Peer-Review comments and authors responses

Dear Reviewers,

Thank you for your valuable comments and insightful suggestions. We have provided detailed, point-by-point responses to each of your comments below. In addition, we have revised the content of our paper based on your advice and clarified certain sections where necessary. We sincerely appreciate your time and consideration. The changes are as follows, and we have highlighted your original comments in bold, followed by our replies. Please kindly check.

Reviewer 1

To the Reviewer 1: Dear Reviewer 1, thank you for your valuable comments and insightful suggestions. We have provided detailed, point-by-point responses to each of your comments below. In addition, we have revised the content of our paper based on your advice and clarified certain sections where necessary. We sincerely appreciate your time and consideration. The changes are summarized as follows:

Comment: "PCC": abbreviation with no full name provided in abstract.
Response: *Thank you for addressing this point. The full name of the abbreviation "PCC", Population, Concept, and Context, has now been provided in the abstract.*

Comment: Outcomes should be described in a more formal way: "primary outcome is post-treatment complication and secondary outcomes are treatment success,
readmission, and hospital stay"
Response: *Thank you for addressing this point. The method section has been revised to formally describe the primary and secondary outcomes as: The primary outcome was 30-day post-treatment complication, and the secondary outcomes were treatment success (90 days and 1 year), readmission (at 30- and 90-day), and hospital stay (median duration).*

Comment: please add how they were measured (e.g., complication rate or hospital stay length).
Response: *Thank you for addressing this point. The method section has been revised to formally describe the primary and secondary outcomes as: The primary outcome was 30-day post-treatment complication, and the secondary outcomes were treatment success (90 days and 1 year), readmission (at 30- and 90-day), and hospital stay (median duration).*

Comment: Keywords: Age should be higher or equal 65.
Response: *Thank you for addressing this point. The keywords section has been updated so that the age criterion is now "≥65 years."*

Comment: The introduction is well written and informative, thus some minor improvements might be needed. Unclear sentence as it lacks context for me: increased since? Please revise: "Furthermore, the demand for institutional care and mortality rates increased by approximately 30% and 20%, respectively (Reinisch et al., 2022). "
Response: *Thank you for addressing this point. We have rephrased this sentence.*

Comment: Term "Conservative treatment:" is unclear.
Response: *Thank you for addressing this point. We removed this term from the introduction, and the terminology has been updated.*

Comment: Frailty is an important term in the manuscript, thus should be mentioned in the Introduction.

Response: *Thank you for addressing this point. We have introduced the importance of frailty in the introduction. "Frailty status is also an important risk factor for negative outcomes in the elderly, increasing the need for institutional care and mortality rates compared with non-frail patients."*

Comment: Materials and Methods - This section is described clearly. Please consider minor comments: You defined non-operative management (NOM) as antibiotic-only treatment. I recommend clarifying whether other NOM strategies are reported in the literature and, if so, briefly justifying why your review focuses exclusively on antibiotics. If antibiotic therapy is indeed the sole intervention classified as NOM in this context, please make that explicit for clarity.

Response: *Thank you for the suggestion. We clarified in the Study Design that, for the purposes of this review, non-operative management (NOM) refers exclusively to antibiotic therapy without surgery. Although other non-surgical strategies (e.g., observation, percutaneous drainage) are described in the literature, they are typically used for complicated appendicitis and are not standard approaches for uncomplicated disease; moreover, antibiotic therapy is the primary NOM strategy studied in this context. Accordingly, our review focuses on antibiotic-based NOM.*

Changes made/ Addition: Definitions. In this review, non-operative management (NOM) refers exclusively to antibiotic therapy without surgery. Other non-surgical strategies (e.g., observation, percutaneous drainage) were not included because they are not standard for uncomplicated appendicitis.

Comment: The term "elderly" is used in the manuscript, however, I recommend adopting a single, consistent term to describe the study population and ensure it reflects the most appropriate terminology (e.g., "geriatric patients"). In addition, please provide a brief justification for selecting 65 years as the cutoff, supported by references to prior literature that defines the geriatric population at this threshold.

