The primary aim of this study was to assess the diagnostic accuracy of joint aspiration culture, serum C-reactive protein (CRP) and serum erythrocyte sedimentation rate (ESR), individually, and in combination for the diagnosis of periprosthetic joint infection (PJI).

A consecutive patient series with pre-operative inflammatory marker levels, an aspiration culture of either hip or knee arthroplasty and intra-operative culture samples from subsequent revision surgery was compiled. This retrospective patient cohort analysis included 128 aspiration. The data were analysed to compare pre-operative aspiration cultures, serum ESR and CRP levels to the chosen gold standard for PJI diagnosis of intra-operative culture samples. A diagnostic algorithm was created using the above tests combined with clinical suspicion index.

The values that had the highest sensitivity and specificity of predicting PJI were >5 for CRP and >16 for ESR. CRP used individually had the highest sensitivity and negative predictive value (NPV) of any test (75.0% and 75.9%, respectively). ESR + aspirate had the highest specificity and positive predictive value (PPV), of 100% for both. Using all three tests together the specificity and PPV were higher than the test individual values (95.3% and 85.0% respectively).

Based on subgroup analyses the combination of ESR or CRP plus joint aspiration has superior PPV compared to individual tests. ESR and CRP had the highest NPV when used in isolation. An algorithm has been developed to guide clinical diagnosis.

Keywords: Periprosthetic joint infection; C-reactive protein (CRP); Erythrocyte sedimentation rate (ESR); joint aspiration; diagnostic algorithm.

INTRODUCTION

Periprosthetic joint infection (PJI) is the second most common complication of joint arthroplasty (1), and has a significant financial burden on healthcare services (2). Early diagnosis may improve outcome. Clinical history in conjunction with inflammatory markers remain the cornerstone of assessment (3,4). Currently, these markers include erythrocyte sedimentation rate (ESR), C-reactive protein (CRP) levels and serum white cell count (5-7). However, serum inflammatory markers may be non-specifically high in non-infectious inflammatory processes (8). Diagnostic pre-operative synovial fluid aspirates have sensitivities ranging from (50-93%) and specificities from (82-97%) (9) combined with a modest to high incidence of false-positive culture results (3-16%) (9,10).

The American Academy of Hip and Knee Surgeons (AAHKS) concluded that in their ex-
perience, the diagnosis of PJI could be made in 90% of patients by getting an ESR and CRP followed by selective aspiration of the joint if these values are elevated or if the clinical suspicion is high (3). The use of a combination of routinely performed tests has been further supported by an evidence-based diagnostic algorithm formulated following a systematic literature review which based the decision to perform joint aspiration on ESR and CRP levels and on radiological findings (11). However, this proposed algorithm failed to examine the value of investigations in combination. Such methodology has previously been shown to improve the diagnostic accuracy for diseases such as rheumatoid arthritis (12).

The primary aim of this study was to evaluate the reliability and sensitivity of a small number of routine pre-operative tests in a large cohort of patients for the diagnosis of PJI. The secondary aim was to create a diagnostic algorithm using a combination of routine pre-operative diagnostic tests and clinical suspicion indices.

MATERIALS AND METHODS

A retrospective cohort study was conducted examining PJI patient factors (age, gender and past medical history) and as well as pre-operative predictors of PJI including routinely performed investigations (pre-operative CRP, ESR, serum WBC count and joint aspiration). These factors were chosen as they are routinely performed in our centre for all patients suspected of PJI. Additional diagnostic tests such as synovial leucocyte count and neutrophil percentage, included in the Musculoskeletal Infection Society (MSIS) criteria (13), were not routinely performed and therefore were not included in the analysis.

The diagnostic accuracy of pre-operative predictors used individually and in combination was then analysed with a secondary aim of creating a new diagnostic algorithm combining these routine investigation results with the level of clinical suspicion of PJI. The gold standard for diagnosis of PJI was one or more positive intra-operative culture samples against which all predictors were compared. Despite the MSIS criteria (13) defining PJI as two or more culture results, a single positive culture result has previously been used as criterion for PJI (14) and patients with single positive cultures at second stage revision have been shown to have higher rates of repeat implant failure (15). The MSIS has also been reported to be incorrect in up to 10% of cases of presumed aseptic joints (16). This study was performed in collaboration with the orthopaedic, microbiology and laboratory teams at a single tertiary care hospital.

