/CRP)

	Periprosthetic infection		pediatric orthopedic infections	
Test	**sens**	**spec**	**sens**	**spec**
ESR	0.92	0.70	0.94	NA
CRP	0.93	0.73	0.95	NA
ESR/CRP	0.96	0.57	0.98	NA

Abbreviations: CRP= C-reactive protein; ESR= erythrocyte sedimentation rate; sens= sensitivity; spec= specificity; NA= not available

Resource use and costs

A survey was conducted by CADTH in order to estimate the costs of ESR and CRP testing across Canada. At least one contact was asked to complete the survey in each jurisdiction. In addition to the cost of each test, each contact as asked to provide the various components that were included in the test costs along with utilization data. A summary of the survey results are presented in Table 11. As shown, there was a lot of variation in the reported test costs and the components that were included in the test costs. For example, the cost of CRP was reported to be $$0.67 in New Brunswick and $1.60 in Nunavut. The cost of CRP was reported to be a high $10.31 in British Columbia. Much of this variation may have been due to variation in the components that were included in the cost estimates. For example only the reagent costs were included in the cost of CRP in New Brunswick and Nunavut. Because of this variation, the costs of the tests in the model were assumed to be that reported by British Columbia, where it was reported that all components were included in the CRP and ESR costs.

Therefore in the model the cost of ESR and CRP were assumed to be equal to $10.61 and $10.31 respectively. The cost of the combined test was assumed to be equal $20.92, the sum of the individual ESR and CRP costs.

Table 11. Results from survey on costs and included components of ESR and CRP tests across jurisdictions in Canada

Jurisdiction	ESR	Cost includes	CRP	Cost includes
Manitoba	varies	test, controls and paper	$1.15	Reagents, quality control, calibration
Yukon	$10.72	supplies and tech time	$5.00	Cost of test
British Columbia	**$10.61**	**All cost factors**	**$10.31**	**All cost factors**
Saskatchewan	$1.47	Technologist time, pipet	$3.13	reagents, controls, calibrators
Nunavut	$ 1.14	supplies only	$1.60	reagent only
NB	$1.50	reagent/collection container	$0.67	reagent only

Source: Lab managers' survey results (2014)

Another source of data for test costs was provincial benefit schedules which provided fees for both ESR and CRP. Table 12 provides the costs derived from these fee schedules. As shown, the average fee for ESR and CRP from these provinces was $5.83 and $9.57. The cost of combined testing ($15.40) was assumed to be the sum of the individual test costs. These alternative costs were assumed in sensitivity analysis.

Table 12- Costs of ESR and CRP across provinces

Province	ESR	CRP	ESR+CRP
British Columbia[33]	$10.61	$10.31	$20.92
Alberta[34]	$3.72	$9.69	$13.41
Saskatchewan[35]	$10	$16.00	$26.00
Manitoba[36]	$3.25	$8.75	$12.00
Ontario[37]	$1.55	$3.10	$4.65
Average	$5.83	$9.57	$15.40

Abbreviations: CRP= C-reactive protein; ESR= erythrocyte sedimentation rate

Sensitivity analysis

Sensitivity analyses were conducted on all variables that affected the model. Specifically one way sensitivity analyses were conducted on the incremental cost of the combined test, the diagnostic accuracy (sensitivity and specificity) of the test strategies, and on the prevalence of conditions for the four different populations. The model was run using alternative testing costs based on the average from various provincial fee schedules. For diagnostic accuracy the model was run using the lower and upper confidence intervals for sensitivity and specificity. Similarly, the model was run using the lower and upper confidence intervals for the prevalence of the condition for each population. In addition, prevalence rates for all conditions were varied from 0.01 to 0.95 as the prevalence rates due to the concern that prevalence rates from the reported studies may not reflect those found in clinical practice.

6.2 Results

Base-case results

Periprosthetic Infection

Tables 13 and 14 present cost-effectiveness results for the periprosthetic infection population when the combined positive test is defined as ESR+CRP and ESR or CRP respectively. False positives and misdiagnoses are presented per 100 patients tested. As shown in Table 11 the number of false positives per 100 patients tested for ESR, CRP and ESR+CRP per patient are estimated to be 17.3, 15.6, and 8.7 respectively. This means a combined ESR+CRP would lead to 8.6 fewer false positives than ESR alone and 6.9 fewer false positives per 100 patients tested than CRP alone. When cost differences are taken into account, the incremental cost per false positive avoided for ESR+CRP compared to ESR alone is estimated to be $118.85. When compared to CRP alone the cost per false positive avoided is $152.89. The cost per total misdiagnosis avoided for ESR+CRP is estimated to be $157.02 and $240.62, respectively.

Table 13- Cost-effectiveness results: periprosthetic infection - ESR+CRP					
Testing Strategy	**Cost**	**False positives (per 100 patients)**	**Total misdiagnoses (per 100 patients)***	**$/False positive avoided**	**$/Misdiagnosis avoided**
ESR	$10.61	17.3	20.7		
CRP	$10.31	15.6	18.6		
ESR+CRP	$20.92	8.7	14.2		
ESR+CRP vs. ESR	$10.31	-8.6	-6.6	$118.85	$157.02
ESR+CRP vs. CRP	$10.61	-6.9	-4.4	$152.89	$240.62

Abbreviations: CRP= C-reactive protein; ESR= erythrocyte sedimentation rate
* total misdiagnoses = false positives + false negatives

Table 14 presents cost-effectiveness results for the periprosthetic infection population when a combined positive test is defined as either ESR or CRP being positive. As shown under this assumption ESR/CRP is estimated to produce 1.7 fewer false negatives compared to ESR alone and 1.3 fewer false negatives per 100 patients tested than CRP alone. The cost per false negative avoided for ESR/CRP is estimated to be $611.22 compared to ESR alone and $838.68 compared to CRP alone. ESR/CRP is dominated by (higher costs, less effectiveness) by both ESR alone CRP alone when all misdiagnoses are considered. The combined test leads to 5.6 more total misdiagnoses per 100 patients tested than CRP alone and 8.3 more misdiagnoses per 100 patients compared to CRP alone.

Table 14- Cost-effectiveness results: periprosthetic infection - ESR/CRP					
Testing Strategy	**Cost**	**False negatives (per 100 patients)**	**Total misdiagnoses (per 100 patients)**	**$/False negative avoided**	**$/Misdiagnosis avoided**
ESR	$10.61	3.4	21.0		
CRP	$10.31	3.0	18.3		
ESR/CRP	$20.92	1.7	26.6		
ESR/CRP vs. ESR	$10.31	-1.7	5.6	$611.22	dominated
ESR/CRP vs. CRP	$10.61	-1.3	8.3	$838.68	dominated

Abbreviations: CRP= C-reactive protein; ESR= erythrocyte sedimentation rate
* total misdiagnoses = false positives + false negatives

Pediatric orthopedic infections

Table 15 presents cost-effectiveness results for the pediatric orthopedic infection population, when the combined positive test is defined as both ESR+CRP being positive (ESR+CRP). As shown in the proportion of false positives for ESR, CRP and ESR+CRP per 100 patients are estimated to be .31.8, 12.1, and 9.3 respectively. This means a combined ESR+CRP would lead to .22.5 fewer false positives than ESR alone and 2.8 fewer false positives per 100 patients than CRP alone. When cost differences are taken into account, the incremental cost per false positive avoided for ESR+CRP compared to ESR alone is estimated to be $45.97. When

compared to CRP alone the cost per false positive avoided is $378.50. The cost per total misdiagnosis avoided for ESR+CRP is estimated to be $49.29 and $440.32 respectively.

Table 15- Cost-effectiveness results: Pediatric orthopedic infections - ESR+CRP

Testing Strategy	Cost	False positives (per 100 patients)	Total misdiagnoses (per 100 patients)*	$/False positive avoided	$/Misdiagnosis avoided
ESR	$10.61	31.8	32.6		
CRP	$10.31	12.1	14.0		
ESR+CRP	$20.92	9.3	11.6		
ESR+CRP vs. ESR	$10.31	-22.5	-20.9	$45.97	$49.29
ESR+CRP vs. CRP	$10.61	-2.8	-2.4	$378.50	$440.32

Abbreviations: CRP= C-reactive protein; ESR= erythrocyte sedimentation rate
* total misdiagnoses = false positives + false negatives

Table 16 presents cost-effectiveness results for the pediatric orthopedic infection population, when a combined positive test is defined as either ESR or CRP being positive. ESR/CRP is estimated to produce 0.3 fewer false negatives compared to ESR alone and 0.2 fewer false negatives than CRP alone. The cost per false negative avoided for ESR/CRP is estimated to be $3,929 compared to ESR alone and $5,391.26 compared to CRP alone. Because specificity was not reported in the single study informing this analysis, total misdiagnoses could not be estimated.

Table 16- Cost-effectiveness results: Pediatric orthopedic infections - ESR/CRP

Testing Strategy	Cost	False negatives (per 100 patients)	Total misdiagnoses (per 100 patients)	$/False negative avoided	$/Misdiagnosis avoided
ESR	$10.61	0.4	NA		
CRP	$10.31	0.3	NA		
ESR/CRP	$20.92	0.1	NA		
ESR/CRP vs. ESR	$10.31	-0.3	NA	$3,929.12	NA
ESR/CRP vs. CRP	$10.61	-0.2	NA	$5,391.26	NA

Abbreviations: CRP= C-reactive protein; ESR= erythrocyte sedimentation rate; NA= not available
* total misdiagnoses = false positives + false negatives

Inflammatory Bowel Disease

Table 17 presents cost-effectiveness results for the inflammatory bowel disease population when the combined positive test is defined as both ESR+CRP being positive (ESR+CRP). The proportion of false positives for ESR, CRP and ESR+CRP per patient are estimated to be 25.1, 22.9, and 12.2 per 100 patients tested respectively. This means a combined ESR+CRP would lead to 12.9 fewer false positives than ESR alone and 10.7 fewer false positives per 100 patients than CRP alone. When cost differences are taken into account, the incremental cost

per false positive avoided for ESR+CRP compared to ESR alone is estimated to be $79.60. When compared to CRP alone the cost per false positive avoided is $99.47. The cost per total misdiagnosis avoided for ESR+CRP is estimated to be $170.49 and $250.36 respectively.

Table 17- Cost-effectiveness results: Inflammatory bowel disease - ESR+CRP

Testing Strategy	Cost	False positives (per 100 patients)	Total misdiagnoses (per 100 patients)*	$/False positive avoided	$/Misdiagnosis avoided
ESR	$10.61	25.1	30.1		
CRP	$10.31	22.9	28.3		
ESR+CRP	$20.92	12.2	24.1		
ESR+CRP vs. ESR	$10.31	-12.9	-6.0	$79.60	$170.49
ESR+CRP vs. CRP	$10.61	-10.7	-04.2	$99.47	$250.36

Abbreviations: CRP= C-reactive protein; ESR= erythrocyte sedimentation rate
* total misdiagnoses = false positives + false negatives

Giant Cell Arteritis

Table 18 presents cost-effectiveness results for patients suspected of having giant cell arteritis using prevalence and diagnostic accuracy data from Kermani et al. The proportion of false positives for ESR, CRP and ESR+CRP in this population are estimated to be 53.3 52.5, and 44.9 per 100 patients tested respectively. Therefore, a combined ESR+CRP would lead to 8 fewer false positives than ESR alone and 7.6 fewer false positives per 100 patients than CRP alone. When cost differences are taken into to account, the incremental cost per false positive avoided for ESR+CRP compared to ESR alone is estimated to be $123.18. When compared to CRP alone the cost per false positive avoided is $139.44. The cost per total misdiagnosis avoided for ESR+CRP is estimated to be $134.73 and $165.43 respectively.

Table 18- Cost-effectiveness results: Giant cell arteritis Giant cell arteritis - ESR+CRP- Kermani et al[19]

Testing Strategy	Cost	False positives (per 100 patients)	Total misdiagnoses (per 100 patients)*	$/False positive avoided	$/Misdiagnosis avoided
ESR	$10.61	53.3	57.1		
CRP	$10.31	52.5	55.8		
ESR+CRP	$20.92	44.9	49.4		
ESR+CRP vs. ESR	$10.31	-8.4	-7.7	$123.18	$134.73
ESR+CRP vs. CRP	$10.61	-7.6	-6.4	$139.44	$165.43

Abbreviations: CRP= C-reactive protein; ESR= erythrocyte sedimentation rate
* total misdiagnoses = false positives + false negatives

Table 19 presents cost-effectiveness results for patients suspected of having giant cell arteritis using prevalence and diagnostic accuracy data from Hayreh et al[28]. The proportion of false positives for ESR, CRP and ESR+CRP in this population are estimated to be 23.4, 12.7, and

5.7 per 100 patients tested respectively. Therefore, a combined ESR+CRP would lead to 17.7fewer false positives than ESR alone and 7.1 fewer false positives per 100 patients than CRP alone. When cost differences are taken into to account, the incremental cost per false positive avoided for ESR+CRP compared to ESR alone is estimated to be $58.250. When compared to CRP alone the cost per false positive avoided is $149.86. The cost per total misdiagnosis avoided for ESR+CRP is estimated to be $57.30 and $163.33 respectively.