Response: *Thank you for addressing this point. We have revised the manuscript to use the term **"elderly"** consistently throughout to ensure clarity and maintain uniform terminology. This approach avoids confusion for readers and aligns with conventions in medical writing, where consistency in terminology is preferred over stylistic variation.*

Comment: Results - This section needs improvements for both primary and secondary outcomes.
Primary analysis: The reader can clearly see the NOM vs OM contrast only in Chehab et al. (2021). For the others, the way you've written it makes it sound like complication rates are being reported in general, without explicitly tying them to NOM or OM. It needs to be corrected as the purpose states: "This scoping review explored the efficacy and safety of NOM compared with OM for acute […]". Thus, in the results you say: "For the primary outcome, reported complications varied in definition and measurement across studies. Chehab et al. (2021) reported higher complication rates in NOM patients (19.8%) compared to OM (10.8%), with NOM-associated risks including perforation (9.5%), sepsis (8.1%), and peritonitis (7.7%). Ashbrook et al. (2024) presented composite outcomes, with a total complication rate of 37.3%, ranging from 28.8% in non-frail patients to 57.1% in frail patients. Lunardi et al. (2024) reported an 18% complication rate. In contrast, Meier et al. (2023) described a broad range of complication categories but did not specify rates. Most complications were documented during hospitalization, except in Chehab et al.(2021), which included follow-up data up to 6 months."

Response: *Thank you for addressing these points. We assimilated your advice and also scrutinized the data again and rewrote the section with updated and more accurate information. We made the following changes:*
"For the primary outcome, three studies (Ashbrook et al., 2024; Chehab et al., 2021, Meier et al., 2023) directly

compared postoperative complications between NOM and OM, whereas Lunardi et al. (2024) focused on age-related differences and outcomes of successful versus failed NOM without a direct OM comparison. Among those comparing management strategies, Ashbrook et al. (2024) presented composite outcomes showing slightly higher complication rates in the NOM group (37.7%) compared with the OM group (36.4%), with an overall complication rate of 37.3%, ranging from 28.8% in non-frail patients to 57.1% in frail patients.
In the same study, complications were further analyzed by stratifying patients into frailty and non-frailty clusters. Within the frail cluster, 66.3% of patients were managed operatively and 33.7% non-operatively, while within the non-frail cluster, 69.8% were managed operatively and 30.2% non-operatively [Appendix B]. Chehab et al. (2021) reported Clostridium difficile infection rates in NOM patients (3%) compared to OM (1%). Lunardi et al. (2024) reported a 10.7% complication rate.
In contrast, Meier et al. (2023) described several complication categories, such as wound infection and abscess formation, but did not specify rates by management type. Most studies reported complications during hospitalization, except Chehab et al. (2021), which included follow-up data up to six months.
The changes above are based on the updated data presented below. We carefully scrutinized and verified the data throughout the manuscript and subsequently made the following changes. That said, the overall findings remain consistent with previous results.
*- **Ashbrook et al., 2024:** Revised total case number and complication percentages (now accurately reflecting 9,065 total cases, with corrected NOM vs. OM and frailty subgroup data). Based on the availability of data, we also computed the rates of operative versus non-operative management among frail patients—P(OM | Frailty) and P(NOM | Frailty)—to allow further analysis of complications within the frailty subgroup in this study, see Appendix B.*
*- **Chehab et al., 2021:** We removed the original complication data. The reasons are as follows: although this study compared complications between NOM and OM, the NOM group included patients who later required surgery due to failed conservative treatment. Therefore, it effectively compared OM with OM', which is not ideal for assessing NOM vs. OM complications and could be misleading. However, we retained Clostridioides difficile infection as a complication, since it was explicitly mentioned in the text and is a legitimate complication relevant to both OM and NOM. These changes do not substantially affect our results or interpretations. We appreciate this opportunity to fine-tune our paper to ensure it accurately reflects the overall evidence.*
*- **Lunardi et al., 2024:** Updated case count and complication rate (1,209 cases; 10.7%).*
*- **Meier et al., 2023:** Removed the previously listed number of complication cases, as the reported figure represented the overall population rather than the number of patients with complications. Post-treatment complication data were not reported.*

