Peer Review File

Reviewers

Comment 1: Determining the utility of various shoulder procedures in hopes of decreasing periprosthetic joint infection is a very important aspect of research and I commend the authors for their efforts. This is a very detailed review of the information extracted from the data they were given. However, I think their limitations that they acknowledge are ultimately too strong to warrant publication. At the basis of this study is PJI; however, it is never mentioned how different studies differentiate initial infection. It is known that C. Acnes represents commensal superficial bacteria and rates of infection differ between superficial and deep samples. In comparing the outcomes of various shoulder procedures, it is important to report on the definition of infection at some point in the paper. The sample size is too small to really make clinical recommendations based on the study. If they want to resubmit, I would suggest expanding the sample to before 2016 for a more robust data set.

Reply 1: We thank you sincerely for allotting your time to review this manuscript in detail. Since most of the studies included have different definitions and some even has not defined it in an objective manner, we focused to reflect those articles without specifically giving their definition. After considering your recommendation, we agreed to mention it in the discussion. In this systematic review, to guide the readers about the best treatment option, we would like to publish only the most recent article's data. We set the cut-off date by looking at the previous systematic reviews. As you mention, it would improve the data quantity, but that would be a reiteration of the previous articles published.

Changes in the text 1: We added text to Lines #287-294

"…and unexpected positive cultures are not equivalent to clinically relevant infections.(40,41) This was also an obscuring factor in the referenced articles about the definition of infection. Furthermore, regarding this issue, Hodakowski et al. noted that definition of infection is variable between different studies and argued that diagnosis and treatment of PSIs are largely variable among different hospitals.(42) This was also valid for the studies investigated in this review. It is important to note that the PSI definition would become more uniform after 2018 PJI meeting consensus on criteria for the diagnosis of PSI, even this clarification is still in its early phases.(43)"

Abstract
Comment 2: Why is periprosthetic joint infection challenging
Reply 2: Thank you for bringing this out. We added a statement to respond that question.
Changes in the text 2: We added text to Line #38
"… challenging because of patient morbidity, poor outcomes and need for reoperations."

Comment 3: Include some inclusion criteria in methods
Reply 3: We added some criteria to address your concern.
Changes in the text 3: We added text to Lines #44-45
"Studies which report on operative treatment and have longer than 2-year follow-up were included in this review."

Comment 4: Delete gender of 3 patients not specified
Reply 4: We deleted that statement.
Changes in the text 4: Deleted text from the manuscript.

Comment 5: 3rd sentence grammatically poor

Reply 5: Thank you for pointing out. We corrected the poor grammar.
Changes in the text 5: We edited the text in Lines #40-42
"This systematic review investigated the most recent articles about various treatment modalities used in the surgical treatment of PSI to find the most effective method in terms of infection clearance and function."

Comment 6: Your major conclusion of the paper that includes 339 patients cannot rest on a sample of 23 of those patients
Reply 6: We edited text to point out the difference between revision and non-revision procedures to meet with your suggestion. We also tried to meet with editorial assessment (5) about the main message of the conclusion.
Changes in the text 6: We edited the text in Lines #61-64
"The revision surgeries (one-stage and two-stage revision) were more effective than the non-revision surgeries in functional outcomes. In terms of infection clearance, revision procedures were more successful. Surgeons should prefer revision methods over non-revision procedures when feasible."

Introduction
Comment 7: Lines 74-76 is this true? Does it make a difference? Depends on the micro-organism
Reply 7: There is inconclusive data in the literature about C. acnes, but we believe the rest of the statement is valid. The debate about C. acnes was discussed in the Discussion.
Changes in the text 7: No change was made in the text.

Comment 8: Lines 84-88 reads as if you are comparing revision to all the other treatment modalities
Reply 8: This was a statement to reflect the previous concept in the treatment. For the last years, revision procedures gained popularity. So we wanted to highlight that point in introduction. We mentioned our investigation in the next paragraph as the main aim of the research.
Changes in the text 8: No change was made in the text.

