December 13, 2023

Letter to the Editor for “Nationwide Results of Microorganism Antigen Testing as a Component of Preoperative Synovial Fluid Analysis”

J Parvizi et al.

Dear Dr. Swiontkowski,

This letter is concerning a recent publication in the journal by O’Shaughnessey et al. (Nationwide Results of Microorganism Antigen Testing as a Component of Preoperative Synovial Fluid Analysis J Bone Joint Surg Am. 2023;105:448-54 d http://dx.doi.org/10.2106/JBJS.22.00807).

The authors should be congratulated for their effort in evaluating the role of a commercial test that has been in use in the United States for many years with no prior publications related to its validation.

In our opinion, the paper suffers from major scientific flaws:

The authors have taken the liberty to alter the International Consensus Meeting (ICM) 2018 criteria by assuming that synovial C-Reactive Protein (CRP) and serum CRP are the same. This assumption has problems on three fronts:

1) Despite the reassurance by the authors that the synovial CRP and serum CRP may be the same (citations noted), the scientific community continues to explore this matter and is not yet convinced. There are numerous studies, not cited in the publication, showing that synovial CRP is not the same as serum CRP. The ICM delegates, in fact, eliminated synovial CRP from the definition that had been included in a prior 2018 paper as a criterion. Incidentally, in the 2018 paper, synovial CRP was given a score of 1 versus a score of 2 for serum CRP.

2) Even if synovial CRP was the same, or superior, to serum CRP the threshold for the test is not yet known. Did the authors have a specific threshold for synovial CRP that they used? Was this validated?

3) The ICM definitions were developed after extensive scientific research, deliberations, discussions, and even voting. We cannot as individuals take the liberty of changing a definition just because the dataset that we have is inadequate to categorize patients into infected versus non-infected. Taking such liberty may alter the dataset substantially, which is the case here, and lead to conclusions that could potentially harm patients.

The fact that authors did not have the ability to categorize patients into infected versus non-infected using the proper ICM definition, makes the entire analyses of the paper and its conclusions invalid.

Thank you for looking into this matter.

Shohat N, Buttaro M, Budhiparama N, Cashman J, Della Valle CJ, Drago L, Gehrke T, Hailer NP, Higuera CA,

References


Conflict of Interest: None Declared

Article Author Response

13 December 2023

Article Author(s) to Letter Writer(s)

11/04/2023

We appreciate the opportunity to address the concerns expressed in the letter to the editor regarding our study entitled “Nationwide Results of Microorganism Antigen Testing as a Component of Preoperative Synovial Fluid Analysis” (J Bone Joint Surg Am. 2023;105:448-54).

Our study retrospectively analyzed laboratory data from a very large cohort of synovial fluid (SF) samples. Every sample had a complete set of SF laboratory results, including the SF-CRP, SF-alpha-defensin, SF-WBC, SF-PMN, and SF-culture, which were collectively utilized to score and classify patients based on the definition of the 2018 International Consensus Meeting (ICM)\(^\text{1}\). Our study transparently described replacement of the serum CRP by the SF-CRP, citing supporting literature, as the only modification to scoring in our study. This modification and its rationale were discussed during the review process and the Journal accepted our study understanding that this modification was made.

We acknowledge the letter writers’ reservations; however, we contend that their critique overlooks established research and diverges from the precedents in the literature, wherein modifications to definitions of PJI have been made by several authors of the 2018 ICM themselves. Furthermore, we provide a sensitivity analysis for our study reassuring the Journal and the readership that the decision to replace serum with SF-CRP had no meaningful impact on our results or conclusions.

Precedent for Modifying Definitions of PJI

The existing literature, notably including works by several authors the 2018 ICM, demonstrates various modifications to recognized definitions of PJI, setting a precedent that is important for scientific creativity and progress. For example, following the establishment of the 2011 MSIS infection definition, Drs. Parvizi and Higuera published one of the first studies ever to use a modified version\(^\text{2}\) in 2012. Their modification of the 2011 MSIS criteria was necessitated by the fact that their data did not have frozen section histology, prompting them to change the scoring system of the definition. Further, in 2023 within JBJS, Dr. Parvizi and his team modified\(^\text{3}\) the 2018 ICM criteria,
notably excluding all serum laboratory tests to suit their study design. These are just a few examples that can be found in the literature clearly demonstrating the modification of a definition of PJI in a manner that could alter the score and eventual classification of patients or samples. We support these examples of studies which utilize transparently modified definitions of PJI, as they provide a reasonable rationale for why the modification was necessary. The precedent of utilizing modifications to definitions of PJI is an important part of the existing and future literature on PJI. We do not believe it is scientifically sound to claim that our study is invalid, solely based on our transparent use of modified criteria, as long as the modification is based on a good rationale.