Comment: Secondary analysis: This paragraph summarizes important findings on treatment success and hospital length of stay, however, since your stated aim was to compare NOM and OM, the opening sentence should explicitly reflect this focus. Without the OM context, the reader may not see how these results address your research question. I also recommend providing results (numbers) for both NOM and OM wherever possible to facilitate comparison. In addition, the reporting of results should be more homogeneous: “OM (5 [IQR 3-9] vs. 4 [IQR 2-7] days” vs. "particularly among frail patients (median 12 days)”. Finally, please revise for structure to improve clarity for a reader.“Secondary outcomes showed that treatment success for NOM varied across studies. Chehab et al. (2021) reported an 82% success rate at six months, while Lunardi et al.
(2024) observed a 62% success rate during hospitalization. Hospital stay duration was generally longer for NOM compared to OM. Chehab et al. (2021) reported a significantly longer median length of stay for NOM compared to OM (5 [IQR 3-9] vs. 4 [IQR 2-7] days, respectively; $p<0.001$). Ashbrook et al. (2024) found that delayed OM resulted in the most extended hospital stays, par,particularly among frail patients (median 12 days), followed by NOM (7 days for frail, 4 days for non-frail) and immediate OM (5 days for frail, 3 days for non-frail). Lunardi et al. (2024) reported that failed NOM was associated with a

significantly longer hospital length of stay (LOS) compared to successful NOM (12 vs. 6 days; $p<0.001$), and Meieret al. (2023) documented an average increase of 3.22 days in LOS for NOM compared to OM."

Response: *Thank you for this helpful comment. We agree that the results section would benefit from clearer emphasis on the comparison between non-operative management and operative management across all included studies.*
Changes made:
- We revised the section to explicitly report NOM and OM outcomes for each study where data were available and restructured the paragraphs for improved clarity and consistency.
- Standardized the reporting format for median values and IQRs, ensured that all primary and secondary outcomes directly reflect the aim of comparing NOM vs OM in geriatric patients with acute appendicitis.

Comment: Discussion - This section provides many interesting insights. However, the flow could be improved to make the results easier to follow and to more directly address your research question, namely the comparison between NOM and OM.

Response: *Thank you for this suggestion. We reorganized the Discussion into four thematic blocks (primary outcomes, frailty findings, secondary outcomes, and methodological considerations). This restructuring improves coherence and ensures the section progresses logically from the main findings to their interpretation.*

Comment: I recommend starting with the primary analysis results, as these represent your key findings, and then proceeding to secondary outcomes.

Response: *Thank you for this suggestion. We revised the opening of the Discussion so that it begins with the primary outcomes. The first paragraph now clearly states that non-operative management (NOM) was consistently associated with higher complication rates and longer hospital stays compared with operative management (OM).*

Comment: For example, the statement "A key finding in our review was the significant variability in post-treatment complication rates" is too vague, as it does not specify how complication rates differed between NOM and OM.

Response: *Thank you for this suggestion. This vague statement was removed. We replaced it with a direct and specific comparison, reporting that NOM was associated with recurrent appendicitis, perforation, sepsis, and peritonitis, whereas OM was linked to wound infections, abscesses, and pulmonary complications. Relevant references were also consolidated here to strengthen the comparison.*

Comment: Please revise this section to highlight those differences more clearly and to ensure the narrative aligns with your stated objective.

Response: *Thank you for this suggestion. We revised sentences across the Discussion so that both NOM and OM are explicitly mentioned throughout. This ensures the comparison remains clear and directly aligned with the study's stated objective in every paragraph.*

Comment: To summarize, the first two paragraphs would benefit from greater clarity and a stronger narrative link to the research question. Please ensure that the framing and flow of these sections consistently highlight the comparison between NOM and OM, so that the reader can clearly see how the evidence addresses your study aim.

Response: *Thank you for this suggestion. The first two paragraphs were substantially revised. The new first paragraph introduces the scope of the review, the number of included studies, and the main comparative finding (higher complication rates and longer hospital stays with NOM). The second paragraph then directly compares complications between NOM and OM and introduces frailty as a key modifying factor, establishing a clearer and*

stronger narrative link to the research aim.

Comment: Conclusions - The conclusions section currently lacks reference to the results obtained in your primary and secondary analyses. To strengthen this section, I recommend including at least one concise sentence that directly addresses the research question and summarizes the key findings regarding the comparison between NOM and OM.

Response: *Thank you for addressing this point. We have made the amendment. Now, it is like "NOM with antibiotics is an increasingly adopted strategy for treating acute, uncomplicated appendicitis. However, its efficacy and safety in the elderly population remain insufficiently understood. This scoping review demonstrated that NOM may be associated with increased risk of complications, lower success rates, and longer hospital stays… "*

Comment: References - Morisa et al.: lack space

Response: *Thank you for addressing this point. We have adjusted it.*

Comment: Error in cita,on. Please check the author's name: R, K. H. (2024). Acute Appendici,s in the Elderly: Current Update on Management: Review Ar,cle. SAR Journal of Medicine, 5(3), 93–97. hFps://doi.org/10.36346/sarjm.2024.v05i03.005

Response: *Thank you for pointing that out! We have already changed it back to Kumar, H. R. (2024). Acute appendicitis in the elderly: Current update on management: Review article. SAR Journal of Medicine, 5(3), 93–97. https://doi.org/10.36346/sarjm.2024.v05i03.005. Please Kindly check.*

Comment: Citation does not look to be proper: (10. Scoping Reviews - JBI Manual for Evidence Synthesis - JBI Global Wiki, n.d.).