Comment 9: You need to discuss defining infection, look at Paziuk 2022 or Hodakowski 2022 for further specifics
Reply 9: Thank you for your suggestion. We need to add, there is an ambiguity in the definition of infection, so we highlighted that in the discussion section. As this work was a systematic review, and inclusion criteria was mentioned in the methods, we opted to further discuss about definition of infection in the discussion section. We hope that will be appropriate for the reviewers.
Changes in the text 9: We added the text to Lines #287-294
"…and unexpected positive cultures are not equivalent to clinically relevant infections.(40,41) This was also an obscuring factor in the referenced articles about the definition of infection. Furthermore, regarding this issue, Hodakowski et al. noted that definition of infection is variable between different studies and argued that diagnosis and treatment of PSIs are largely variable among different hospitals.(42) This was also valid for the studies investigated in this review. It is important to note that the PSI definition would become more uniform after 2018 PJI meeting consensus on criteria for the diagnosis of PSI, even this clarification is still in its early phases.(43)"

Methods
Comment 10: Fine for most part
Reply 10: We appreciate it. We tried to keep up with the PRISMA guideline.

Changes in the text 10: No change was made in the text.

Comment 11: Could include definition of infection here to if you prefer
Reply 11: Thank you for your suggestion. We need to add, there is an ambiguity in the definition of infection, so we highlighted that in the discussion section. As this work was a systematic review, and inclusion criteria was mentioned in the methods, we opted to further discuss about definition of infection in the discussion section. We hope that will be appropriate for the reviewers.
Changes in the text 11: We added the text to Line #287-294
"…and unexpected positive cultures are not equivalent to clinically relevant infections.(40,41) This was also an obscuring factor in the referenced articles about the definition of infection. Furthermore, regarding this issue, Hodakowski et al. noted that definition of infection is variable between different studies and argued that diagnosis and treatment of PSIs are largely variable among different hospitals.(42) This was also valid for the studies investigated in this review. It is important to note that the PSI definition would become more uniform after 2018 PJI meeting consensus on criteria for the diagnosis of PSI, even this clarification is still in its early phases.(43)"

Results
Comment 12: Should not include why certain studies did not make it in 136-143
Reply 12: We omitted that part.
Changes in the text 12: Deleted the text from the manuscript.
Deleted the references of the articles 12-13. References were renumbered accordingly.

Comment 13: 145-146 is either two sentences or include "and" in the middle
Reply 13: Thank you for correcting.
Changes in the text 13: We edited the text in Line #154
".. The mean follow-up duration was 53.3±19.5 months."

Comment 14: Delete 148-153
Reply 14: We omitted that part.
Changes in the text 14: Deleted the text from the manuscript.

Comment 15: Line 170 need to mention the small DAIR sample size
Reply 15: A statement was added to meet with your suggestion.
Changes in the text 15: We added text to Line #172
"..group had a small sample size of 23 shoulders, but had the.."

Comment 16: Check standard deviation in line 182-183, I doubt it is 0 but it could be
Reply 16: This was "0" due to fact that mentioned study had not given separate results, but only the mean result.
Changes in the text 16: No change was made in the text.

Discussion
Comment 17: 215-219 reads as a limitation to the study but it is your first main paragraph of your major conclusion
Reply 17: To meet with the 6th comment from reviewers, we modified the conclusion section. We appreciate your suggestion here to clarify the research's conclusion.
Changes in the text 17: We edited the text in Lines #61-64
"The revision surgeries (one-stage and two-stage revision) were more effective than the non-revision surgeries in functional outcomes. In terms of infection clearance, revision procedures

were more successful. Surgeons should prefer revision methods over non-revision procedures when feasible."

Comment 18: 251-252, ok antibiotics was more prolonged, but tell me how long it was prolonged for and why that should matter. Have patients gotten pre-surgery antibiotics in any of these studies? That of course would seemingly make a difference in PJI
Reply 18: As you pointed out, this even can be a gamechanger in the treatment of PJI. Data from different articles were inconclusive about the previous antibiotic utilization to summarize in this manuscript. We believe that point is needed to discuss in a future work, so we decided only to reference the mentioned articles for further reading.
Changes in the text 18: No change was made in the text.