Table 19 Cost-effectiveness results: Giant cell arteritis Giant cell arteritis - ESR+CRP- Hayreh et al[28]

Testing Strategy	Cost	False positives (per 100 patients)	Total misdiagnoses (per 100 patients)*	$/False positive avoided	$/Misdiagnosis avoided
ESR	$10.61	23.4	24.2		
CRP	$10.31	12.7	12.7		
ESR+CRP	$20.92	5.7	6.2		
ESR+CRP vs. ESR	$10.31	-17.7	-18.0	$58.25	$57.30
ESR+CRP vs. CRP	$10.61	-7.1	-6.5	$149.86	$163.33

Abbreviations: CRP= C-reactive protein; ESR= erythrocyte sedimentation rate
* total misdiagnoses = false positives + false negatives

Sensitivity analysis

Table 20 presents cost-effectiveness results using cost estimates derived from provincial benefit fee schedules. As shown, the incremental cost per false positive avoided from the combined test compared to ESR alone is very similar using the alternate cost source compared to the base case assumption across the various populations. However compared to CRP alone, the incremental cost per false positive avoided is nearly half of what it is using the base case assumption. A similar pattern emerges for the incremental cost per false negative for ESR/CRP compared to ESR alone and CRP alone.

Table 20- Sensitivity analysis of cost-effectiveness using alternative estimates for ESR ($5.83), CRP ($9.57) and combined testing ($15.40) based on provincial fee schedules

	Periprosthetic Infections		Pediatric Orthopedic Infections		Inflammatory Bowel Disease		Giant Cell Arteritis (Kermani)		Giant Cell Arteritis (Hayreh)	
	vs. ESR	vs. CRP	vs. ESR	vs. CRP	vs ESR	vs. CRP	vs. ESR	vs. CRP	vs. ESR	vs. CRP
Test Cost source	Cost per false positive avoided of ESR+CRP vs. ESR or CRP alone									
Base-case	$119	$153	$46	$378	$80	$99	$123	$139	$58	$150
Alternative Costs (Provincial fee schedules)	$110	$84	$43	$208	$74	$55	$114	$77	$54	$82
Test Cost source	Cost per false negative avoided of ESR/CRP vs. ESR or CRP alone									
Base-case	$611	$839	$3,929	$5,391						
Alternative Costs (Provincial fee schedules)	$567	$461	$3,647	$2,960						

Abbreviations: CRP= C-reactive protein; ESR= erythrocyte sedimentation rate; vs.= versus

Cost-effectiveness results are shown in Table 21 when the lower and upper confidence intervals for sensitivity and specificity for the testing strategies are used in the model. The confidence intervals for sensitivity and specificity for each test by population were based on data provided in Appendix 12 and Table 2. As shown, varying diagnostic accuracy produces small variations in the incremental cost per false positive avoided for ESR+CRP in most populations. The cost per false positive avoided is impacted the most in the Giant Cell Arteritis population when using sensitivity and specificity estimates from Kermani et al[19]. More variation is seen is seen for the incremental cost per false negative avoided for ESR/CRP compared to ESR or CRP alone. In the pediatric orthopedic population, the cost per false negative avoided for ESR/CRP compared to ESR alone is estimated to be $2,619 when the lower confidence intervals for diagnostic accuracy for all test strategies are used. The cost per false negative avoided is estimate to be $5,239 if the upper 95% confidence intervals for diagnostic accuracy are used. The cost per false negative avoided for combined ESR/CRP compared to CRP alone is estimated to be $3,235 if the lower 95% confidence intervals for diagnostic accuracy are used as estimates for diagnostic accuracy. If the high 95% confidence interval for sensitivity is used for all testing strategies, the incremental cost per false negative for ESR/CRP compared to ESR alone avoided becomes $8,087.

Table 21- Sensitivity analysis of cost-effectiveness using lower and upper confidence intervals for diagnostic accuracy for all testing strategies

	Periprosthetic Infections		Pediatric Orthopedic Infections		Inflammatory Bowel Disease		Giant Cell Arteritis (Kermani)		Giant Cell Arteritis (Hayreh)	
	vs. ESR	vs. CRP	vs. ESR	vs. CRP	vs ESR	vs. CRP	vs. ESR	vs. CRP	vs. ESR	vs. CRP
ESR+CRP diagnostic accuracy	Cost per false positive avoided of ESR+CRP vs. ESR or CRP alone									
Base-case	$119	$153	$46	$378	$80	$99	$123	$139	$58	$150
Lower C.I.	$119	$167	$42	$378	$68	$93	$123	$139	$66	$125
Upper C.I.	$111	$153	$50	$378	$97	$139	$452	$697	$61	$187
ESR/CRP diagnostic accuracy	Cost per false negative avoided of ESR/CRP vs. ESR or CRP alone									
Base-case	$611	$839	$3,929	$5,391						
Lower C.I.	$489	$839	$2,619	$3,235						
Upper C.I.	$611	$1,258	$5,239	$8,087						

Abbreviations: CI= confidence interval; CRP= C-reactive protein; ESR= erythrocyte sedimentation rate; vs.=versus

Table 22 presents cost-effectiveness results when the lower and upper confidence intervals for prevalence of disease for each population are used in the model. Little variation in the cost per false positive avoided are found for ESR+CRP. Lower prevalence leads to lower cost per false positive avoided for all populations. This is because lower prevalence rates leads to more false positive test results and therefore more false positives avoided for ESR+CRP because of its higher specificity. More variation in costs is found in the incremental cost per false negative avoided for ESR/CRP. Higher prevalence leads to lower cost per false negative avoided for all populations. This is because higher prevalence rates leads to less false negative test results and therefore fewer false negatives avoided for ESR/CRP.

Table 22- Sensitivity analysis of cost-effectiveness using lower and upper confidence intervals of condition prevalence

	Periprosthetic Infections		Pediatric Orthopedic Infections		Inflammatory Bowel Disease		Giant Cell Arteritis (Kermani)		Giant Cell Arteritis (Hayreh)	
	vs. ESR	**vs. CRP**	**vs. ESR**	**vs. CRP**	**Vs. ESR**	**vs. CRP**	**vs ESR**	**vs. CRP**	**vs ESR**	**vs. CRP**
ESR+CRP prevalence data used	Cost per false positive avoided of ESR+CRP vs. ESR or CRP alone									
Base-case (mean)	$119	$153	$46	$378	$80	$99	$123	$139	$58	$150
Lower CI	$113	$145	$45	$367	$70	$87	$118	$134	$56	$144
Upper CI	$126	$162	$48	$391	$92	$115	$128	$145	$61	$157
ESR/CRP prevalence data used	Cost per false negative avoided of ESR/CRP vs. ESR or CRP alone									
Base-case(mean)	$611	$839	$3,929	$5,391						
Lower C.I.	$660	$905	$7,271	$9,977						
Upper C.I.	$569	$781	$2,692	$3,694						

Abbreviations: CI= confidence interval; CRP= C-reactive protein; ESR= erythrocyte sedimentation rate; vs.=versus

Table 23 presents sensitivity analysis using the lower 95% confidence intervals for both diagnostic accuracy and condition prevalence as model values for all testing strategies and using upper 95% confidence intervals for both condition prevalence and diagnostic for all testing strategies. As shown, there is not much difference in the cost per false positive avoided for ESR+CRP between the lower and upper confidence sensitivity analysis except in the Giant Cell Arteritis population when using diagnostic accuracy data from Kermani et al[19]. There is also not much difference in the cost per false negative of ESR/CRP vs either CRP or ESR alone between the lower and upper 95% confidence interval sensitivity analysis.

Table 23- Sensitivity analysis of cost-effectiveness using lower and upper confidence intervals of both diagnostic accuracy and condition prevalence

	Periprosthetic Infections		Pediatric Orthopedic Infections		Inflammatory Bowel Disease		Giant Cell Arteritis (Kermani)		Giant Cell Arteritis (Hayreh)	
	vs. ESR	**vs. CRP**	**vs. ESR**	**vs. CRP**	**Vs. ESR**	**vs. CRP**	**vs ESR**	**vs. CRP**	**vs ESR**	**vs. CRP**
ESR+CRP prevalence and diagnostic accuracy data used	Cost per false positive avoided of ESR+CRP vs. ESR or CRP alone									
Base-case	$119	$153	$46	$378	$80	$99	$123	$139	$58	$150
Lower CI	$113	$158	$41	$367	$59	$82	$118	$134	$63	$120
Upper CI	$118	$162	$52	$391	$112	$162	$470	$726	$63	$196
ESR/CRP prevalence and diagnostic accuracy data used	Cost per false negative avoided of ESR/CRP vs. ESR or CRP alone									
Base-case)	$611	$839	$3,929	$5,391						
Lower C.I.	$528	$905	$4,848	$5,986						
Upper C.I.	$569	$1,172	$3,589	$5,540						

Abbreviations: CI= confidence interval; CRP= C-reactive protein; ESR= erythrocyte sedimentation rate; vs.=versus

Table 24 presents the cost per false positive avoided (ESR+CRP) for each population varying the prevalence rate from, 0.01 to 0.95.A sensitivity was conducted varying prevalence rate because of the concern that condition prevalence found in the studies included in this analysis may not reflect prevalence rates found in clinical practice. As shown, as prevalence rates increase the cost per false positive avoided also increases. This is because as prevalence rates increase, the number of potential false positives avoided by using a combined CRP and ESR test decreases.

Table 24- Sensitivity analysis of cost-effectiveness of ESR+CRP varying the condition prevalence for each population

	Periprosthetic Infections		Pediatric Orthopedic Infections		Inflammatory Bowel Disease		Giant Cell Arteritis (Kermani)		Giant Cell Arteritis (Hayreh)	
	vs. ESR	vs. CRP	vs. ESR	vs. CRP	vs. ESR	vs. CRP	vs. ESR	vs. CRP	vs. ESR	vs. CRP
Prevalence rate										
0.01	$69	$89	$43	$357	$61	$77	$95	$107	$42	$107
0.05	$72	$93	$45	$372	$64	$80	$99	$112	$43	$112
0.10	$76	$98	$48	$393	$67	$84	$104	$118	$46	$118
0.15	$81	$104	$51	$416	$71	$89	$110	$125	$49	$125
0.20	$86	$111	$54	$442	$76	$95	$117	$133	$52	$133
0.25	$92	$118	$57	$472	$81	$101	$125	$141	$55	$141
0.30	$98	$126	$61	$505	$87	$108	$134	$152	$59	$152
0.35	$106	$136	$66	$544	$93	$117	$144	$163	$63	$163
0.40	$115	$147	$72	$589	$101	$126	$156	$177	$69	$177
0.45	$125	$161	$78	$643	$110	$138	$170	$193	$75	$193
0.50	$137	$177	$86	$707	$121	$152	$187	$212	$82	$212
0.55	$153	$196	$95	$786	$135	$168	$208	$236	$92	$236
0.60	$172	$221	$107	$884	$152	$189	$234	$265	$103	$265
0.65	$196	$253	$123	$1,010	$173	$217	$268	$303	$118	$303
0.70	$229	$295	$143	$1,179	$202	$253	$312	$354	$137	$354
0.75	$275	$354	$172	$1,415	$243	$303	$375	$424	$165	$424
0.80	$344	$442	$215	$1,768	$303	$379	$469	$531	$206	$531
0.85	$458	$589	$286	$2,358	$404	$505	$625	$707	$275	$707
0.90	$687	$884	$430	$3,537	$606	$758	$937	$1,061	$412	$1,061
0.95	$1,375	$1,768	$859	$7,073	$1,213	$1,516	$1,875	$2,122	$825	$2,122

Table 25 presents the cost per false negative avoided (ESR/CRP) for each population varying the prevalence rate from, 0.01 to 0.95. As shown, as prevalence rates increase the cost per false positive avoided decreases. This is because as prevalence rates increase, the number of potential false negatives avoided by using a combined CRP and ESR test increases. Higher false negative avoided results in lower incremental cost per false negative avoided. In the base case analysis, a prevalence rate of 42.2% was used for periprosthetic infections. A prevalence rate of 6.6% was used for the pediatric orthopedic infections. Note that the results are identical for the two populations for each prevalence rate sued in the sensitivity analysis. This is because the difference in sensitivity between the combined ESR/CRP test and the individual tests are the same for these two populations. Therefor when both the diagnostic accuracy and prevalence assumed are the same, the number of false negatives avoided will also be the same for the two populations.