Inclusion criteria were all patients who had routine blood tests, pre-operative intra-articular aspiration (of prosthetic hip or knee joint) and subsequent intra-operative culture samples from the same prosthetic joint. Figure 1 shows how the final cohort was selected from the initial dataset. The final cohort consisted of 128 patient-aspiration-operation pairs from 122 patients, including six patients with two prosthetic joints.

The patient cohort was identified by interrogating the Laboratory Information Management System (LIMS), from January 2010 to December 2016.
Using the Cognos search tool (17) all samples logged from patients who had been admitted to an orthopaedic ward at the time of receipt of any culture samples identified 4451 individual sample codes. These codes were then cross referenced using a separate microbiology laboratory system, to identify the unique patient numbers linked to each sample. The data was then sorted according to patient numbers to identify individuals who had both an aspiration sample and intra-operative samples with recorded culture results.

One hundred and twenty-two patients were then entered into the Trusts’ electronic patient system and individual records were reviewed, to first corroborate that the individual had had culture samples taken pre-operatively and intra-operatively and samples were consistent with samples logged by the microbiology laboratory system. Once this had been confirmed the eRecord results section was reviewed and combined with the microbiology sample information from LIMS each individual patient had data collected for: date of birth, sex, past medical history (PMH) with particular reference to co-morbidities increasing the risk of PJI (diabetes, peripheral vascular disease, rheumatological conditions, vascular disease, renal failure, hepatic failure, chronic pulmonary disease (18)), previous surgery to the involved joint, pre-operative CRP, ESR and serum white cell count (WCC), date of aspiration and organisms isolated (if culture positive), date and type of surgery performed at time of sampling, number of intra-operative samples taken and organisms isolated (if culture positive).

Statistical analysis was performed using Statistical Package for Social Sciences version 17.0 (SPSS Inc., Chicago, IL, USA). A Student’s t-test, unpaired and paired was used to compare linear variables between groups. Dichotomous variables were assessed using a Chi square test. Receiver operating characteristic (ROC) curve analysis was used to identify predictors and thresholds (cut points) for the CRP and ESR values according to the overall diagnosis. The area under the curve (AUC) ranges from 0.5, indicating a test with no accuracy, to 1.0 where the test is perfectly accurate by identifying all PJI. The threshold is equivalent to the point at which the sensitivity and specificity are maximal in predicting PJI. Multivariate logistic regression analyses were used to identify pre-operative independent predictors of PJI. A p-value of <0.05 was defined as statistically significant.

The diagnostic accuracy of the independent pre-operative predictors of PJI were analysed and compared with intra-operative culture sample to calculate sensitivities, specificities, positive predictive values (PPV) and negative predictive values (NPV). The diagnostic predictors were analysed individually and in combination. PPV and NPV are useful diagnostic markers as they predict the diagnosis (looking forward) rather than retrospectively analysing the accuracy of a test (looking backwards) as with sensitivity and specificity (19).

The authors conducted a retrospective service evaluation, as such there was no additional patient contact and no requirement for formal ethical approval. The project was registered with the institutional audit department (registration number 7851) and was conducted in accordance with the Declaration of Helsinki and the guidelines for good clinical practice.

**RESULTS**

The 128 patient-aspiration pairs had a mean age of 67 years (range 41-89), 66 were female and 62 male. Fifty-nine patient-aspiration pairs had medical co-morbidities that could increase their risk of infection, i.e. an ‘at risk’ PMH (diabetes, peripheral vascular disease, rheumatological conditions, vascular disease, renal failure, hepatic failure, chronic pulmonary disease (18)). Sixty patient-aspirations pairs had samples taken from a knee and 68 were taken from a hip.

Table 1 shows the analysis of the dichotomous variables; gender, PMH, joint type and joint aspirate results, to determine those which were predictors of PJI. The presence of an ‘at risk’ PMH and a positive joint aspirate were significant predictors of PJI (p=0.04 and <0.001 respectively) but gender and joint type were not predictive of PJI (p=0.31 and 0.93 respectively).