Comment 19: 278-288 you finally mention C. Acnes infection but do not mention how they defined it in your studies, if you agree with the fact it is confusing, you should mention if it is heterogenous in your study as well. You can look at Hodakowski 2022 for this exact discussion
Reply 19: Thank you for your suggestion. We decided to discuss the definition in detail.
Changes in the text 19: We added text to line #287-294
"…and unexpected positive cultures are not equivalent to clinically relevant infections.(40,41) This was also an obscuring factor in the referenced articles about the definition of infection. Furthermore, regarding this issue, Hodakowski et al. noted that definition of infection is variable between different studies and argued that diagnosis and treatment of PSIs are largely variable among different hospitals.(42) This was also valid for the studies investigated in this review. It is important to note that the PSI definition would become more uniform after 2018 PJI meeting consensus on criteria for the diagnosis of PSI, even this clarification is still in its early phases.(43)"

Tables
Comment 20: Table 2/3 you should make some type of marking for statistically significant findings. Both are good tables
Reply 20: We tried to keep the tables simple for easy comprehension. We added * for sig. findings in comparison.
Changes in the text 20: Table 2 was added * indicating statistically significant values between different treatments
Table 3 was removed to meet with the editorial suggestion.

AOJ Editorial Office
Comment 1: Since the search is for articles from 2016 onwards, the time frame (2016-2022) should be indicated in the title to better inform readers. Please rearrange the Title.
Reply 1: Thank you for your suggestion.
Changes in the text 1:Changes was made in the text Line #4
"Periprosthetic Shoulder Infection: Comparison of Surgical Treatment Options A systematic

review from 2016 to 2022"

Comment 2: Lines 39-40, "This systematic review investigated the most recent articles on the various treatment modalities used in the surgical treatment of PSI". The statement is not complete, and the objective of the paper should be clearly stated.
Reply 2: We added a sentence as this review's main objective.
Changes in the text 2: We added text to Line #41-42
"…to find the most effective method in terms of infection clearance and function."

Comment 3: Line 47, it's stated "A total of 339 patients (139 female, 197 male)". But the total number is 336 (139 plus 197) not 339. Please check and revise.

Reply 3: In Glanzman 2016, gender of the patients were not stated, so 3 patients were gender not-specified. To meet with the comment of reviewers (4) though, we deleted the gender of not specified patients.

Changes in the text 3: Deleted text from the manuscript.

Comment 4: Please specify the methods used to assess risk of bias in the included studies and summarize their risk of bias briefly.

Reply 4: Methods and phases were explained in Methods section.

Changes in the text 4: We added text to Lines #114-120

"ROBIS(12) tool was used to assess risk of bias. For study eligibility criteria, deifinitions were specific and review had a clear objective. For identification and selection of the studies, a meticulous search was performed by three different authors individually to minimize the bias and errors in the process. For data collection and study appraisal, data extraction was performed structurally to minimize the errors. For synthesis and findings, study heterogeneity was taken into account and missing data about the variables were reported. Overall, risk of bias in this review was rated as low."

Comment 5: Conclusion is not a repetition of the Results, but a thoughtful and conclusive summary or implication based on the Results. The concluding sentence should wrap up the entire work with a synthesis of key points. Please modify the conclusion in the abstract.

Reply 5: Thank you for that recommendation, we edited conclusion to address your concern accordingly.

Changes in the text 5: We edited the text in Line #61-64

"The revision surgeries (one-stage and two-stage revision) were more effective than the non-revision surgeries in functional outcomes. In terms of infection clearance, revision procedures were more successful. Surgeons should prefer revision methods over non-revision procedures when feasible."

Comment 6: Please report how many reviewers collected data from each report, whether multiple reviewers worked independently or not (for example, data collected by one reviewer and checked by another), and any processes used to resolve disagreements between data collectors.

Reply 6: All data was reviewed by the authors independently, as stated in the 2nd paragraph. Method of data collection was explained in more detail.

Changes in the text 6: We edited text in Line #104-107

"Following the PRISMA 2020 guidelines, three authors reviewed all the titles and abstracts (or full texts if needed) independently to reduce the risk of bias, removed duplicates, and included studies on PSIs.(11) Articles were investigated even if that individual paper was selected by one author only, then re-evaluated according to inclusion criteria by three authors."