Table 25- Sensitivity analysis of cost-effectiveness of ESR/CRP varying the condition prevalence for each population

	Periprosthetic Infections		Pediatric Orthopedic Infections		Inflammatory Bowel Disease		Giant Cell Arteritis	
	vs. ESR	vs. CRP	vs. ESR	vs. CRP	vs. ESR	vs. CRP	vs. ESR	vs. CRP
Prevalence rate								
0.01	$25,775	$35,367	$25,775	$35,367				
0.05	$5,155	$7,073	$5,155	$7,073				
0.10	$2,578	$3,537	$2,578	$3,537				
0.15	$1,718	$2,358	$1,718	$2,358				
0.20	$1,289	$1,768	$1,289	$1,768				
0.25	$1,031	$1,415	$1,031	$1,415				
0.30	$859	$1,179	$859	$1,179				
0.35	$736	$1,010	$736	$1,010				
0.40	$644	$884	$644	$884				
0.45	$573	$786	$573	$786				
0.50	$516	$707	$516	$707				
0.55	$469	$643	$469	$643				
0.60	$430	$589	$430	$589				
0.65	$397	$544	$397	$544				
0.70	$368	$505	$368	$505				
0.75	$344	$472	$344	$472				
0.80	$322	$442	$322	$442				
0.85	$303	$416	$303	$416				
0.90	$286	$393	$286	$393				
0.95	$271	$372	$271	$372				

7 Discussion

7.1 Summary of Main findings

The systematic review of studies evaluating the diagnostic value of ESR and CRP tests, individually and in combination, for the diagnosis of suspected inflammatory conditions or severe infections, included a total of nine primary studies. As expected, the results of this review indicated that ESR/CRP tests had consistently higher sensitivity values and lower specificity values relative to ESR and CRP tests alone. On the contrary, sensitivity values of ESR+CRP tests were consistently lower and their specificity values were consistently higher than those of individual ESR and CRP tests. However, in our analysis, the performance of combined ESR and CRP tests (ESR+CRP or ESR/CRP) versus ESR and CRP varied across the conditions.

The results of this review showed that ESR+CRP was statistically more specific than both ESR and CRP (p-values<0.001) in diagnosis of periprosthetic infections and giant cell arteritis, and more specific than ESR in the diagnosis of orthopedic infections in children (p-value<0.001). In terms of relative sensitivity, the ESR/CRP combination was shown to be statistically superior to

ESR (p-values <0.05), and comparable, if not superior, to CRP (p-values= 0.05) in diagnosis of periprosthetic infections in adults and orthopedic infections in children. Our analysis revealed no statistical differences in diagnostic performance measures between ESR+CRP versus ESR and CRP alone in detecting inflammatory bowel diseases, and differentiating viral from bacterial bronchiolitis. No data were available on ESR/CRP testing in inflammatory bowel diseases or giant cell arteritis. The variability we found in our results could be attributed to the paucity of data and differences in study populations (eligibility criteria) or diagnostic thresholds.

The comparison of the estimates of ODA suggested that ESR+CRP could be significantly overall more accurate than both ESR and CRP tests for diagnosis of periprosthetic infections and giant cell arteritis, and overall more accurate than ESR, but not CRP, in diagnosis of orthopedic infections in children. No statistical differences were found between the ODA of ESR+CRP and that of ESR or CRP alone in inflammatory bowel diseases. It should be noted that a higher ODA estimate means that using the candidate test can result in more correctly diagnosed subjects. The measure uses the absolute number of correctly diagnosed participants (TP+TN), In addition, ODA increases as disease prevalence decreases, but sensitivity and specify remain unaffected.[38] Therefore, this measure should be interpreted considering other measures of diagnostic performance such as sensitivity, specificity, and predictive values. In addition, ODA increases as disease prevalence decreases, but sensitivity and specify remain unaffected.[38]

One published economic evaluation[29] was identified that evaluated the cost-effectiveness of a combined test of ESR and CRP vs. ESR alone to CRP alone. In this 2003 study the cost per useful result of combined ESR and CRP testing was found to be ¥1737 ($CAN26.67) compared to CRP alone and ¥2001 ($CAN 30.72) respectively However, this was a Japanese-based study and the generalizability of their results to the Canadian setting is questionable, given the methods used and difference in populations. Additionally, “useful result” was not an outcome of interest in our clinical review. Therefore a primary economic analysis was undertaken in which the cost-effectiveness outcomes were cost per false positive avoided and cost per false negative avoided, based on the diagnostic accuracy data and prevalence data collected as part of our clinical review.

In our primary economic analysis, the base case incremental cost per false positive avoided for ESR+CRP compared to ESR alone ranged from $46 for pediatric infections to $123 for the detection of giant cell arteritis when using prevalence and diagnostic accuracy data from Kermani et al[19]. The incremental cost per false positive avoided for ESR+CRP vs. CRP alone ranged from $99 for inflammatory bowel disease to $378 for pediatric orthopedic infections. The cost effectiveness varied across populations due to both differences in prevalence and due to the differences in added specificity of ESR+CRP vs. either test alone.

The base case cost per false negative avoided for ESR/CRP was estimated to be $611 compared to ESR alone and $830 compared to CRP alone for the detection of perioperative infection. For the detection of pediatric infection, the base case cost per false negative of ESR/CRP compared to ESR was estimated to be $3,929 compared to ESR alone and $5,391 compared to CRP alone. The cost per false positive negative avoided for ESR/CRP compared to either test alone was higher in the pediatric orthopedic infection population than in the periprosthetic infection population. This is likely due to the differences in the assumed prevalence for these conditions. Cost-effectiveness results were very sensitive to the assumed prevalence of the condition in the population. For example if the prevalence of pediatric infection

in the tested population was 25%, the cost per false negative avoided of CRP/ESR becomes $1,289 and $1,760 compared to ESR alone and CRP alone respectively.

Drawing conclusions about whether combined testing is cost-effective based on this analysis is difficult. There are not common cost-effectiveness thresholds for the incremental cost per false negative or false positive avoided to reference. Furthermore, whether a cost per false test result is cost-effective is likely to differ across populations, as the consequences of a false test also would likely differ across populations.

7.2 Strengths and Limitations

The clinical review provides a comprehensive review of available comparative evidence on accuracy of ESR, CRP, ESR+CRP and ESR/CRP tests for diagnosis of inflammatory and infectious conditions. The review highlights the limitations of the existing evidence by critically appraising the quality of the included studies. In addition to the direct comparison of the tests of interest to the reference standard employed in each study, we also indirectly compared the relative accuracy the combined ESR and CRP testing to either ESR or CRP alone.

As with all research, our review has a number of limitations.. Very few studies were identified for each condition. As a result, we were not able to provide sufficient data and perform pooled analysis on the diagnostic performance of the study tests for various clinical conditions. Data from a larger number of studies would particularly be helpful to perform a more efficient indirect meta-analysis.[39] Additionally, the studies included in the systematic review were highly heterogeneous in terms of setting, prevalence of target condition, and inclusion criteria. Therefore, we are limited in drawing firm conclusions about our findings.

The economic evaluation also has limitations. The outcome of incremental cost per false test result (positive or negative) is intermediate and makes it difficult to draw conclusions on cost-effectiveness. In order to translate this into a more common cost-effectiveness outcome such as the incremental cost per QALY, the cost and health consequences of both false positives and false negatives would need to be incorporated. These would differ according to the population being evaluated. Because of the large number of different potential populations, and a paucity of lack of knowledgeinformation of regarding which populations for which there would be diagnostic data would be available, available; more final cost-effectiveness outcomes were not evaluated.

Additionally, the analysis was limited to pairwise comparisons of a combined testing strategy to ESR alone or CRP alone. Therefore cost-effectiveness analysis considering all three testing strategies simultaneously could not take place. Therefore conclusions of which of the three strategies (ESR alone, CRP alone, combined testing) for a given willingness to pay threshold could not be made.

7.3 Generalizability of Findings

The majority of the studies included in our review were performed in the US or UK and were mainly based in academic or referral hospitals. Although the results seem to be generalizable to other similar centres, their generalizability to general practice in Canada may be limited. In addition, the prevalence of the target conditions (e.g., periprosthetic infection) in some of the included studies was likely to be higher than it would be seen in general practice. Thus, some of the results presented in this review are based on patient populations with high pre-test probabilities.

7.4 Other Considerations

Overall diagnostic accuracy measures, that were estimated based on the area under ROC curves, indicated that ESR and CRP, regardless if they were conducted in pair or individually, have a maximum overall diagnostic accuracy of 0.93. Therefore, these tests should not be relied on to rule in or rule out inflammatory or infectious conditions and their results need to be considered in the context of other clinical findings.

REFERENCES

1. Pagana KD, Pagana TJ, MacDonald S. *Mosby's Canadian manual of diagnostic and laboratory tests*. Toronto: Mosby Canada; 2013.

2. *CRP vs ESR assessing & measuring the inflammatory response*. Dunedin (New Zealand): bpac[nz]; 2005. Available from: http://www.bpac.org.nz/resources/campaign/crp_esr/crp_esr_poem.asp; www.bpac.org.nz/resources/other/audits/bpac_crp_audit_wv.pdf

3. Kushner I. *Acute phase reactants*. In: Furst DE, Romain PL, editors. UpToDate. Waltham, MA:2014 [cited 2014 Aug 8]. Available from: www.uptodate.com.

4. Morris SA. C-Reactive Protein (CRP): the best lab test for inflammation/infection. *Alberta Health Serv Lab Rep*. 2013 Apr 1;3(1). Available from: http://www.albertahealthservices.ca/LabServices/wf-lab-nls-2013-04-crp.pdf

5. Chandrashekara S, Renuka P, Suresh KP. ESR or CRP, which inflammatory measure can accurately replace clinical measures in rheumatoid arthritis? *Ind J Rheumatol*. 2012;7(2):69-73.

6. Crowson CS, Rahman MU, Matteson EL. Which measure of inflammation to use? A comparison of erythrocyte sedimentation rate and C-reactive protein measurements from randomized clinical trials of golimumab in rheumatoid arthritis. *J Rheumatol*. 2009;36(8):1606-10.

7. Watson J, Round A, Hamilton W. Raised inflammatory markers. *BMJ*. 2012;344:e454.

8. AHS Laboratory Services. C-reactive protein (CRP) and Erythrocyte Sedimentation Rate (ESR) Testing. *Alberta Health Serv Lab Bull*. 2014 Mar 19. Available from: http://www.albertahealthservices.ca/LabServices/wf-lab-bulletin-esr-crp-testing.pdf

9. Saenger A, Block D. *C-Reactive Protein and Erythrocyte SedimentationRate Testing: Should Both Be Ordered?* In: Utilization Spotlight. Mayo Medical Laboratories; 2012 Nov [cited 2014 Aug 8]. Available from: http://www.mayomedicallaboratories.com/articles/communique/2012/11utiliz.html.

10. Kainth MK, Gigliotti F. Simultaneous Testing of Erythrocyte Sedimentation Rate and C-Reactive Protein: Increased Expenditure without Demonstrable Benefit. *J Pediatr*. 2014 Jun 21.

11. Colombet I, Pouchot J, Kronz V, Hanras X, Capron L, Durieux P, et al. Agreement between erythrocyte sedimentation rate and C-reactive protein in hospital practice. *Am J Med*. 2010 Sep;123(9):863-13.

12. 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.

13. Walter SD. Properties of the summary receiver operating characteristic (SROC) curve for diagnostic test data. *Stat Med*. 2002;21(9):1237-56.

14. Wells GA, Sultan SA, Chen L, Khan M, Coyle D. *Indirect treatment comparison. Version 1.0 [computer program]*. Ottawa: Canadian Agency for Drugs and Technologies in Health (CADTH); 2009.

15. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol*. 1997;50(6):683-91.

16. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. *Analysing and presenting results*. Version 1.0. In: Deeks JJ, Bossuyt PM, Gatsonis C, editors. Cochrane handbook for systematic reviews of diagnostic test accuracy. The Cochrane Collaboration; 2010. Chapter 10 [cited 2012 Oct 29]. Available from: http://srdta.cochrane.org/.