While we champion the spirit of scientific adaptation and transparent modifications, we are against misrepresentations or ambiguous descriptions that could mislead the academic community. There are studies in esteemed orthopedic journals, JBJS included, where several prominent authors purport to use the 2018 ICM definition without qualification. Yet, they actually use a scoring system which is notably different than the 2018 ICM definition. This misrepresentation of the 2018 ICM criteria is particularly apparent in one study that includes the 2018 ICM definition in its title, yet erroneously misrepresents a figure of the scoring method as the 2018 ICM criteria, when in fact it illustrates an alternative scoring system. The discrepancies in these studies are not made clear to readers and are not addressed in the manuscript, leading to potential corruption of the scientific record. We believe that our transparent use of a modified ICM definition is fundamentally different in nature than these examples of misrepresentation.

The Use of SF-CRP Instead of Serum CRP

The review process for our study included questioning of the use of SF-CRP instead of serum CRP, supported in our study by two citations. However, given the concern expressed in the letter that the serum CRP and SF-CRP could have meaningful differences, we reinforce our rationale for modification demonstrating robustly that the serum and SF-CRP are highly correlated and that if there is a difference, the SF-CRP is likely more accurate.

A large number of studies in the literature confirm a significant direct correlation between serum and SF-CRP. Specifically, we identified five studies that assessed this relationship: three highlighted a strong correlation with coefficient values between 0.7 and 0.8, while the other two demonstrated a moderate-to-strong correlation ranging from 0.6 to 0.7. For instance, Tetreault and Della Valle et al. argued that “synovial fluid CRP level in an inflamed or infected joint often mimics the serum CRP level.” Additionally, Baker and Parvizi et al. asserted “There was a good correlation between serum and synovial fluid CRP levels.” We agree with these direct quotes from the literature, demonstrating the consistent and well-established principle that serum and SF-CRP are highly correlated.

We additionally identified six studies that compared the diagnostic performance of serum and SF-CRP in diagnosing PJI. Across the board, these studies reported comparable levels of sensitivity, specificity, and area under the curve (AUC) for both tests. Intriguingly, in five out of these six studies, the AUC for synovial fluid CRP surpassed that of serum CRP, primarily attributed to enhanced specificity, which translates to a reduced rate of false positives. The sixth study discerned no disparity between the two test types. Baker and Parvizi et al., who conducted the largest of these studies, observed, “While [synovial CRP] diagnostic performance was higher than its serum counterpart, there was still good correlation between serum and synovial CRP levels in revision arthroplasty patients.” It is reasonable to conclude from the existing literature that the SF-CRP is either equal to or superior to the diagnostic performance of the serum CRP.

It can be observed that there are now a large number of studies evaluating the comparison of serum and SF-CRP, all in support of the general idea that these tests are highly correlated and that any differences between these tests generally
favor SF-CRP. We believe that the letter writers’ concern, that differences between serum and synovial fluid CRP invalidate our study, is not consistent with a very full, multi-institutional, and multi-national literature on these biomarkers.

Qualitative Reasons to Prefer SF-CRP instead of Serum CRP

While we did not have the space to explain in our manuscript, there are many reasons why SF-CRP could be considered a qualitatively better biomarker than serum CRP. Serum CRP assays are usually conducted in local settings where test variations can occur based on different commercial kits used for CRP testing. Because the serum test is not designed to diagnose PJI, laboratories can report differing normal ranges using differing units, complicating result interpretation and potentially requiring unit conversion to align with thresholds used for PJI definitions\(^{17}\). This process may not be uniformly or properly executed across studies, and examples can be found where authors\(^{18}\) erroneously choose to use the generic laboratory interpretation instead of the PJI-optimized 10 mg/L. Additionally, serum CRP is often measured at a different time than the synovial fluid aspirate, introducing potential temporal discrepancies with other test results. Another critical point is that serum CRP data is typically reported separately from synovial fluid results, demanding high data integrity as researchers manually input these details into study datasets, which is likely accompanied by a background rate of data transfer error.