Response: *Thank you for pointing that out! We have already changed it back to Peters, M. D. J., Godfrey, C., McInerney, P., Munn, Z., Tricco, A. C., & Khalil, H. (2024). Scoping reviews. In E. Aromataris, C. Lockwood, K. Porritt, B. Pilla, & Z. Jordan (Eds.), JBI manual for evidence synthesis. JBI. https://doi.org/10.46658/JBIMES-24-0*

Comment: Please verify if referencing to tables is proper (PPCR journal guidelines and other papers published in PPCR journal): e.g.: "(age, sex, comorbidity [Table 1]), interventions [Table 3], and outcomes [Table 4 and Table 5].

Response: *Thank you for your advice. We have made amendments to ensure that the table stylization is more consistent. We also reviewed the current PPCR Author's Guide; however, there are no specific guidelines related to this matter. Therefore, we have done our best to clearly present the figures and ensure the references to the tables are appropriate.*

Comment: a third reviewer. [Table 2].

Response: *Thank you for your kind suggestion. We believe that a third reviewer is not required, as the two existing reviewers have provided consistent and aligned feedback.*

Comment: Abbreviations without full names and citation: Primary and secondary outcomes were defined a priori according to international guidelines (WSES Jerusalem Guidelines 2020; ACS NSQIP) and key clinical trials (CODA; APPAC).

Response: *Thank you for your comment. We have revised the text to include the full names and appropriate citations for all abbreviations. The sentence now reads as follows:*

Primary and secondary outcomes were defined a priori based on international guidelines and relevant clinical trials (World Society of Emergency Surgery Jerusalem Guidelines, 2020; American College of Surgeons National Surgical Quality Improvement Program, 2021; Comparison of Outcomes of Antibiotic Drugs and Appendectomy Trial, 2020; Antibiotic Therapy vs Appendectomy for Treatment of Uncomplicated Acute Appendicitis Trial, 2015).

Comment: Citation: (n.d.).

Response: *Thank you for addressing this point. We have already made changes to it.*

Comment: Review style: e.g.: “One possible explanation lies […]” / “One potential explanation is that conservative […]”

Response: *Thank you for addressing this point. We have already made changes to it.*

Comment: same phrase in 2 paragraphs close to each other.

Response: *Thank you for addressing this point. We have already made changes to it.*

Comment: Punctuation errors: Regarding primary outcomes, patients managed with NOM showed high rates of complications such as recurrent appendicitis, perforation, sepsis, and peritonitis, especially among frail individuals who fail the initial conservative approach.(Chehab et al., 2021; Meier et al., 2023) However, operative management is not without risks in this population, and has been associated with wound infections, abscesses, and pulmonary complications.(Cimino et al., 2024; Weinandt et al., 2020).

Response: *Thank you for addressing this point. We have already made changes to it.*

Once again, we sincerely thank you for your efforts in helping us improve. Your valuable comments and encouragement have enabled us to produce a stronger paper, and we look forward to hearing positive news from you.

Yours faithfully,
Group 9 PPCR team

Reviewer 2

To the Reviewer 2: Dear Reviewer 2, thank you for your valuable comments and insightful suggestions. We have provided detailed, point-by-point responses to each of your comments below. In addition, we have revised the content of our paper based on your advice and clarified certain sections where necessary. We sincerely appreciate your time and consideration. The changes are as follows, and we have highlighted your original comments in bold, followed by our replies. Please kindly check.

Comment: Please, check the last citation (and corresponding reference) from the third paragraph.

Response: *Thank you for addressing this point. It has been updated.*

Comment: In the last paragraph, considering that this is a scoping review, the expression “to address this issue” may not be entirely appropriate. Scoping reviews typically aim to explore broader research questions rather than to directly address a specific issue. Their purpose is to map existing evidence, identify knowledge gaps, and highlight areas for future investigation.