17. *Imputing standard deviations*. In: Higgins JPT, Green S, editors, editors. Cochrane handbook for systematic reviews of interventions 5.1.0 [updated March 2011]. The Cochrane Collaboration; 2011. Chapter 16 [cited 2012 Oct 29]. Available from: www.cochrane-handbook.org.

18. Costa CR, Johnson AJ, Naziri Q, Maralunda GA, Delanois RE, Mont MA. Efficacy of erythrocyte sedimentation rate and C-reactive protein level in determining periprosthetic hip infections. *Am J Orthop*. 2012 Apr;41(4):160-5.

19. Kermani TA, Schmidt J, Crowson CS, Ytterberg SR, Hunder GG, Matteson EL, et al. Utility of erythrocyte sedimentation rate and C-reactive protein for the diagnosis of giant cell arteritis. *Semin Arthritis Rheum*. 2012 Jun;41(6):866-71.

20. Robinson S, Leonard P. C reactive protein, erythrocyte sedimentation rate, or both, in the diagnosis of atraumatic paediatric limb pain? *Emerg Med J*. 2012 Dec;29(12):969-71.

21. Johnson AJ, Zywiel MG, Stroh A, Marker DR, Mont MA. Serological markers can lead to false negative diagnoses of periprosthetic infections following total knee arthroplasty. *Int Orthop*. 2011 Nov;35(11):1621-6.

22. 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 Mar;468(3):861-6.

23. Austin MS, Ghanem E, Joshi A, Lindsay A, Parvizi J. A simple, cost-effective screening protocol to rule out periprosthetic infection. *J Arthroplasty*. 2008 Jan;23(1):65-8.

24. Dolwani S, Metzner M, Wassell JJ, Yong A, Hawthorne AB. Diagnostic accuracy of faecal calprotectin estimation in prediction of abnormal small bowel radiology. *Aliment Pharmacol Ther*. 2004 Sep 15;20(6):615-21.

25. Ghanem E, Antoci V, Jr., Pulido L, Joshi A, Hozack W, Parvizi J. The use of receiver operating characteristics analysis in determining erythrocyte sedimentation rate and C-reactive protein levels in diagnosing periprosthetic infection prior to revision total hip arthroplasty. *Int J Infect Dis*. 2009 Nov;13(6):e444-e449.

26. Jellema P, van Tulder MW, van der Horst HE, Florie J, Mulder CJ, van der Windt DA. Inflammatory bowel disease: a systematic review on the value of diagnostic testing in primary care. *Colorectal Dis*. 2011;13(3):239-54. Available from: http://onlinelibrary.wiley.com/doi/10.1111/j.1463-1318.2009.02131.x/abstract

27. Grzesk E, Koltan S, Grzesk G, et al. Value of erythrocyte sedimentation rate, c-reactive protein and procalcitonin concentration versus multimarker strategy in management of bronchiolitis in pediatric emergency. *Med Biol Sci*. 2012;26.2:11-7. Available from: http://www.degruyter.com/view/j/mbs.2012.26.issue-2/v10251-012-0025-7/v10251-012-0025-7.xml

28. Hayreh SS, Podhajsky PA, Raman R, Zimmerman B. Giant cell arteritis: validity and reliability of various diagnostic criteria. *Am J Ophthalmol*. 1997 Mar;123(3):285-96.

29. Takemura Y, Ishida H, Inoue Y. Utilization of common inflammatory markers in new, symptomatic, primary care outpatients based on their cost-effectiveness. *Clin Chem Lab Med*. 2003 May;41(5):668-74.

30. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ*. 1996 Aug 3 [cited 2015 Mar 12];313(7052):275-83. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2351717/

31. *Daily currency converter*. Ottawa: Bank of Canada; 2015. [cited 2015 Mar 12]. Available from: http://www.bankofcanada.ca/rates/exchange/daily-converter/

32. *Consumer Price Index, health and personal care, by province (monthly) (Newfoundland and Labrador)*. In: CANSIM, table 326-0020 and Catalogue nos. 62-001-X and 62-010-X. Ottawa: Statistics Canada; 2015 [cited 2015 Mar 19]. Available from: http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/cpis13a-eng.htm.

33. Medical Services Commission. *Laboratory medicine*. Victoria: Province of British Columbia; 2013 Aug 1. [cited 2014 Sep 8]. Available from: https://www.health.gov.bc.ca/msp/infoprac/physbilling/payschedule/pdf/40-laboratory-medicine.pdf

34. *Alberta Health Care insurance plan: medical procedure list as of 01 April 2014*. Edmonton: Government of Alberta; 2014 Apr 1. [cited 2014 Sep 8]. Available from: www.health.alberta.ca/documents/SOMB-Medical-Procedures-2014-04.pdf

35. *Payment schedule for insured services provided by a physician*. Regina: Government of Saskatchewan; 2014 Apr 1. [cited 2014 Sep 8]. Available from: http://www.health.gov.sk.ca/physician-documents

36. Minister of Health, Manitoba Health, Healthy Living and Seniors. *Manitoba physician's manual*. Winnipeg: Province of Manitoba; 2014 Apr 1. [cited 2014 Sep 8]. Available from: www.gov.mb.ca/health/documents/physmanual.pdf

37. Ontario Ministry of Health and Long-Term Care. *Schedule of benefits for laboratory services*. Toronto: Queen's Printer for Ontario; 1999 Apr 1. [cited 2014 Sep 8]. Available from:

http://www.health.gov.on.ca/english/providers/program/ohip/sob/lab/lab_services_sched_01_19990401.pdf

38. Eusebi P. Diagnostic accuracy measures. *Cerebrovasc Dis*. 2013;36(4):267-72.

39. Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, et al. Indirect comparisons of competing interventions. *Health Technol Assess*. 2005 [cited 2013 Jul 15];9(26):1-134, iii-iv. Available from: http://www.hta.ac.uk/execsumm/summ926.htm

40. Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. *J Clin Epidemiol*. 1991;44(8):763-70.

APPENDIX 1: LITERATURE SEARCH STRATEGY

OVERVIEW	
Interface:	National Library of Medicine
Databases:	PubMed **Note:** Subject headings have been customized for each database.
Date of Search:	June 11, 2014
Alerts:	Monthly search updates began June 11, 2014 and ran until August 2014.
Study Types:	Studies with diagnostic outcomes of interest and economic literature.
Limits:	Publication years 2004-June 2014 Humans

SYNTAX GUIDE	
[mesh]	Medical Subject Heading
[majr]	MeSH major topic
*	After a word, a truncation symbol (wildcard) to retrieve plurals or varying endings
[ti]	Title
[tiab]	Title or Abstract
[pt]	Publication type
[rn]	CAS registry number

Search Strategy

Line #	Search Strategy
#13	Search **(#4 AND #7) NOT #10** Filters: **Publication date from 2004/01/01 to 2014/12/31; English**
#11	Search **(#4 AND #7) NOT #10**
#12	Search **(#4 AND #7) NOT #10** Filters: **English**
#10	Search **#8 OR #9**
#9	Search **((Animals[MESH] OR Animal Experimentation[MESH] OR "Models, Animal"[MESH] OR Vertebrates[MESH]) NOT (Humans[MESH] OR Human experimentation[MESH])) OR (((animals[tiab] OR animal model[tiab] OR rat[tiab] OR rats[tiab] OR mouse[tiab] OR mice[tiab] OR rabbit[tiab] OR rabbits[tiab] OR pig[tiab] OR pigs[tiab] OR porcine[tiab] OR swine[tiab] OR dog[tiab] OR dogs[tiab] OR hamster[tiab] OR hamsters[tiab] OR chicken[tiab] OR chickens[tiab] OR sheep[tiab]) AND (publisher[sb] OR inprocess[sb] OR pubmednotmedline[sb])) NOT (human[ti] OR humans[ti] OR people[ti] OR children[ti] OR adults[ti] OR seniors[ti] OR patient[ti] OR patients[ti]))**
#8	Search **editorial[pt] OR comment[pt] OR letter[pt] OR newspaper article[pt]**
#7	Search **#5 OR #6**
#6	Search **9007-41-4[rn] OR c-reactive protein*[tiab] OR creactive protein*[tiab] OR**

	circulating reactive protein*[tiab] OR CRP[tiab] OR POC-CRP*[tiab] OR CRP-test*[tiab] OR POCCRP-test*[tiab] OR hsCRP[tiab] OR hs-CRP[tiab]
#5	Search **C-Reactive Protein[majr]**
#4	Search **#1 OR #2 OR #3**
#3	Search **erythrocyte sedimentation*[tiab] OR erythrocyte sediment[tiab] OR blood sedimentation*[tiab] OR blood sediment[tiab]**
#2	Search **westergren[tiab] OR Biernacki's reaction[tiab] OR ESR[tiab] OR ESR-test*[tiab] OR sed rate[tiab]**
#1	Search **Blood Sedimentation[majr]**

OTHER DATABASES	
Cochrane Library Issue 6, 2014	Same MeSH, keywords, and date limits used as per Medline search, excluding study types and Human restrictions. Syntax adjusted for Cochrane Library databases.

Grey Literature

Dates for Search:	June 2014
Keywords:	Included terms for erythrocyte sedimentation rate and c-reactive protein.
Limits:	Publication years 2004-present

Relevant websites from the following sections of the CADTH grey literature checklist, "Grey matters: a practical tool for evidence-based searching" (http://www.cadth.ca/en/resources/finding-evidence-is/grey-matters) were searched:

- Health Technology Assessment Agencies
- Health Economics
- Clinical Practice Guidelines
- Databases (free)
- Internet Search

APPENDIX 2: TITLE AND ABSTRACT SCREENING CHECKLIST

Ref#:
Author:
Year:

Population	Any population
Intervention	a) Erythrocyte Sedimentation Rate (ESR) b) C-reactive Protein (CRP) c) A combination of ESR and CRP
Comparator	Any comparator
Outcome	Diagnostic performance measures -sensitivity -specificity -positive predictive value (PPV) -negative predictive value (NPV) -receiver operator curve (ROC) -area under the curve (AUC)
Study type	Any study type

APPENDIX 3: FULL TEXT SCREENING CHECKLIST

Ref#:
Author:
Year:

	Include	Exclude
Population	□ Any undifferentiated population not being tested to monitor an existing condition	• Studies without a defined population/condition
Intervention	ESR and/or CRP used in combination: □ ESR and CRP □ ESR OR CRP	• No results on the ESR and CRP test combination • ESR, CRP and other test combinations (≥3 tests)
Comparators	□ ESR or CRP alone	• High-sensitivity CRP • ESR test performed using Wintrobe method • Tests that occur outside a central laboratory (e.g., in a physician's office) • Point-of-care tests
Outcomes	□ Diagnostic performance (sensitivity, specificity, PPV, NPV, AUC, ODA) □ Cost-effectiveness (ICUR, ICER, other)	Clinical utility
Study types	□ RCTs, prospective or retrospective observational (non-randomized) studies (cross-sectional, cohort, case-control)	Other types of study design (e.g. case reports)

Abbreviations: AUC= area under the curve; NPV=negative predictive value; ODA=Overall diagnostic accuracy; PPV=positive predictive value; ROC=receiver operator curve

Reason for exclusion: ----------

APPENDIX 4: QUADAS-2 TOOL FOR THE QUALITY ASSESSMENT OF DIAGNOSTIC ACCURACY STUDIES [12]

State the review question

Patients (setting, intended use of index test, presentation, prior testing): .
Index test(s): .
Reference standard and target condition: .