In contrast, the SF-CRP, as used in our study, surpasses serum CRP in every qualitative aspect mentioned above. All 67,441 patients in our study underwent the same validated laboratory test for SF-CRP in a CLIA-certified laboratory. The units and normal/abnormal thresholds were consistently reported in alignment with the literature and the laboratory’s internal validation. This uniformity eliminates errors in interpretation arising from variable normal ranges and cut-offs. Additionally, SF-CRP was measured contemporaneously with other synovial fluid tests, capturing a unified snapshot of the patient’s condition. Finally, all data were directly transferred from the laboratory instruments to the dataset, ensuring maximal data integrity without the risk of manual entry errors. These factors constitute real advantages of the SF-CRP in the setting of research, lending to a higher level of data integrity.

Scenario-Based Sensitivity Analysis Validates Our Study’s Results and Conclusions

While the prevailing literature and qualitative analysis robustly supports our methodologic decision to utilize SF-CRP in place of serum CRP, we are also able to provide a sensitivity analysis which mathematically proves that our results and conclusions were not measurably altered by our decision to utilize SF-CRP.

For a meaningful change to occur in our study’s outcomes, a combination of hypothetical conditions must simultaneously converge: 1) A meaningful segment of our samples must have divergent serum CRP and SF-CRP values. 2) When discrepancies arise, the SF-CRP result must be the inaccurate one. 3) The discrepancy must occur in samples which are dependent on the CRP result for classification.

The 2018 ICM scoring system assigns specific points to various criteria, categorizing them as: Not Infected (0-2 points), Inconclusive (3-5 points), and Infected (≥6 points). An examination of this scoring methodology uncovers several critical insights. A substantial portion of the samples from our study were assigned a classification that was mathematically independent of the CRP result. In other words, even if we were to replace our SF-CRP results with serum CRP results that changed the CRP criteria points from 0 to 2 or from 2 to 0, these sample classifications would not change category. Specifically, samples with a point tally of 0 or those exceeding 7 have classifications that the CRP result in our study cannot alter, given that a shift of more than 2 points would be required to modify their
categorization. Similarly, samples scoring 2 points (accompanied by a positive SF-CRP) or 7 points (with a negative SF-CRP) maintain classifications mathematically independent of the CRP result; any change induced by replacing our data with the serum CRP could only reinforce their existing category. These observations are direct mathematical consequences of the 2018 ICM scoring structure. A substantial 84.8% of our study’s samples held classifications that were mathematically independent of the CRP result (whether positive or negative), having no exposure to the letter writers’ concerns. Thus, opting for serum CRP over SF-CRP could not, mathematically speaking, alter their classification. Grasping this principle is pivotal to comprehending the results of the subsequent sensitivity analysis, which underscores that our study’s findings are unaltered by the CRP component of the ICM score.

In the methods of our study, we transparently reported the use of 6.6 mg/L as the cutoff for SF-CRP, in close alignment with both our internal validation and reports from the literature. To conduct our first sensitivity analysis for this response, we recalculated our results with two extreme thresholds for SF-CRP, which bound the range reported in the literature. We found that changing the SF-CRP threshold to either 2 mg/L or to 11 mg/L had no meaningful impact on any of our results. The reason for this is simple: 84.8% of our samples could not mathematically be affected by these changes; and of the remaining 15.2% of samples, there were insufficient numbers of samples between the CRP thresholds of 2 mg/L and 11 mg/L for any meaningful impact on our results. Therefore, the results and conclusions of our study are not at all sensitive to differences in the SF-CRP thresholds that are within the range described in the literature (Table 1).

For the second sensitivity analysis, we calculated the results of a worst-case scenario, which assumes an absolute inverse relationship between serum and SF-CRP, a premise that is not even realistic but provides insightful revelations. To conduct this analysis, we reversed the clinical interpretation of the CRP result in our study for every sample. The consequence is that every sample in our study which received two points for the SF-CRP would now be attributed 0 points, and every sample that was attributed 0 points for the SF-CRP would be attributed 2 points, replicating the assumption that the serum CRP contradicted the SF-CRP result for every sample in the data set. Conducting this worst-case sensitivity analysis resulted in no meaningful alteration to our results or conclusions (Table 1).