Comment 7: Subheadings in the Methods and Results sections. This version is hard to read. The authors could refer to the section/topic of the PRISMA checklist.

Reply 7: Thank you for your suggestion. We edited the text accordingly.

Changes in the text 7: We added texts to Lines #96, 127, 143

"Studies Selection

Evaluation Criteria

Statistical Analysis"

Comment 8: Lines 107-108, the authors mentioned references to each selected article were checked to identify any missed articles. But I couldn't find this information in In Figure 1. In addition, the number of excluded papers should also be reported in the flow diagram. The authors are recommended to use the templated flow diagram in PRISMA 2020 as it has already encompassed such sources.
Reply 8: We used the template to meet with your concern.
Changes in the text: Changes were made in Figure 1

Comment 9: In the "Eligibility Search for Full Text Articles" process in Figure 1, it's stated the number is 81, but 204 minus 124 is equal to 80. And the final articles included in the systematic review were not 16 accordingly. All numbers have to be checked thoroughly and adapted accordingly.
Reply 9: Thank you for your correction. Incorrect numbers were edited.
Changes in the text: Changes were made in Figure 1

Comment 10: For ease of reading and double-checking, it would be better to add the reference number of the included studies in the tables.
Reply 10: Reference numbers were added to tables.
Changes in the text: Tables 1 and 2 were edited.

Comment 11: Please specify the methods used to assess risk of bias in the included studies, including details of the tool used, how many reviewers assessed each study and whether they worked independently.
Reply 11: In line with your previous suggestion method of data collection was explained in more detail.
Changes in the text 11: We edited text in Line #104-107
"Following the PRISMA 2020 guidelines, three authors reviewed all the titles and abstracts (or full texts if needed) independently to reduce the risk of bias, removed duplicates, and included studies on PSIs.(11) Articles were investigated even if that individual paper was selected by one author only, then re-evaluated according to inclusion criteria by three authors."

Comment 12: This is related to comment 11 above regarding the quality appraisal. Please present tables or figures indicating for each study the risk of bias in each domain/component/item assessed and overall study-level risk of bias.
Reply 12: We tried to address your concern(4) by explaining briefly in the text.
Changes in the text: We added text to Line #114-120
"ROBIS(12) tool was used to assess risk of bias. For study eligibility criteria, deifinitions were specific and review had a clear objective. For identification and selection of the studies, a meticulous search was performed by three different authors individually to minimize the bias and errors in the process. For data collection and study appraisal, data extraction was performed structurally to minimize the errors. For synthesis and findings, study heterogeneity was taken into account and missing data about the variables were reported. Overall, risk of bias in this review was rated as low."

Comment 13: Lines 164-166, "217 (63.4%) two-stage revisions, 59 (17.2%) one-stage revisions, 37 (10.8%) definitive spacer applications, 23 (6.7%) DAIR, and 6 (1.7%) resection arthroplasty procedures were applied", the authors could consider moving it after the statement "The most common persistent infection agent was C. acnes, with nine infected shoulders (39.1%) (Figure

5)" in line 169.

Reply 13: Statement was moved to suggested section.

Changes in the text: Text was moved to Lines #169-171

"217 (63.4%) two-stage revisions, 59 (17.2%) one-stage revisions, 37 (10.8%) definitive spacer applications, 23 (6.7%) DAIR, and 6 (1.7%) resection arthroplasty procedures were applied."

Comment 14: Please add the number of shoulders of each study.

Reply 14: Number of shoulders was added.

Changes in the text: Table 1 was edited.

Comment 15: The study design should also include "case series" not just retrospective.

Reply 15: They were all retrospective case series. This was added to the table mentioned to address your concern.

Changes in the text: Table 1 was edited.

Comment 16: The CMS was reported in Table 2 with the specific data, then, the authors could remove it from Table 1.

Reply 16: We removed the column mentioned.

Changes in the text:Table 1 was edited.

Comment 17: Does "Scientific level mean "the journal rank (quartiles)"? If not, please make corresponding explanations.