For the scalar variables; pre-operative CRP, ESR, WBC and patient age, table 2 shows the ROC
Once the independent predictors of PJI had been identified the sensitivity, specificity, PPV and NPV were calculated for each to determine their diagnostic accuracy, compared with intra-operative positive cultures. Despite the significance of ‘at risk PMH’ as an independent predictor of PJI it had low sensitivity, specificity, PPV and NPV (52.0%, 62.3%, 66.1% and 47.8% respectively) and was therefore excluded from the model. The remaining predictors were analysed individually, in pair combinations and as a combination of all three.

Table 3 shows a full summary of the diagnostic accuracy results for each predictor individually and for all possible combinations and figure 3 shows the comparative values between all combinations. CRP used individually had the highest sensitivity and NPV of any individual or combination of tests (75.0% and 75.9%, respectively). ESR individually had a comparable NPV of 75.0% but a lower sensitivity of 66.7%. Aspirate individually had the highest specificity and PPV of 86.1% and 74.4% respectively.

Combining diagnostic tests has previously been shown to increase the probability of true positivity in the diagnosis (12). When considered in pair combinations, ESR + CRP increased their specificity to 82.8% but the sensitivity, PPV and NPV decreased compared to the individual values of ESR and CRP used alone, CRP + aspirate increased the specificity to 92.9% but the sensitivity, PPV and NPV decreased compared to the individual values of CRP and aspirate separately. ESR + aspirate had

---

**Table 1. — Pre-operative non-linear predictors of PJI**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Positive</th>
<th>Negative</th>
<th>OR (95% CI)</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Male</td>
<td>30</td>
<td>32</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>26</td>
<td>40</td>
<td>(0.34 to 1.40)</td>
</tr>
<tr>
<td>PMH</td>
<td>No</td>
<td>36</td>
<td>33</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>39</td>
<td>20</td>
<td>(0.23 to 0.96)</td>
</tr>
<tr>
<td>Joint</td>
<td>Knee</td>
<td>26</td>
<td>34</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>Hip</td>
<td>30</td>
<td>38</td>
<td>(0.51 to 2.08)</td>
</tr>
<tr>
<td>Aspirate</td>
<td>Negative</td>
<td>27</td>
<td>62</td>
<td>6.66</td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>29</td>
<td>10</td>
<td>(2.85 to 15.6)</td>
</tr>
</tbody>
</table>

*Chi square. Significant p values in bold.

## Table 2. — ROC curve analysis of linear predictors of PJI

<table>
<thead>
<tr>
<th>Variable</th>
<th>AUC</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.516</td>
<td>0.407 to 0.625</td>
<td>0.77</td>
</tr>
<tr>
<td>WCC</td>
<td>0.603</td>
<td>0.497 to 0.709</td>
<td>0.07</td>
</tr>
<tr>
<td>CRP</td>
<td>0.717</td>
<td>0.616 to 0.818</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>ESR</td>
<td>0.741</td>
<td>0.648 to 0.834</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Significant p values in bold.

---

**Fig. 2 — ROC curve predicting PJI using age, C-reactive protein (CRP), white cell count (WCC) and erythrocyte sedimentation rate (ESR)**

Curve analysis of linear variables, CRP and ESR were demonstrated to be predictors of PJI with an AUC of more than 70% (p<0.001 for both CRP and ESR), whereas age and WBC were not predictive of PJI (p=0.77 and 0.07 respectively). Figure 2 shows the cut off points (threshold values) for the linear variables. The values that had the highest sensitivity and specificity of predicting PJI were >5 for CRP and >16 for ESR.

Once the independent predictors of PJI had been identified the sensitivity, specificity, PPV and NPV were calculated for each to determine their diagnostic accuracy, compared with intra-operative positive cultures. Despite the significance of ‘at risk PMH’ as an independent predictor of PJI it had low sensitivity, specificity, PPV and NPV (52.0%, 62.3%, 66.1% and 47.8% respectively) and was therefore excluded from the model. The remaining predictors were analysed individually, in pair combinations and as a combination of all three.