Response: *Thank you for addressing this point. We have adjusted the phrasing to be more adequate for a scoping review.*

Comment: Methods section - In the first paragraph of the Study Design section, third line, the author has included the full reference instead of the in-text citation format recommended by the journal. Please revise this by including the proper in-text citation and ensuring the corresponding reference is correctly listed in the References section.

Response: *Thank you for the suggestion. We corrected in-text citation formatting and capitalization, clarified the database rationale, described the screening/extraction workflow (Covidence for screening; a piloted Excel form for*

data-charting with access controls), and pre-specified mortality as a secondary outcome—recorded as 'not reported' when absent. Citations for PRISMA-ScR (Tricco et al., 2018) and the PCC framework (Peters et al., 2024) now follow the journal's text format. References from the methods section have been revised.

Comment: In the second line, please review the word "Also" with an inappropriate upper letter.
Response: *Thank you for addressing this point. It has been updated. Corrected capitalization of "also" and edited for flow.*

Comment: If word count is an issue, consider moving the search strategy to a supplementary document.
Response: *Thank you for your concern. Word count is not an issue for us at this stage, so we have retained the search strategy within the main manuscript.*

Comment: Why did the authors search only two databases? Why were additional databases, such as Scopus or Web of Science, not included? The limited choice of databases should be clearly justified.

Response: *Thank you for addressing this point. PubMed and Embase were selected for comprehensive and complementary coverage (MeSH/Emtree); preliminary scoping indicated minimal incremental yield from Scopus/Web of Science relative to available resources.We acknowledge that institutional access limitations contributed to the decision not to include other databases, which were evaluated during preliminary scoping.*

Comment: Selection, extraction, and synthesis sub-section: Why was a Microsoft Excel tool developed instead of using Covidence? It is unclear why Covidence, a platform specifically designed for systematic reviews, was used for screening but not for data extraction. If there is a valid reason for this choice, please clarify it in the Methods section. Additionally, the use of Microsoft Excel should be justified, as it is not considered a secure database for data management.

Response: *Thank you for addressing this point. We would like to provide further clarification. We used Covidence for the screening process, as it is specifically designed to facilitate systematic review screening and collaboration. However, for data extraction, we employed a standardized and piloted Microsoft Excel form because it allowed us to capture heterogeneous and study-specific variables that were not easily supported within our Covidence configuration.*
All Excel files were stored on an institutionally managed, access-controlled drive to ensure data security and integrity. Importantly, no patient-identifiable data were collected at any stage. This approach allowed for greater flexibility in data management while maintaining confidentiality and methodological rigor.

Comment: Outcomes sub-section: Please review the in-text citations, as they do not comply with the journal's formatting requirements. Additionally, it is difficult to read the Outcomes subsection without seeing mortality listed as a primary or, at minimum, a secondary outcome. Why was mortality not included? Was the data unavailable? Even if it was not reported in the studies included, mortality could still be listed as an outcome in the Methods section, and the Results section could then note that none of the selected articles provided this information. This is because the research question (and, consequently, the selected outcomes) must be defined prior to the search. Therefore, an outcome should not be excluded simply because data are not available in existing literature.

Response: *Thank you for addressing these points. We recognize that two issues were raised and have responded to each below.*
First, regarding in-text citations, we referred to the Author's Guide on the Harvard PPCR website, which states:
"APA guidelines for referencing are explained below:

- In the text, the references should appear as the first author's last name followed by the date of publication. Examples: 'as determined (Doe, 2010a, 2010b; Smith and Johnson, 2015). Smith et al. (2010) have shown...'"
Based on this guideline, our in-text citations comply with the required APA-style referencing format. If there are specific instances where our formatting deviates from the journal's expectations, we would greatly appreciate further clarification so that we can make the necessary adjustments.
Second, regarding mortality, it was not considered an isolated outcome of interest in our study or in the initial study designs, primarily due to its ambiguous nature. Elderly patients may die from causes unrelated to the disease course or its treatment, particularly among frail populations. We believe that focusing on mortality as a single outcome could introduce bias; therefore, it was not selected as a primary outcome. Instead, we aimed to explore outcomes measured within 3- , 6-months, and up to 1 year in our predefined analysis.
During data extraction, we observed that some papers mentioned mortality but did not specify whether it referred to all-cause or cause-specific mortality. Even Ashbrook et al. (2024) and Lunardi et al. (2024) referred to in-hospital mortality, which is relevant but still lacked definitional clarity. Deaths resulting from comorbidities such as dementia or unrelated events (e.g., falls) could confound the interpretation of mortality outcomes, as they may not reflect treatment effects. Therefore, including overall mortality without distinguishing cause-specific mortality could lead to misleading conclusions when comparing NOM and OM in elderly or frail populations.