Draw a flow diagram for the

Risk of bias and applicability judgments

Domain 1: Patient selection

A. Risk of bias	
Describe methods of patient selection:	
Was a consecutive or random sample of patients enrolled?	**Yes / No / Unclear**
Was a case-control design avoided?	**Yes / No / Unclear**
Did the study avoid inappropriate exclusions?	**Yes / No / Unclear**
Could the selection of patients have introduced bias? **Risk: Low / High / Unclear**	
B. Concerns regarding applicability	
Describe included patients (prior testing, presentation, intended use of index test and setting): . . .	
Is there concern that the included patients do not match the review question? **Concern: Low / High / Unclear**	

Domain 2: Index test(s)

A. Risk of bias	
Describe the index test and how it was conducted and interpreted:	
Were the index test results interpreted without knowledge of the results of the reference standard?	**Yes / No / Unclear**
If a threshold was used, was it pre-specified?	**Yes / No / Unclear**
Could the conduct or interpretation of the index test have introduced bias? **Risk: Low / High / Unclear**	
B. Concerns regarding applicability	
Is there concern that the index test, its conduct, or interpretation differ from the review question? **Concern: Low / High / Unclear**	

Domain 3: Reference standard

A. Risk of bias	
Describe the reference standard and how it was conducted and interpreted:	
Is the reference standard likely to correctly classify the target condition?	**Yes / No / Unclear**
Were the reference standard results interpreted without knowledge of the results of the index test?	**Yes / No / Unclear**
Could the reference standard, its conduct, or its interpretation have introduced bias? **Risk: Low / High / Unclear**	
B. Concerns regarding applicability	
Is there concern that the target condition as defined by the reference standard does not match the review question? **Concern: Low / High / Unclear**	

Domain 4: Flow and timing

A. Risk of bias	
Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): . . . Describe the time interval and any interventions between index test(s) and reference standard:	
Was there an appropriate interval between index test(s) and reference standard?	**Yes / No / Unclear**
Did all patients receive a reference standard?	**Yes / No / Unclear**
Did patients receive the same reference standard?	**Yes / No / Unclear**
Were all patients included in the analysis?	**Yes / No / Unclear**
Could the patient flow have introduced bias? **Risk: Low / High / Unclear**	

APPENDIX 5: CLINICAL DATA ABSTRACTION FORM

STUDY	
Ref ID	
Author	
Publication Year	
Country	
Funding	

METHODOLOGY	
Study design	
Setting	
Total sample size	
Condition(s) the tests are used for	
Other inclusion criteria	
Exclusion criteria	

INTERVENTION/COMPARATOR			
	ESR	**CRP**	**Reference standard ----------**
Manufacturer			
Sample size			
Cut-off point(s) for a positive test			

POPULATION CHARACTERISTICS				
	ESR	**CRP**	**ESR+CRP**	**Reference standard ----------**
Mean age, year (SD)				
Gender (% female)				
Ethnicity (% white)				
Concurrent conditions				
Other important variables(unit) **1.----------( )** **2.----------(……..)**				

RESULTS

Outcome	CRP	ESR	ESR+CRP	ESR/CRP	P-value (comparison)
Diagnostic test performance					
Total number tested					
No true positives (%)					
No true negatives (%)					
No false positives (%)					
No false negatives (%)					
sensitivity (95% CI)					
Specificity (95% CI)					
Positive likelihood ratio (95% CI)					
Negative likelihood ratio (95% CI)					
Positive predictive value (95% CI)					
Negative predictive value (95% CI)					
Area under ROC Curve (95% CI)					
Overall diagnostic accuracy(95% CI)					

APPENDIX 6: DETAILS OF OUTCOME MEASURESFOR ASSESSMENT OF DIAGNOSTIC TEST PERFORMANCE

		Reference standard		
		Positive (Disease +)	Negative (Disease –)	Total
Index test	Positive	TP	FP	TP+FP
	Negative	FN	TN	FN+TN
	Total	TP+FN	FP+TN	TP+FP+TN+FN

TP= True positives; when the positive index test agrees with the positive reference standard.

FP= False positives; when the positive index test disagrees with the negative reference standard.

FN= False negatives; when the negative index test disagrees with the positive reference standard.

TN= True negatives; when the negative index test agrees with the negative reference standard.

From this 2 x 2 table, several tests of diagnostic performance can be made with confidence intervals.[40]

Sensitivity: TP/(TP+FN): the proportion of persons with the disease who are correctly identified by a test. i.e, a test with a high sensitivity is useful for 'ruling out' a disease if a person tests negative.

Confidence interval: $p \pm Z * \sqrt{\frac{p*(1-p)}{TP+FN}}$

Specificity: TN/(TN+FP). the proportion of persons without a disease who are correctly identified by a test. High specificity is important when the treatment or diagnosis is harmful to the patient

Confidence interval: $p \pm Z * \sqrt{\frac{p*(1-p)}{TN+FP}}$

Positive Predictive Value (PPV): TP/(TP+FP): the proportion of patients with positive test results who are correctly diagnosed.

Confidence interval: $p \pm Z * \sqrt{\frac{p*(1-p)}{TP+FP}}$

Negative Predictive Value (NPV): TN/(TN+FN): proportion of patients with negative test results who are correctly diagnosed

Confidence interval: $p \pm Z * \sqrt{\frac{p*(1-p)}{TN+FN}}$

Positive Likelihood Ratio (LR+): Indicates how much more likely it is to get a positive test in the diseased as opposed to the non-diseased group

Confidence interval: $LR+ = \exp(\ln \frac{sensitivity}{1-specificity} \pm 1.96 * \sqrt{\frac{1-sensitivity}{TP} + \frac{specificity}{FP}})$

Negative Likelihood Ratio (LR-): Indicates how much more likely it is to get a negative test in the non-diseased as opposed to the diseased group.

Confidence interval: $LR- = \exp(\ln \frac{1-sensitivity}{specificity} \pm 1.96 * \sqrt{\frac{sensitivity}{FN} + \frac{1-specificity}{TN}})$

Receiver Operator Curve Analysis

AUC analysis will be performed for the patient level analysis. Because the estimates of sensitivity and specificity will be constructed for the full patient population, only one estimate of sensitivity and one estimate of specificity will be generated. With only one estimate the sensitivity/specificity graphical methods to derive AUC are not applicable. Instead, the accepted method of estimating AUC will be determined by the nonparametric Wilcoxon approximation of the 2x2 table (which is statistically equivalent to the AUC generated with the trapezoid rule, and the Mann-Whitney U Test).

The degree of precision of the estimate of the AUC estimated will be reported by generating the standard error and 95% confidence interval around the estimate.

Area Under the Curve (AUC): represents the probability that a randomly chosen diseased subject is correctly diagnosed with greater suspicion than a randomly chosen non-diseased subject.

Wilcoxon AUC = $\frac{TN \times TP + 0.5 \times TN \times FN + 0.5 \times FP \times TP}{N_N \times N_A}$

Standard error (Hanley and McNeil method):

$$SE(A) = \sqrt{\frac{A(1-A) + (N_A - 1) * (Q_1 - A^2) + (N_N - 1) * (Q_2 - A^2)}{N_A * N_N}}$$

where A = AUC, area under the curve

N_A = number of positive disease cases

N_N = number of negative disease cases

$$Q1 = \frac{TN \times [TP^2 + TP \times FN + \frac{1}{3} \times FN^2] + FP \times [\frac{1}{3} \times TP^2]}{N_N \times N_A^2}$$

$$Q2 = \frac{FN \times [\frac{1}{3} \times TN^2] + TP \times [TN^2 + TN \times TP + \frac{1}{3} xFP^2]}{N_A \times N_N^2}$$

Example:

Overall			Total
	CICA: D+	CICA: D-	
64 CT: + test	183	22	205
64 CT: - test	2	219	221
Total	185	241	426

TP = 183, FP= 22, FN = 2, TN = 219

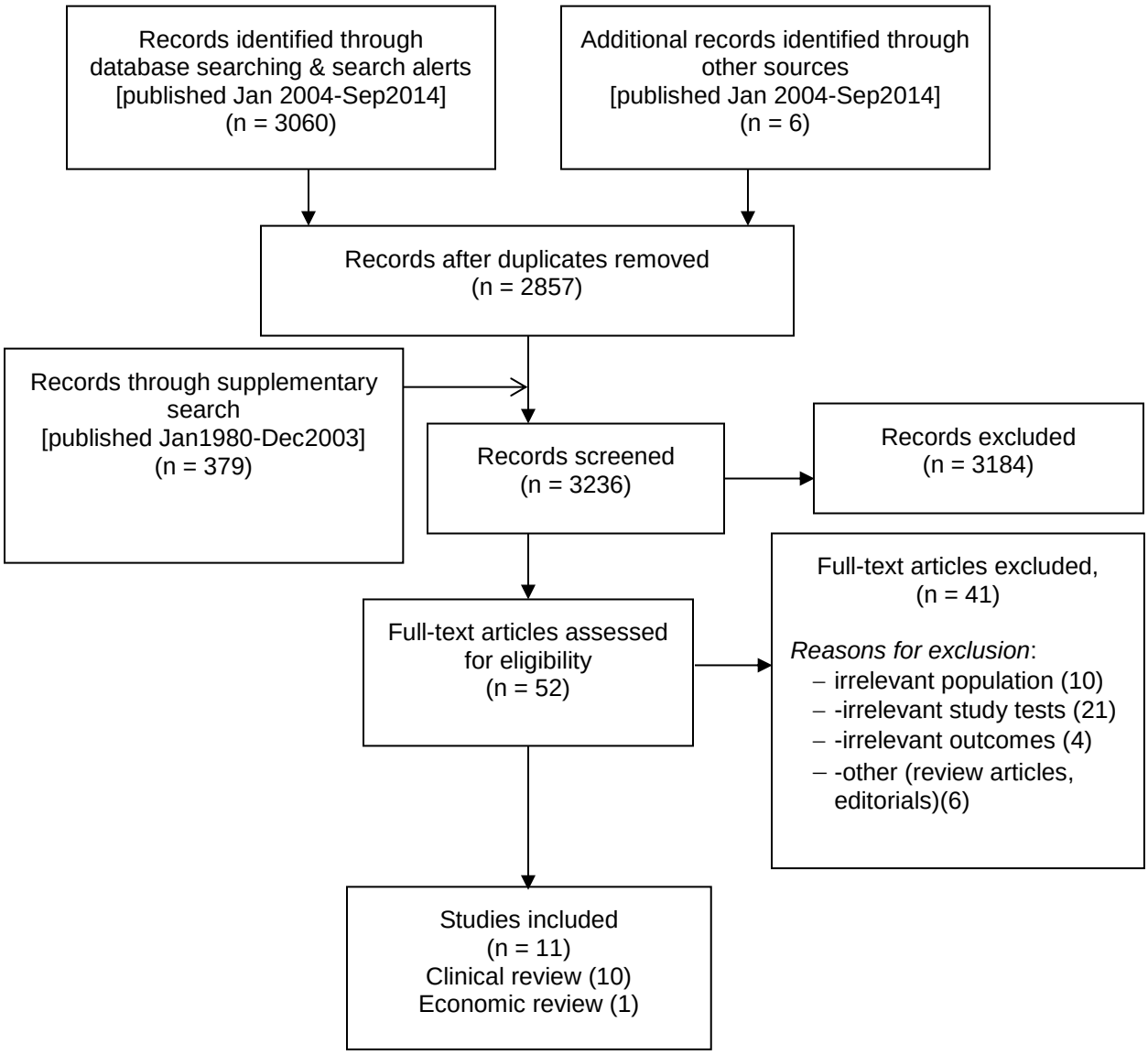
D+: disease positive, D-: disease negative (absent)

AUC = (219 x 183 + 0.5 x 219 x 2 + 0.5 x 22 x 183) / (185 x 241) = 0.9490
Similarly, Q1= 0.9287, Q2= 1.5051, SE =0.0581.
95% CI = (0.9490 – 1.96*0.0581, 0.9490 + 1.96* 0.0581) = (0.8351, 1)

Overall diagnostic accuracy: (TP/TN)/(TP+FP+TN+FN) The proportion of correctly classified subjects among all study participants.

Confidence interval: $p \pm Z * \sqrt{\frac{p * (1-p)}{TN + TP + + FN + FP}}$

APPENDIX 7: SELECTION OF INCLUDED STUDIES



APPENDIX 8: LIST OF INCLUDED AND EXCLUDED STUDIES

Included Studies:

Clinical review

- Costa CR, Johnson AJ, Naziri Q, Maralunda GA, Delanois RE, Mont MA. Efficacy of erythrocyte sedimentation rate and C-reactive protein level in determining periprosthetic hip infections. Am J Orthop (Belle Mead NJ). 2012 Apr;41(4):160-5.
- Kermani TA, Schmidt J, Crowson CS, Ytterberg SR, Hunder GG, Matteson EL, et al. Utility of erythrocyte sedimentation rate and C-reactive protein for the diagnosis of giant cell arteritis. Semin Arthritis Rheum. 2012 Jun;41(6):866-71.
- Robinson S, Leonard P. C reactive protein, erythrocyte sedimentation rate, or both, in the diagnosis of atraumatic paediatric limb pain? Emerg Med J. 2012 Dec;29(12):969-71.
- Grzesk E, Koltan S, Grzesk G, et al. Value of erythrocyte sedimentation rate, c-reactive protein and procalcitonin concentration versus multimarker strategy in management of bronchiolitis in pediatric emergency. Med Biol Sci. 2012;26.2:11-7.
- Johnson AJ, Zywiel MG, Stroh A, Marker DR, Mont MA. Serological markers can lead to false negative diagnoses of periprosthetic infections following total knee arthroplasty. Int Orthop. 2011 Nov;35(11):1621-6.
- 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 Mar;468(3):861-6.
- Ghanem E, Antoci V, Jr., Pulido L, Joshi A, Hozack W, Parvizi J. The use of receiver operating characteristics analysis in determining erythrocyte sedimentation rate and C-reactive protein levels in diagnosing periprosthetic infection prior to revision total hip arthroplasty. Int J Infect Dis. 2009 Nov;13(6):e444-e449.
- Austin MS, Ghanem E, Joshi A, Lindsay A, Parvizi J. A simple, cost-effective screening protocol to rule out periprosthetic infection. J Arthroplasty. 2008 Jan;23(1):65-8.
- Dolwani S, Metzner M, Wassell JJ, Yong A, Hawthorne AB. Diagnostic accuracy of faecal calprotectin estimation in prediction of abnormal small bowel radiology. Aliment Pharmacol Ther. 2004 Sep 15;20(6):615-21.
- Hayreh SS, Podhajsky PA, Raman R, Zimmerman B. Giant cell arteritis: validity and reliability of various diagnostic criteria. Am J Ophthalmol 1997;123(3):285-96.