Table 3 shows a full summary of the diagnostic accuracy results for each predictor individually and for all possible combinations and figure 3 shows the comparative values between all combinations. CRP used individually had the highest sensitivity and NPV of any individual or combination of tests (75.0% and 75.9%, respectively). ESR individually had a comparable NPV of 75.0% but a lower sensitivity of 66.7%. Aspirate individually had the highest specificity and PPV of 86.1% and 74.4% respectively.

Combining diagnostic tests has previously been shown to increase the probability of true positivity in the diagnosis (12). When considered in pair combinations, ESR + CRP increased their specificity to 82.8% but the sensitivity, PPV and NPV decreased compared to the individual values of ESR and CRP used alone, CRP + aspirate increased the specificity to 92.9% but the sensitivity, PPV and NPV decreased compared to the individual values of CRP and aspirate separately. ESR + aspirate had
Table 3. — Diagnostic accuracy of individual tests and in combination

<table>
<thead>
<tr>
<th>Test Combination</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRP &gt; 5</td>
<td>75.0</td>
<td>62.9</td>
<td>61.8</td>
<td>75.9</td>
</tr>
<tr>
<td>ESR &gt; 16</td>
<td>66.7</td>
<td>70.3</td>
<td>61.2</td>
<td>75.0</td>
</tr>
<tr>
<td>Positive aspirate</td>
<td>51.8</td>
<td>86.1</td>
<td>74.4</td>
<td>69.7</td>
</tr>
<tr>
<td>CRP + ESR</td>
<td>20.0</td>
<td>82.8</td>
<td>45.0</td>
<td>59.6</td>
</tr>
<tr>
<td>CRP + Aspirate</td>
<td>14.3</td>
<td>92.9</td>
<td>61.5</td>
<td>57.5</td>
</tr>
<tr>
<td>ESR + Aspirate</td>
<td>4.4</td>
<td>100.0</td>
<td>100.0</td>
<td>59.8</td>
</tr>
<tr>
<td>CRP + ESR + Aspirate</td>
<td>41.5</td>
<td>95.3</td>
<td>85.0</td>
<td>71.8</td>
</tr>
</tbody>
</table>

Highest values for each diagnostic accuracy test in bold.

Fig. 3 — Diagnostic accuracy as per test combination.

Fig. 4 — Diagnostic algorithm for periprosthetic joint infection.
100% specificity and PPV, the highest value of any pair or triple combination or individual tests.

Using all three tests together the specificity and PPV were higher than the test individual values (95.3% and 85.0 respectively) but lower than ESR + aspirate used without CRP. The NPV (71.8%) was higher than the tests in any pair combinations but lower than CRP or ESR individually. The sensitivity (41.5%) was higher than the tests in any pair combinations but lower than the values for each test used individually.

Using the diagnostic accuracy results for each independent predictor, considered individually, in pair combinations and as a combination of all three; a diagnostic algorithm was created using the PPVs and NPVs integrated with the level of clinical suspicion of PJI.

Figure 4 shows the diagnostic algorithm. The algorithm utilises the initial level of clinical suspicion of PJI which would involve a combination of the patients presenting history and clinical examination findings as clinical history remains the cornerstone of initial assessment (3,4). For patients with a low clinical index of suspicion for PJI the algorithm focuses on the investigation combinations with the highest PPV values as the aim is to prove the presence of PJI. For patients with a high index of clinical suspicion the algorithm focuses on the investigation combinations with the highest NPVs as the aim is to disprove the presence of infection.

DISCUSSION

The principal finding of this study is that routinely performed pre-operative inflammatory markers and joint aspiration in combination offer enhanced diagnostic reliability and sensitivity for PJI compared to use in isolation. A diagnostic algorithm is presented to support the use of these simple tests individually and in combination when considered with the level of clinical suspicion of PJI.