> **Comment:** Results section - In the Results section, a potential methodological issue appears. From reading the Methods, it was understood that the inclusion of only complicated appendicitis was an exclusion criterion. However, if a study included both complicated and uncomplicated appendicitis, it should be eligible for inclusion in this scoping review. If the authors are excluding studies with both types of appendicitis, this may introduce substantial bias and could explain the small number of studies ultimately selected.

Response: *Thank you for your helpful comments! Studies that included both complicated and uncomplicated appendicitis were legible for inclusion if they provided separate or extractable data for the uncomplicated subgroup. We revised the results section to clarify this point and ensure that our inclusion approach is transparent and consistent with the methods.*

> **Comment:** "Frailty was stratified by two studies": I did not fully understand this sentence; please clarify it.

Response: *Thank you for your helpful comments. We revised the sentence for clarity. It now reads "Frailty was assessed as a key variable and used to stratify outcomes in two studies (Ashbrook et al., 2024; Chehab et al., 2021)."*
Ashbrook et al. (2024) compared overall post-operative rates of NOM vs OM groups by frailty and non-frail patients. Hence we have: 4 sets of data, in terms of comparing post-operative complication: NOM VS OM, Frailty VS non-frail and P(OM | Frailty) vs P(NOM | Frailty) and P(OM | non-Frailty) and P(NOM | non-Frailty)
Overall management comparison (NOM vs. OM):
Frailty stratification (frail vs. non-frail):
- Complication rates were substantially higher in frail patients (57.1%) compared with non-frail patients (28.8%).
- Indicates frailty strongly influences postoperative vulnerability, regardless of management type.
Conditional probabilities:
- P(OM | Frailty) = 66.3%; P(NOM | Frailty) = 33.7%
- P(OM | ¬Frailty) = 69.8%; P(NOM | ¬Frailty) = 30.2%
- Both groups favor operative management, but complication rates rise markedly with frailty.
Interpretation of effect modification:
The impact of treatment type (OM vs. NOM) on complications differs by frailty status.
Frailty modifies, rather than confounds, the relationship between management and outcomes.

Comment: Discussion section - "These findings might appear counterintuitive, as surgery is typically associated with higher perioperative risk in frail elderly patients." – It is acceptable to provide the authors' opinions, but the Discussion should primarily be a dialogue between the study's findings (in this case, the scoping review) and the existing literature, including studies on other abdominal diseases or even on complicated appendicitis. Please, provide literature information here.

Response: *Thanks for your comment. Here we are based on the paper by Felhmann et al, which we added as the reference at the end of the sentence. This study found that frailty status was directly associated with worse outcomes in emergency surgeries.*

Comment: "R, 2024" – again, is this citation corrected?

Response: *Thank you for pointing that out! We have already changed it back to Kumar, H. R. (2024). Acute appendicitis in the elderly: Current update on management: Review article. SAR Journal of Medicine, 5(3), 93–97. https://doi.org/10.36346/sarjm.2024.v05i03.005.*
Please Kindly check.

Comment: Fourth paragraph: is it not repeated information? Authors can try to join and summarize this paragraph with the second paragraph.

Response: *Thanks for your comments and suggestions. Here we believe there are two distinct roles for the second and the fourth paragraphs. In the second one, we are highlighting the possible clinical outcomes when comparing the OM with the NOM, which could be influenced by frailty status as well (that is, the paragraph is focused on the success of different treatment options in a patient-centered approach). As for the the fourth paragraph, the idea is to show how diverse the literature is about the definition of treatment success/failure with OM or NOM, mainly when considering the failure after treatment, which varied importantly throughout the studies in definitions and duration for follow-up (that is, this paragraph focus on the necessity to standardize the definition of treatment failure in order to get more homogeneous studies and more comparable results in future researches.*

Comment: Strengths and limitations subsection: In the Strengths section, the authors mention that the use of large national datasets enhances external validity, whereas in the Limitations section they state that generalizability is limited because all studies were conducted in the United States. When this type of situation occurs, it is possible to acknowledge that, writing that although using data from a single country may restrict generalizability, the inclusion of large datasets can partly offset this limitation by improving external validity. This balanced approach is recommended whenever applicable. In addition, I suggest presenting the limitations first and concluding the Discussion with the study's strengths.