Economic review

- Takemura Y, Ishida H, Inoue Y. Utilization of common inflammatory markers in new, symptomatic, primary care outpatients based on their cost-effectiveness. Clin Chem Lab Med 2003;41(5):668-74.

Excluded Studies

Irrelevant study population

- Jee WH, McCauley TR, Lee SH, Kim SH, Im SA, Ha KY. Sacroiliitis in patients with ankylosing spondylitis: association of MR findings with disease activity. Magn Reson Imaging. 2004 Feb; 22(2):245-50.
- Lindqvist E, Eberhardt K, Bendtzen K, Heinegard D, Saxne T. Prognostic laboratory markers of joint damage in rheumatoid arthritis. Ann Rheum Dis. 2005 Feb;64(2):196-201.
- Jung SY, Park MC, Park YB, Lee SK. Serum amyloid a as a useful indicator of disease activity in patients with ankylosing spondylitis. Yonsei Med J. 2007 Apr 30;48(2):218-24.

- Romero-Sanchez C, Robinson WH, Tomooka BH, Londono J, Valle-Onate R, Huang F, et al. Identification of acute phase reactants and cytokines useful for monitoring infliximab therapy in ankylosing spondylitis. Clin Rheumatol. 2008 Nov;27(11):1429-35.
- Gonzalez-Gay MA, Garcia-Unzueta MT, Gonzalez-Juanatey C, Miranda-Filloy JA, Vazquez-Rodriguez TR, De Matias JM, et al. Anti-TNF-alpha therapy modulates resistin in patients with rheumatoid arthritis. Clin Exp Rheumatol. 2008 Mar;26(2):311-6.
- Crowson CS, Rahman MU, Matteson EL. Which measure of inflammation to use? A comparison of erythrocyte sedimentation rate and C-reactive protein measurements from randomized clinical trials of golimumab in rheumatoid arthritis. J Rheumatol. 2009 Aug;36(8):1606-10.
- de Vries MK, van Eijk IC, van der Horst-Bruinsma IE, Peters MJ, Nurmohamed MT, Dijkmans BA, et al. Erythrocyte sedimentation rate, C-reactive protein level, and serum amyloid a protein for patient selection and monitoring of anti-tumor necrosis factor treatment in ankylosing spondylitis. Arthritis Rheum. 2009 Nov 15;61(11):1484-90.
- Pongprasobchai S, Jianjaroonwong V, Charatcharoenwitthaya P, Komoltri C, Tanwandee T, Leelakusolvong S, et al. Erythrocyte sedimentation rate and C-reactive protein for the prediction of severity of acute pancreatitis. Pancreas. 2010 Nov;39(8):1226-30.
- Turner D, Mack DR, Hyams J, LeLeiko N, Otley A, Markowitz J, et al. C-reactive protein (CRP), erythrocyte sedimentation rate (ESR) or both? A systematic evaluation in pediatric ulcerative colitis. J Crohns Colitis. 2011 Oct;5(5):423-9.
- Ye M, Yu H, Yu W, Zhang G, Xiao L, Zheng X, et al. Evaluation of the significance of circulating insulin-like growth factor-1 and C-reactive protein in patients with chronic obstructive pulmonary disease. J Int Med Res. 2012;40(3):1025-35.

Irrelevant Study Tests

- Jeremiah Z, Leonard I, Ezinma A. Discordantly elevated Erythrocyte Sedimentation Rate (ESR) and depressed C-Reactive Protein (CRP) values in early diagnosis of pulmonary tuberculosis patients in Maiduguri, Nigeria. Open J Blood Dis. 2013;3(2):74-7. Available from: http://www.scirp.org/journal/PaperInformation.aspx?paperID=33616
- CRP vs ESR assessing & measuring the inflammatory response. Dunedin (New Zealand): bpacnz; 2005. Available from: http://www.bpac.org.nz/resources/campaign/crp_esr/crp_esr_poem.asp; www.bpac.org.nz/resources/other/audits/bpac_crp_audit_wv.pdf
- Laboratory tests in general practice. Brussels: Belgian Health Care Knowledge Centre; 2007. Available from: https://kce.fgov.be/publication/report/laboratory-tests-in-general-practice
- Wang C, Ao Y, Fan X, Wang J, Cui G, Hu Y, et al. C-Reactive Protein and Erythrocyte Sedimentation Rate Changes After Arthroscopic Anterior Cruciate Ligament Reconstruction: Guideline to Diagnose and Monitor Postoperative Infection. Arthroscopy. 2014 May 23.
- Parikh M, Miller NR, Lee AG, Savino PJ, Vacarezza MN, Cornblath W, et al. Prevalence of a normal C-reactive protein with an elevated erythrocyte sedimentation rate in biopsy-proven giant cell arteritis. Ophthalmology. 2006 Oct;113(10):1842-5.
- Kofla-Dlubacz A, Matusiewicz M, Krzystek-Korpacka M, Iwanczak B. Correlation of MMP-3 and MMP-9 with Crohn's disease activity in children. Dig Dis Sci. 2012 Mar;57(3):706-12.
- Khan MH, Smith PN, Rao N, Donaldson WF. Serum C-reactive protein levels correlate with clinical response in patients treated with antibiotics for wound infections after spinal surgery. Spine J. 2006 May;6(3):311-5.

- Iyengar KP, Vinjamuri S. Role of 99mTc Sulesomab in the diagnosis of prosthetic joint infections. Nucl Med Commun. 2005 Jun;26(6):489-96.
- Ghanem E, Parvizi J, Burnett RS, Sharkey PF, Keshavarzi N, Aggarwal A, et al. Cell count and differential of aspirated fluid in the diagnosis of infection at the site of total knee arthroplasty. J Bone Joint Surg Am. 2008 Aug;90(8):1637-43.
- Costello F, Zimmerman MB, Podhajsky PA, Hayreh SS. Role of thrombocytosis in diagnosis of giant cell arteritis and differentiation of arteritic from non-arteritic anterior ischemic optic neuropathy. Eur J Ophthalmol. 2004 May;14(3):245-57.
- Bottner F, Wegner A, Winkelmann W, Becker K, Erren M, Gotze C. Interleukin-6, procalcitonin and TNF-alpha: markers of peri-prosthetic infection following total joint replacement. J Bone Joint Surg Br. 2007 Jan;89(1):94-9.
- Ayazi P, Mahyar A, Daneshi MM, Jahani HH, Pirouzi M, Esmailzadehha N. Diagnostic Accuracy of the Quantitative C-Reactive Protein, Erythrocyte Sedimentation Rate and White Blood Cell Count in Urinary Tract Infections among Infants and Children. Malays J Med Sci. 2013 Oct; 20(5):40-6.
- Schinsky MF, Della Valle CJ, Sporer SM, Paprosky WG. Perioperative testing for joint infection in patients undergoing revision total hip arthroplasty. J Bone Joint Surg Am. 2008 Sep;90(9):1869-75.
- Mondal SK, Nag DR, Bandyopadhyay R, Chakraborty D, Sinha SK. Neonatal sepsis: Role of a battery of immunohematological tests in early diagnosis. Int J Appl Basic Med Res. 2012 Jan;2(1):43-7.
- Korppi M. Non-specific host response markers in the differentiation between pneumococcal and viral pneumonia: what is the most accurate combination? Pediatr Int. 2004 Oct;46(5):545-50.
- Don M, Valent F, Korppi M, Canciani M. Differentiation of bacterial and viral community-acquired pneumonia in children. Pediatr Int. 2009 Feb;51(1):91-6.
- Ahmed Z, Ghafoor T, Waqar T, Ali S, Aziz S, Mahmud S. Diagnostic value of C- reactive protein and haematological parameters in neonatal sepsis. J Coll Physicians Surg Pak. 2005 Mar;15(3):152-6.

Irrelevant Outcomes

- Houshian S, Seyedipour S, Wedderkopp N. Epidemiology of bacterial hand infections. Int J Infect Dis. 2006 Jul;10(4):315-9.
- Gonzalez-Juanatey C, Llorca J, Miranda-Filloy JA, Amigo-Diaz E, Testa A, Garcia-Porrua C, et al. Endothelial dysfunction in psoriatic arthritis patients without clinically evident cardiovascular disease or classic atherosclerosis risk factors. Arthritis Rheum. 2007 Mar 15;57(2):287-93.
- Kainth MK, Gigliotti F. Simultaneous Testing of Erythrocyte Sedimentation Rate and C-Reactive Protein: Increased Expenditure without Demonstrable Benefit. J Pediatr. 2014 Jun 21.
- Schwarzkopf R, Mikhael B, Wright E, Estok DM, Katz JN. Treatment failure among infected periprosthetic total hip arthroplasty patients. Open Orthop J. 2014;8:118-24.
- *Irrelevant Study Type*
- Azhar SS, Tang RA, Dorotheo EU. Giant cell arteritis: diagnosing and treating inflammatory disease in older adults. Geriatrics. 2005 Aug;60(8):26-30.
- Batlivala SP. Focus on diagnosis: the erythrocyte sedimentation rate and the C-reactive protein test. Pediatr Rev. 2009 Feb;30(2):72-4.
- da Mota LM, dos Santos Neto LL, de Carvalho JF. Autoantibodies and other serological markers in rheumatoid arthritis: predictors of disease activity? Clin Rheumatol. 2009 Oct;28(10):1127-34.

- Jellema P, van Tulder MW, van der Horst HE, Florie J, Mulder CJ, van der Windt DA. Inflammatory bowel disease: a systematic review on the value of diagnostic testing in primary care. Colorectal Disease. 2011;13(3):239-54.
- Looking back, looking forward: key messages 2005 - 2011. Dunedin (New Zealand): bpacnz; 2012. Available from: http://www.bpac.org.nz/Report/2012/March/reminders.aspx
- Guerado E, Cervan AM. Surgical treatment of spondylodiscitis. An update. Int Orthop. 2012 Feb;36(2):413-20.