We are not surprised by the apparent redundancy of the CRP result in our data set. Every sample in our data set has a very complete set of synovial fluid data, which is generally more data than is needed to arrive at a definitive diagnosis. And the vast majority of samples in our study had such a high or low ICM score, that the CRP result simply made no difference to the diagnostic category. From a mathematical point of view, the results of the CRP test included in our study had no meaningful impact on the results or conclusions. This presumed redundancy was even suggested in the 2018 ICM document which specifically stated that “Not all tests are needed to use this proposed definition and a preoperative diagnosis can be made without the need for intraoperative findings.”

Conclusion

We have supported our rationale to make a transparent modification to the 2018 ICM criteria for the purposes of our study, not only relying on precedent set by some of the leaders of the 2018 ICM, but also supported by the existing literature demonstrating a high correlation between serum and SF-CRP and the likely diagnostic superiority of SF-CRP. We further demonstrated, using the strategy of a sensitivity analysis, that it is simply mathematically incorrect to suggest that our decision to use SF-CRP led to an alteration of our results or conclusion.

We believe that this sensitivity analysis and review of the literature should reassure the Journal and its readers that our
methodological choice was reasonable and constitutes valid methodology for this and future studies.

Krista O'Shaughnessy Toler, MS, MBA, PMP
Pearl Ravindra Paranjape, MS
Alex McLaren, MD
Brett Levine, MD, MS
Alvin Ong, MD
Carl Deirmengian, MD

References


<table>
<thead>
<tr>
<th>Diagnostic Results</th>
<th>Manuscript Results with Threshold CRP = 6.6 mg/L</th>
<th>Sensitivity Analysis with Threshold CRP = 2 mg/L</th>
<th>Sensitivity Analysis with Threshold CRP = 11 mg/L</th>
<th>Worst-Case Sensitivity Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>MID Panel Sensitivity</td>
<td>94.2%</td>
<td>93.9%</td>
<td>94.2%</td>
<td>94.0%</td>
</tr>
<tr>
<td>MID Panel Specificity</td>
<td>98.4%</td>
<td>98.5%</td>
<td>98.3%</td>
<td>98.4%</td>
</tr>
<tr>
<td>MID Panel False(+) Rate</td>
<td>1.6%</td>
<td>1.5%</td>
<td>1.7%</td>
<td>1.6%</td>
</tr>
<tr>
<td>MID Panel PPV</td>
<td>91.7%</td>
<td>92.9%</td>
<td>91.1%</td>
<td>91.3%</td>
</tr>
<tr>
<td>MID Panel NPV</td>
<td>93.8%</td>
<td>92.3%</td>
<td>94.5%</td>
<td>94.4%</td>
</tr>
<tr>
<td>Staphylococcus Panel Sensitivity</td>
<td>93.0%</td>
<td>92.8%</td>
<td>93.1%</td>
<td>92.9%</td>
</tr>
<tr>
<td>Candida Panel Sensitivity</td>
<td>92.3%</td>
<td>90.5%</td>
<td>92.6%</td>
<td>88.9%</td>
</tr>
<tr>
<td>Enterococci Panel Sensitivity</td>
<td>97.2%</td>
<td>97.2%</td>
<td>97.4%</td>
<td>97.4%</td>
</tr>
<tr>
<td>Staphylococcus Panel Specificity</td>
<td>98.9%</td>
<td>98.9%</td>
<td>98.9%</td>
<td>98.9%</td>
</tr>
<tr>
<td>Candida Panel Specificity</td>
<td>99.7%</td>
<td>99.7%</td>
<td>99.7%</td>
<td>99.7%</td>
</tr>
<tr>
<td>Enterococci Panel Specificity</td>
<td>99.6%</td>
<td>99.6%</td>
<td>99.6%</td>
<td>99.6%</td>
</tr>
</tbody>
</table>

**Table legend:** Results of various sensitivity analyses compared to manuscript results. CRP = C-reactive protein.; MID = microorganism antigen immunoassay detection; PPV = positive predictive value; NPV = negative predictive value.