Reply 17: Scientific level was Levels of Evidence. We changed it accordingly to clarify and address your concern.

Changes in the text: Table 1 was edited.

Comment 18: The data of CMS and FE were already reported in Table 2. Could the authors consider combining the two tables?

Reply 18: We tried to highlight some studies having incomplete data by giving 2 separate tables. After considering your comment, we combined 2 tables and removed Table 3.

Changes in the text: Changes were made in Table 2. Removed Table 3

Comment 19: The data in Figures 2-5 should better report either in Table 1 or Table 2. For example, report the type and number of initial arthroplasty indication for each included study. Therefore, Figures 2-5 could either delete or they are regarded as supplementary materials, since the qualitative descriptions are enough.

Reply 19: We tried to keep tables basic for easy comprehension. This suggestion might further complicate already crowded tables with too much information. Figures might be regarded as supplementary materials if that is more appropriate for the editorial.

Changes in the text: No change was made in the text.

Comment 20: Readers couldn't get informative information from Figure 6. And there are some incorrect numbers for the DAIR and two-stage revision groups. It's suggested to remove it.

Reply 20: Thank you for your suggestion.

Changes in the text: Figure 6 was removed.

No change was made in the text.

Re-review

Comment 1: As for the Highlight Box, "shoulder arthroplasty" or "periprosthetic shoulder infection" should be also mentioned in the "Key findings" section. Also, it seems the authors repeated the findings again in the "what is new" section. It means what this manuscript adds

and what the added value of this study is. Therefore, the knowledge gap or the main purpose of this review could be summarized here.

Reply 1: Suggestions were applied.

Changes in the text:

Highlight box was modified in Lines #35-40

Comment 2: Regarding Table 1, how did the authors rate the levels of evidence of these included studies? Please specify the tool or system used to assess certainty in the body of evidence in the Methods.

Reply 2: Evidence tool was specified.

Changes in the text: Line #128-134

"After exclusions, a total of sixteen studies that met the inclusion criteria were identified for final evaluation. All articles were retrospective case series. Levels of evidence rated according to the Oxford 2011 Levels of Evidence.(12)

Five surgical treatment options included from the studies were; DAIR, resection arthroplasty, definitive spacer, one-stage revision, and two-stage revision. Multiple staged revisions were also identified as two-stage revisions."

Comment 3: "Risk of bias assessment" and "Data collection" should be also added as the subheadings in the Methods. The Results section still needs to be improved. Please add some subtitles for clarity.

Reply 3: Subheadings were added for clarification.

Changes in the text: Following subheadings were added:

Data collection, risk of bias assessment were added to Methods.

Initial surgery, causative microorganisms, interventions for infection clearance, functional assessment, additional surgeries were added to Results.

Comment 4: ROBIS is a tool for assessing the risk of bias in systematic reviews (rather than in primary studies). Therefore, the revise statement in line 127 is incorrect. For the authors' reference, here is the method used to assess risk of bias of case series studies: JBI Critical Appraisal Checklist for Case Series. Please refer to the weblink to download the checklist of case series: https://jbi.global/critical-appraisal-tools. The authors could also read the guidance about how to use this checklist on it.

Reply 4: JBI's tool was added.

Changes in the text:

Text was deleted from the manuscript.

Comment 5: Subsequently, please specify the methods used to assess risk of bias in the included studies. It's also highly recommended to use a table to summarize the risk of bias of each study in each question and overall study-level risk of bias.

Reply 5: Table 1 was added.

Changes in the text: Line #136-145

"Joanna Briggs Institute's tool of Critical Appraisal Checklist(13) was used to assess risk of bias. Clear criteria for inclusion and exclusion were questioned for studies. Infection descriptions and methods for the diagnosis of infection were assessed for validity. The studies were investigated for consecutive and complete inclusion of participants. Patient demographics, clinical information and outcomes were reviewed for clarification. Statistical analyses for each study were investigated for appropriateness. Risks of bias of each study were summarized in Table 1. Overall, in this systematic review, study-level risk of bias was rated low, since all

questions were addressed conveniently and there were only few unclear answers with mixed variables."

Former Tables 1 and 2 were renamed as Tables 2 and 3.