APPENDIX 9: CHARACTERISTICS OF THE INCLUDED STUDIES

Author, Year	Country	Study Design	Setting	Study Population				Target Condition
				N	Eligibility criteria	Mean Age	Gender (%female)	
Costa, 2012[18]	USA	Diagnostic Accuracy (prospective)	Single hospital	77	Inclusion: Patients undergoing revision THA between 2000 and 2008 Exclusion: Not stated	61 Year (19-89 range)	58.4%	Periprosthetic infection (after THA)
Kermani, 2012[19]	USA	Diagnostic Accuracy (retrospective)	University hospital	764	Inclusion: patients undergoing temporal artery biopsy (TAB) between 2000 and 2008 Exclusion: Pts who didn't have both ESR and CRP measured 6weeks before TAB	72.7 Year (9.27 SD)	65%	Giant cell arteritis
Robinson, 2012[20]	UK	Diagnostic Accuracy (retrospective)	Teaching hospital-pediatrics emergency room	259	Inclusion: Patients <=13 years presenting with atraumatic limb pain who had both ESR and CRP measured at presentation Exclusion: Previous orthopedic infection, chronic recurrent multifocal osteomyelitis, fractures and chronic disease causing limb pain	NA	NA	Pediatric orthopedic infection
Grzesk, 2012[27]	Poland	Diagnostic Accuracy (prospective) Internal comparison of two subgroups	Single hospital	149	Inclusion: children hospitalized with clinical presentation of bronchiolitis. Exclusion: bronchial asthma, cystic fibrosis, pulmonary bronchodysplasts, congenital heart diseases, abnormalities of chest and lung, history of treatment with bronchodilators and anti-inflammatory drugs, and gastro-esophageal reflux,	7 Months 1-24 range	31.5%	Pediatric bronchiolitis
Johnson, 2011[21]	USA	Diagnostic Accuracy (prospective)	Single hospital	113	Inclusion: Patients undergoing revision of TKA between 2000 and 2007 Exclusion: Patients with strong evidence of an infected joint	61 Year (28-89 range)	46%	Periprosthetic infection (after TKA)
Paakkonen, 2010[22]	Finland	Diagnostic Accuracy (prospective)	Multiple referral hospitals	265	Inclusion: Patients 3months to 15 years old presenting with signs and symptoms suggesting acute osteoarticular infection Exclusion: Culture negative cases; less than 3 months of age; and immune deficient patients	NA	NA	Pediatrics septic bone and joint infection
Ghanem, 2009[25]	USA	Diagnostic Accuracy (retrospective)	Single hospital	479	Inclusion: Patients undergoing revision THA between 2000 and 2005 Exclusion: Patients with comorbidities that could elevate ESR and CRP values, such as inflammatory diseases, chronic renal failure, hepatitis, active malignancy or infection in other regions of the body	66 Year (23-93 range)	53%	Periprosthetic infection (after THA)
Austin, 2008[23]	USA	Diagnostic Accuracy (prospective)	University hospital	296	Inclusion: Patients undergoing revision TKA between 2000 and 2005, who had pre-	66 Year	57.8%	Periprosthetic infection after (TKA)

Author, Year	Country	Study Design	Setting	Study Population				Target Condition
				N	Eligibility criteria	Mean Age	Gender (%female)	
					operative ESR and CRP tests and pre- or intra-operative cultures performed. Exclusion: Patients with confounding factors that could elevate ESR and CRP values.			
Dolwani, 2004[24]	UK	Diagnostic Accuracy (Case control design- 63 undifferentiated patients with 25 controls with active Crohn's and 26 healthy individuals)	University hospital	63	Inclusion: Out-patient or in-patient cases, presenting with abdominal pain and diarrhea, who underwent a small bowel barium follow through test. Exclusion: Patients with known malignancy, those on NSAIDS or steroids, known cases of celiac disease, severe cardiomyopathy, renal or hepatic impairment, significant psychiatric disease, and alcohol or drug dependency. Patients with abnormal rigid sigmoidoscopy or positive stool culture were excluded from the "case" group.	47 Year (17-86 range)	68%	Prediction of abnormal small bowel radiology for inflammatory bowel diseases
Hayreh, 1997[28]	USA	Diagnostic Accuracy (Case control design- 363 cases and 749 controls)	University hospitals and clinics	363 (total, ESR, sample) 223 (ESR and CRP sub-sample) §	Inclusion: Cases: patients referred for temporal artery biopsy from 1973-1994; Controls: otherwise healthy patients with non-arteritic anterior ischemic optic neuropathy or retinal nerve occlusion. Exclusion: patients with any systemic abnormalities that might elevate ESR or CRP values.	72 Year (20-95 range)	66%	Giant cell arteritis

Abbreviations: CRP= C-reactive protein; ESR= erythrocyte sedimentation rate; NA= not available; NSAIDs= nonsteroidal anti-inflammatory drugs; SD= standard deviation; THA= total hip arthroplasty; TKA= total knee arthroplasty; UK= the United Kingdom; USA= the United States of America

†Proportion of the patients who received confirmation of the diagnosis (disease positive) by the reference standard.

*This study only analyzed the cases that were diagnosed as "disease positive" by the reference standard.

§CRP testing was available for a subset of patients who were recruited from 1985 to 1994 (n=223)

APPENDIX 10: RISK OF BIAS AND APPLICABILITY IN THE INCLUDED DIAGNOSIS STUDIES (RESULTS OF QUADAS 2 QUALITY ASSESSMENT)[12]

Study	Risk of Bias				Applicability Concerns			Potential issues
Author, Year	Patient selection	Index Test	Reference Standard	Follow & Timing	Patient selection	Index Test	Reference Standard	
Costa, 2012[18]	Unclear	L	L	L	H	L	L	• Not all of potentially eligible cases were included in the study. • Cases with a lower likelihood of infection were excluded from ESR and CRP testing and, therefore, from the analysis. • 47% of the non-infected cases had other disorders elevating CRP results.
Kermani, 2012[19]	Unclear	L	Unclear	L	L	L	L	• Based on the author's discussion, there was a risk that patient's with negative biopsy GCA had been misclassified as non-disease. • Authors also stated that due to the performing the study in a referral hospital, the study population might not be the representative of all patients with GCA.
Robinson, 2012[20]	Unclear	H	H	L	H	L	Unclear	• Retrospective chart review: only patients who had blood test results recorded were included in the study. • No pre-specified cut-off-points were noted for ESR and CRP. • No gold standard was used to confirm the final diagnosis. Therefore, the reference standard included a description of patient's complaint; and a history of referral to the orthopedic team, diagnosis of infection and antibiotic-therapy. • It is unclear whether the reference test was interpreted independent of the index tests (ESR and CRP).
Grzesk, 2012[27]	Unclear	L	L	Unclear	H	L	unclear	• Not all of potentially eligible cases were included in the study. • It is unclear whether the reference test was interpreted independent of the index tests (ESR and CRP). • Broad exclusion criteria were applied to exclude patients with disorders that might elevate inflammatory markers. • It is unclear whether the reference standard was performed in the same way for all study participants.

Study	Risk of Bias				Applicability Concerns			Potential issues
Author, Year	Patient selection	Index Test	Reference Standard	Follow & Timing	Patient selection	Index Test	Reference Standard	
Johnson, 2011[21]	Unclear	L	L	L	L	L	L	• Not all of potentially eligible cases were included in the study.
Paakkonen, 2010[22]	H	L	L	Unclear	Unclear	L	L	• Study population consisted of culture (reference standard)-positive cases, i.e. culture-negative cases were excluded from ESR/CRP testing and, hence, from the analysis. • The study excluded immune-deficient patients. • The final diagnosis of various infection types was made using different clinical tests. • The schedules of measuring ESR and CRP differed.
Ghanem, 2009[25]	Unclear	L	L	L	L	L	L	• Retrospective chart review • Broad exclusion criteria were applied to exclude patients with disorders that might elevate inflammatory markers.
Austin, 2008[23]	Unclear	L	L	L	L	L	L	• Not all of potentially eligible cases were included in the study, i.e. only patients who had both ESR and CRP measurements were included. • Broad exclusion criteria were applied to exclude patients with disorders that might elevate inflammatory markers.
Dolwani, 2004[24]	H	L	L	Unclear	L	L	L	• The study used a case-control design to estimate the diagnostic performance of the study tests. • Patients who were diagnosed as positive by rigid sigmoidoscopy or stool culture tests were excluded from the "case" group, with no explanation of as to why. • Unclear description of scheduling of study tests in each study group. • In addition to a control group of patients with Crohn's disease, the study included a control group of healthy volunteers for whom ESR and CRP tests were not performed.
Hayreh, 1997[28]	H	L	L	H	L	L	L	• The study used a case-control design to estimate the diagnostic performance of the study tests. • Patients with disorders that might elevate inflammatory markers were excluded. • Unclear description of scheduling of study tests in each study group. • Only a subset of patients, who received both ESR

Study	Risk of Bias				Applicability Concerns			Potential issues
Author, Year	Patient selection	Index Test	Reference Standard	Follow & Timing	Patient selection	Index Test	Reference Standard	
								and CRP tests, were included in the analysis of combined ESR and CRP test results.

Abbreviations: CRP= C-reactive protein; ESR= erythrocyte sedimentation rate; GCA= giant cell arteritis; H= high; L= low

APPENDIX 11: TEST CHARACTERISTICS AND ESTIMATED TEST RESULTS USING DATA FROM THE INCLUDED STUDIES

Study	Test	Cut-off point	Reference Standard	Total sample size	Disease frequency (%)†	TP	FP	TN	FN
Diagnosis of Periprosthetic Infection									
Costa, 2012[18]	ESR	>30mm/h	Synovial fluid bacterial culture	77	89.61	61	3	5	8
	CRP	>10mg/L				64	5	3	5
	ESR+CRP	Same as above				58	2	6	11
	ESR/CRP	Same as above				65	8	2	2
Johnson, 2011[21]	ESR	>30mm/h	At least one of the following: (1) ≥2 positive cultures with the same organism, (2) positive histological findings for acute inflammatory response on intraoperative frozen section, (3) macroscopic purulence, or (4) a draining sinus tract communicates with the joint space.	113	84.07	86	12	6	9
	CRP	>10mg/L				90	14	4	5
	ESR+CRP	Same as above				85	13	5	10
	ESR/CRP	Same as above				90	11	7	5
Ghanem, 2009[25]	ESR	>30mm/h	At least one of the following: (1) an abscess or sinus tract that communicates with the joint space, (2) positive preoperative aspiration culture, (3) ≥2 positive intra-operative cultures or one positive culture and the presence of other indicators of infection includes macroscopic purulence or an elevated cell count and differential of the aspirate fluid.	479	26.50	120	105	247	7
	CRP	>10mg/L				116	82	270	11
	ESR+CRP	Same as above				111	42	310	16
	ESR/CRP	Same as above				124	145	207	3
Austin, 2008[23]	ESR	>30mm/h	Synovial fluid culture	296	39.19	106	50	130	10
	CRP	>10mg/L				109	47	133	7
	ESR+CRP	Not applicable				NA	NA	NA	NA
	ESR/CRP	Same as above				111	79	101	5
Diagnosis of Pediatric Orthopedic Infections									
Robinson, 2012[20]	ESR	>12mm/h	Diagnosis of infection by an orthopedic consultant and history of antibiotic-therapy	259	6.56	15	83	159	2
	CRP	>7mg/L				12	32	210	5
	ESR+CRP	Same as above				11	25	217	6
	ESR/CRP	Same as above				NA	NA	NA	NA
Paakkonen, 2010[22]	ESR	>20mm/h	Blood culture, bone/joint needle aspiration and culture	NA*	NA*	248	NA	NA	17
	CRP	>20mg/L				251	NA	NA	14

Study	Test	Cut-off point	Reference Standard	Total sample size	Disease frequency (%)†	TP	FP	TN	FN
	ESR+CRP	Not applicable				NA	NA	NA	NA
	ESR/CRP	Same as above				261	NA	NA	4
Diagnosis of Pediatric Bronchiolitis									
Grzesk, 2012[27]	ESR	>15mm/h	Clinical symptoms, chest x-ray, white blood cell count, and bacteriological tests (for suspected lower respiratory tract infections)	149	11 (bacterial infection)	NA	NA	NA	NA
	CRP	>15mg/L				NA	NA	NA	NA
	ESR+CRP	Same as above				NA	NA	NA	NA
	ESR/CRP	Not applicable				NA	NA	NA	NA
Diagnosis of Inflammatory Bowel Diseases									
Dolwani, 2004[24]	ESR	>=10mm/h	Barium Follow Through (BaFT)	63	23.81	12	16	32	3
	CRP	>=6mg/L				12	14	34	3
	ESR+CRP	Same as above				8	8	40	7
	ESR/CRP	Not applicable				NA	NA	NA	NA
Diagnosis of Giant Cell Arteritis									
Kermani, 2012[19]	ESR	>22mm/h (men); >29mm/h (women)	Temporal artery biopsy	764	23.17	149	414	173	28
	CRP	>8mg/L				153	408	179	24
	ESR+CRP	Same as above				143	345	242	34
	ESR/CRP	Not applicable				NA	NA	NA	NA
Hayreh, 1997[28]	ESR	>10mm/h (men); >20mm/h (women)	Temporal artery biopsy	855 (ESR sample)§ 181 (CRP and ESR sub-sample)‡	29.20 (total, ESR, sample) 19.28 (CRP and ESR sub-sample)	103	247	502	3
	CRP	>5mg/L				43	25	113	0
	ESR+CRP	Same as above				42	11	127	1
	ESR/CRP	Not applicable				NA	NA	NA	NA

Abbreviations: CRP= C-reactive protein; ESR= erythrocyte sedimentation rate; FN= false negative; FP= false positive; mg/L= milligram per liter; mm/h= millimeter per hour; NA= not available; TN= true negative; TP= true positive

†Proportion of the patients who received confirmation of the diagnosis (disease positive) by the reference standard.

*****This study only analyzed the cases who were diagnosed as "disease positive" by the reference standard.

§ Diagnostic accuracy measures were calculated using the data from patients with positive temporal artery biopsy (n=106 out of 363) and the control group (n=749).

‡ Diagnostic accuracy measures were calculated using the data from patients with positive temporal artery biopsy (n=43 out of 223) and the corresponding control group (n=138).