There are key limitations to this work. It is retrospective and as such may not have captured all of the patient information but did allow for both ante and retro analysis of data contained within the datasets. In some instances the authors were unable to obtain detailed information regarding antibiotic administration and for this reason it was left out of the final analysis so as to improve the completeness of the final dataset. The use of intra-operative culture results as the chosen gold standard does have limitations, due to reported cases of culture negative PJI (20). An alternative gold standard for PJI diagnosis are the internationally accepted diagnostic criteria from the MSIS Workgroup (21), however this was not used as the gold standard as the majority of patients did not have all investigations completed to use these criteria. Honkanen et al. (22) suggested that true PJI cases could be missed by the MSIS criteria as a result of exclusion of clinical decision making. The algorithm recommends in cases with low level of clinical suspicion and a negative aspiration that further investigations be conducted or alternative diagnoses be considered. The algorithm does not offer further guidance on what investigations should be performed and this topic was deemed to vast to be covered satisfactorily in this paper.

The current study found that CRP had the highest rate of sensitivity, consistent with Pohlig et al.’s (23). For ESR the current study reports a sensitivity and specificity lower than that reported by Pohlig et al. (23), however they used an increased cut off point of 30mm/hr, as recommended by the MSIS International Consensus (13) which could explain their higher rate of specificity. Some authors argue that the MSIS cut-off values are arbitrary with limited evidence (10,24) and indeed thresholds ranging from 12 to 40mm/hr have been proposed (24). The current study calculated 16mm/hr for ESR as the value with the highest sensitivity and specificity for predicting PJI. The reported diagnostic accuracy of aspiration of the joint in the current study is consistent with previously reported ranges of sensitivity and specificity (9).

By combining the assessed routine pre-operative tests an investigative pathway was offered without the need for multiple costly tests with limited accessibility as required by other diagnostic models (13). The authors used predictive values of all individual and combinations of investigations to create a diagnostic algorithm. With the highest NPVs, CRP and ESR were recommended for use in
isolation for patients with a high index of clinical suspicion of PJI without the need for joint aspiration i.e. those patients in which investigation were being used to disprove the diagnosis. For those patients in whom the index of clinical suspicion was low (i.e. investigated to prove infection) the algorithm recommends obtaining an aspirate and using in combination with ESR due to the high PPV. Joint aspiration is an invasive procedure with reported complications including pain, nerve palsy and infection (25,26). Therefore, selective aspiration as described in the algorithm, is recommended (9,27-29).

Whilst newer diagnostic tools such as biomarkers and isotope labelled scans have been developed, the results are variable (30,31). Alpha defensin, a biomarker for PJI, initially had impressive results (32) however, the novel peri-operative alpha defensin test, Synovasure has been shown to have inferior diagnostic accuracy for PJI (33) and still requires a joint aspiration to be performed. The use of isotope-labelled radionucleotide scans have been shown to have limitations differentiating infection from aseptic inflammation (34). Several other markers such as procalcitonin, tumour necrosis factor-alpha, lipopolysaccharide-binding protein, and CD64 have also been evaluated as a marker of PJI but showed inconsistent results (35). These newer methods have the added disadvantages of increased cost, limited accessibility and the inability to identify the causative organism.

The algorithm proposed from this work utilises routine pre-operative investigations, which are readily available in orthopaedic centres and also incorporates clinical decision making, a potential advantage over the previously discussed MSIS International Consensus Meeting diagnostic criteria for PJI (13). This is supported by the recommendations of the AAHKS (2017) (3). By using retrospective clinical data, the findings and recommendations of the current study are directly comparable to current practice and relevant to treating clinicians. The authors believe that the current paper reports an investigative pathway that is realistically applicable to current routine orthopaedic practice. As such, this has the potential to improve pre-operative planning and patient counselling thereby optimising surgical outcome for infected revision hip or knee arthroplasty.

**CONCLUSION**

The current study evaluated the accuracy of aspiration culture, serum CRP and ESR used in isolation and combinations, in the diagnosis of PJI. It has shown serum CRP and ESR levels used in isolation have superior NPV values compared to aspiration alone. ESR in combination with aspiration had the highest PPV value. From these results a new diagnostic algorithm has been created incorporating the NPV and PPV values with clinical indices of suspicion.

**REFERENCES**