APPENDIX 12: DIAGNOSTIC PERFORMANCE OF ESR AND CRP INDIVIDUALLY AND IN COMBINATION AS COMPARED TO THE REFERENCE STANDARD

Study	Test	Sensitivity* (95% CI)	Specificity* (95% CI)	PPV* (95% CI)	NPV* (95% CI)	LR(+)* (95% CI)	LR(-)* (95% CI)	AUC* (95% CI)	ODA* (95% CI)
Diagnosis of Periprosthetic Infection									
Costa, 2012[18]	ESR	**0.89**** (0.81, 0.99)	0.63** (0.29, 0.96)	0.95 0.90, 1.00)	0.38 (0.12, 0.65)	2.36 (1.01, 5.48)	0.19 (0.08, 0.46)	0.755** (0.273, 1.236)	0.86 (0.78, 0.94)
	CRP	**0.93**** (0.87, 0.99)	0.38** (0.40, 0.71)	0.93 (0.87, 0.99)	0.38 (0.04, 0.71)	1.48 (0.43, 5.08)	0.19 (0.11, 0.33)	0.651** (0.241, 1.062)	0.87 (0.80, 0.95)
	ESR+CRP	**0.84**** (0.75, 0.98)	0.75** (0.45, 1.00)	0.97 (0.92, 1.01)	0.35 (0.13, 0.58)	3.36 (1.71, 6.59)	0.21 (0.06, 0.71)	0.795 (0.303, 1.288)	0.83 (0.75, 0.91)
	ESR/CRP	**0.97**** **(0.93, 1.00)**	0.20** (0.00, 0.45)	0.89 (0.82, 0.96)	0.50 (0.01, 0.99)	1.21 (0.19, 7.67)	0.15 (0.11, 0.20)	0.585 (0.280, 0.890)	0.87 (0.80, 0.95)
Johnson, 2011[21]	ESR	**0.91 (0.83, 0.95)**	**0.33 (0.12, 0.65)**	0.88 (0.81, 0.94)	0.40 (0.15, 0.65)	1.36 (0.55, 3.35)	0.28 (0.20, 0.40)	0.619 (0.355, 0.884)	0.81 (0.74, 0.89)
	CRP	**0.95 (0.88, 0.98)**	**0.20 (0.06, 0.50)**	0.87 (0.80, 0.93)	0.44 (0.12, 0.77)	1.22 (0.36, 4.10)	0.24 (0.18, 0.30)	0.585 (0.347, 0.822)	0.83 (0.76, 0.90)
	ESR+CRP	**0.89 (0.87, 0.97)**	**0.29 (0.10, 0.59)**	0.87 (0.80, 0.93)	0.33 (0.09, 0.57)	1.24 (0.48, 3.20)	0.38 (0.28, 0.51)	0.586 (0.341, 0.831)	0.80 (0.72, 0.87)
	ESR/CRP	**0.95 (0.89, 0.98)**	**0.38 (0.14, 0.69)**	0.89 (0.83 ,0.95)	0.58 (0.30, 0.86)	1.55 (0.55, 4.35)	0.14 (0.09, 0.20)	0.668 (0.377, 0.959)	0.86 (0.79, 0.92)
Ghanem, 2009[25]	ESR	**0.94 (0.89, 0.98)**	**0.70 (0.65, 0.75)**	**0.56 (0.48, 0.62)**	**0.97 (0.94, 0.99)**	**3.10 (0.70, 7.90)**	**0.09 (0.03, 0.27)**	0.823 (0.776, 0.871)	0.77 (0.73, 0.80)
	CRP	**0.91 (0.85, 0.95)**	**0.77 (0.72, 0.81)**	**0.61 (0.53, 0.68)**	**0.96 (0.92, 0.98)**	**4.10 (2.30, 7.20)**	**0.13 (0.05, 0.38)**	0.840 (0.781, 0.899)	0.81 (0.77, 0.84)
	ESR+CRP	**0.88 (0.81, 0.93)**	**0.88 (0.84, 0.92)**	**0.75 (0.66, 0.81)**	**0.95 (0.92, 0.97)**	**7.50 (2.40, 18.30)**	**0.20 (0.09, 0.38)**	0.877 (0.800, 0.954)	0.88 (0.85, 0.91)
	ESR/CRP	**0.98 0.93, 0.99)**	**0.59 (0.53, 0.64)**	**0.48 (0.42, 0.54)**	**0.98 (0.95, 0.99)**	**2.50 (1.50, 4.40)**	**0.04 (0.01, 0.16)**	0.782 (0.759, 0.806)	0.69 (0.65, 0.73)

Study	Test	Sensitivity* (95% CI)	Specificity* (95% CI)	PPV* (95% CI)	NPV* (95% CI)	LR(+)* (95% CI)	LR(-)* (95% CI)	AUC* (95% CI)	ODA* (95% CI)
Austin, 2008[23]	ESR	**0.91** (0.86, 0.97)	**0.72** (0.66, 0.79)	**0.68** (0.61, 0.75)	**0.93** (0.89, 0.97)	3.29 (1.81, 5.99)	0.12 (0.09, 0.15)	0.818 (0.722, 0.914)	0.80 (0.75, 0.84)
	CRP	**0.94** (0.90, 0.98)	**0.74** (0.68, 0.80)	**0.70** (0.63, 0.77)	**0.95** (0.91, 0.99)	3.60 (1.75, 7.42)	0.08 (0.06, 0.10)	0.839 (0.739, 0.939)	0.82 (0.77, 0.8)
	ESR+CRP	NA	NA	NA	NA	NA	NA	NA	NA
	ESR/CRP	**0.96** (0.92, 0.99)	**0.56** (0.49, 0.63)	**0.58** (0.51, 0.65)	**0.95** (0.91, 0.99)	2.18 (0.92, 5.19)	0.08 (0.06, 0.09)	0.759 (0.681, 0.837)	0.72 (0.66, 0.77)
Diagnosis of Pediatric Orthopedic Infections									
Robinson, 2012[20]	ESR	**0.88** (0.73, 1.04)	**0.66** (0.60, 0.72)	0.15 (0.08, 0.22)	0.99 (0.97, 1.00)	**2.57** (0.70, 0.949)	0.18 (0.14, 0.23)	0.770 (0.591, 0.948)	0.67 (0.61, 0.73)
	CRP	**0.71** (0.45, 0.92)	**0.87** (0.83, 0.91)	0.27 (0.14, 0.40)	0.98 (0.96, 1.00)	**5.34** (2.55, 11.17)	0.34 (0.22, 0.53)	0.787 (0.743, 0.831)	0.86 (0.81, 0.90)
	ESR+CRP	0.65 **(0.42, 0.87)**	0.90 (0.86, 0.94)	0.31 (0.16, 0.46)	0.97 (0.95, 0.99)	**6.26** (3.29, 11.94)	0.39 (0.24, 0.66)	0.772 (0.708, 0.835)	0.88 (0.84, 0.92)
	ESR/CRP	NA	NA	NA	NA	NA	NA	NA	NA
Paakkonen, 2010[22]	ESR	**0.94 (0.90, 0.96)**	NA	NA	NA	NA	NA	NA	NA
	CRP	**0.95 (0.91, 0.97)**	NA	NA	NA	NA	NA	NA	NA
	ESR+CRP	NA	NA	NA	NA	NA	NA	NA	NA
	ESR/CRP	**0.98 (0.96, 0.99)**	NA	NA	NA	NA	NA	NA	NA
Diagnosis of Pediatric Bronchiolitis									
Grzesk, 2012[27]	ESR	NA	NA	NA	NA	NA	NA	0.710 (0.600, 0.830)	NA
	CRP	NA	NA	NA	NA	NA	NA	0.630 (0.510, 0.750)	NA
	ESR+CRP	NA	NA	NA	NA	NA	NA	0.74 (0.600, 0.880)	NA

Study	Test	Sensitivity* (95% CI)	Specificity* (95% CI)	PPV* (95% CI)	NPV* (95% CI)	LR(+)* (95% CI)	LR(-)* (95% CI)	AUC* (95% CI)	ODA* (95% CI)
	ESR/CRP	NA	NA	NA	NA	NA	NA	NA	NA
Diagnosis of Inflammatory Bowel Diseases									
Dolwani, 2004[24]	ESR	**0.79** (0.60, 1.00)	**0.67** (0.53, 0.80)	**0.42** (0.25, 0.61)	**0.91** (0.82, 1.01)	2.40 (0.86, 6.73)	0.30 (0.19, 0.48)	0.733 (0.662, 0.805)	0.70 (0.59, 0.81)
	CRP	**0.77** (0.60, 1.00)	**0.70 (0.58, 0.84)**	0.46† (0.27, 0.65)	**0.91** (0.83, 1.01)	2.74 (0.98, 7.67)	0.28 (0.17, 0.47)	0.754 (0.659, 0.850)	0.73 (0.62, 0.84)
	ESR+CRP	**0.50** (0.28, 0.79)	**0.84** (0.73, 0.94)	**0.5** (0.26, 0.75)	**0.84** (0.75, 0.95)	3.20 (1.84, 5.58)	0.56 (0.25, 1.23)	0.683 (0.531, 0.836)	0.76 (0.66, 0.87)
	ESR/CRP	NA	NA	NA	NA	NA	NA	NA	NA
Diagnosis of Giant Cell Arteritis									
Kermani, 2012[19]	ESR	**0.84** (0.79, 0.90)	**0.30** (0.26, 0.33)	**0.26** (0.23, 0.30)	0.86 (0.81, 0.91)	1.19 (0.83, 1.71)	0.54 (0.49, 0.58)	0.568 (0.551, 0.586)	0.42 (0.39, 0.46)
	CRP	**0.86** (0.81, 0.92)	**0.31** (0.27, 0.34)	**0.27** (0.24, 0.31)	0.88 (0.84, 0.93)	1.24 (0.84, 1.84)	0.44 (0.41, 0.48)	0.585 (0.564, 0.606)	0.43 (0.40, 0.47)
	ESR+CRP	**0.81** (0.75, 0.87)	**0.41** (0.37, 0.45)	**0.29** (0.25, 0.33)	0.88 (0.84, 0.92)	1.37 (1.00, 1.89)	0.47 (0.42, 0.51)	0.610 (0.588, 0.632)	0.50 (0.47, 0.54)
	ESR/CRP	NA	NA	NA	NA	NA	NA	NA	NA
Hayreh, 1997[28]	ESR	0.97 (0.92, 0.99)	0.67 (0.64, 0.72)	0.29 (0.25, 0.34)	0.99 (0.98, 1:00)	2.95 (0.96, 9.00)	0.04 (0.04, 0.05)	0.821 (0.821, 0.821)	0.71 (0.68, 0.74)
	CRP	1.00 (0.93, 1:00)	0.82 (0.74, 0.88)	0.63 (0.51, 0.75)	1:00 (0.97, 1:00)	5.52 (NA, NA)	0.00 (NA, NA)	0.909 (0.909, 0.909)	0.86 (0.81, 0.91)
	ESR+CRP	0.98 (0.88, 1:00)	0.92 (0.86, 0.96)	0.93 (0.81, 0.94)	0.99 (0.96, 1:00)	12.25 (1.76, 85.07)	0.03 (0.01, 0.04)	0.949 (0.949, 0.949)	0.93 (0.90, 0.97)
	ESR/CRP	NA	NA	NA	NA	NA	NA	NA	NA

Abbreviations: AUC= area under the curve; CRP= C-reactive protein; CI= confidence interval; ESR= erythrocyte sedimentation rate; LR(+)= positive likelihood ratio; LR(-)= negative likelihood ratio; NPV= negative predictive value; NA= not available; ODA= overall diagnostic accuracy; PPV= positive predictive value.

*****The estimates are calculated using available parameters from the included studies

Note: The estimates reported by the authors of the articles are presented in bold font, and the calculated measures are presented in normal font. When the reported values were different from our calculated values, only calculated values are in shown in the table, for consistency and reproducibility of the data.

****** The reported sensitivity and specificity values and their 95% confidence intervals (CI; calculated using Wilson score method) were reported as follows: ESR- sensitivity 0.89 (0.80, 0.94) and specificity =0.69 (0.51, 0.83); CRP- sensitivity= 0.93 (0.85, 0.97), and specificity = 0.40 (0.23, 0.59); ESR+CRP- sensitivity= 0.84 (0.73, 0.91), and specificity= 0.77 (0.57, 0.90), and ESR/CRP- sensitivity= 0.97 (0.90, 0.99), and specificity= 0.23 (0.10, 0.43).
† The specificity value for CRP, reported by the authors, was 0.42