Response: *Thanks for your suggestion, it was already adjusted in order to provide more precise information regarding the limitations and strengths of using a large national database.*

Apart from the advised studies, we have carefully scrutinized and verified the data throughout the manuscript and decided to make the following changes. We have inspected all the data once again, discussed extensively, and made the following revisions:
Key updates include:
- **Ashbrook et al., 2024:** Revised total case number and complication percentages (now accurately reflecting 9,065 total cases, with corrected NOM vs. OM and frailty subgroup data). Based on the availability of data, we also computed the rates of operative versus non-operative management among frail patients—P(OM | Frailty) and P(NOM | Frailty)—to allow further analysis of complications within the frailty subgroup in this study. We now have

"...", see appendix A.
- Lunardi et al., 2024: Updated case count and complication rate (1,209 cases; 10.7%).
- Meier et al., 2023: Removed the previously listed number of complication cases, as the reported figure represented the overall population rather than the number of patients with complications. Post-treatment complication data were not reported.
- Chehab et al., 2021: We removed the original complication data. The reasons are as follows: although this study compared complications between NOM and OM, the NOM group included patients who later required surgery due to failed conservative treatment. Therefore, it effectively compared OM with OM′, which is not ideal for assessing NOM vs. OM complications and could be misleading. However, we retained Clostridioides difficile infection as a complication, since it was explicitly mentioned in the text and is a legitimate complication relevant to both OM and NOM. These changes do not substantially affect our results or interpretations. We appreciate this opportunity to fine-tune our paper to ensure it accurately reflects the overall evidence.
Once again, we sincerely thank you for your efforts in helping us improve. Your valuable comments and encouragement have enabled us to produce a stronger paper, and we look forward to hearing positive news from you.

Yours faithfully,
Group 9 PPCR team

Reviewer 3

To the Reviewer 3: Dear Reviewer 3, thank you for your valuable comments and insightful suggestions. We have provided detailed, point-by-point responses to each of your comments below. In addition, we have revised the content of our paper based on your advice and clarified certain sections where necessary. We sincerely appreciate your time and consideration. The changes are as follows, and we have highlighted your original comments in bold, followed by our replies. Please kindly check.

Comment: The rationale for focusing exclusively on geriatric patients is strong, but the introduction could better distinguish how treatment considerations differ from younger adults (e.g., differences in immune response, risks of loss of independence post-surgery).

Response: *Thank you for addressing this point. We have adjusted the introduction to highlight expected differences in outcomes for the elderly.*

Comment: The results are presented clearly, but the heterogeneity of outcome definitions (e.g., treatment success, complications) is a major limitation that deserves stronger emphasis.

Response: *We appreciate the observation regarding the heterogeneity of outcome definitions across studies and agree that this represents an important limitation of our review.*
- We have further strengthened this point by explicitly noting how variations in the definitions of "treatment success" and "complications" limit comparability and synthesis of findings across studies.
- We have also briefly reiterated this as part of the Limitations paragraph to highlight its methodological relevance.

Comment: The authors appropriately highlight frailty as a major modifier. However, the conclusion that "frailty significantly modifies risks and benefits" overstates the available evidence, as only two included studies formally stratified by frailty. The conclusion should be tempered to reflect this.

Response: *Thank you for the valuable comments and insightful suggestions. We have changed to "One of the most important findings in our study is that frailty may be a key effect modifier influencing outcomes. When evaluating the treatment options for uncomplicated appendicitis for older patients, it may be important to stratify these patients*

according to functional reserve and comorbidity burden, and make individualized treatment plans, rather than treating this population as a homogeneous group. However, further research is still needed to determine the conditions under which NOM may be the most appropriate treatment option, especially among frail elderly patients. To enable meaningful comparisons between NOM and appendectomy, future studies should incorporate standardized functional status and frailty assessments, clearly defined and commonly accepted outcome measures, and detailed antibiotic protocols.". Please kindly check.

> **Comment:** Limitations should be expanded. Beyond database restriction and lack of antibiotic details, the following should be explicitly acknowledged: All included studies were retrospective and U.S.-based, limiting generalizability; No randomized trials included geriatric-only cohorts; Potential publication bias, as only English-language studies were considered. How do you believe this heterogeneity affects the comparability and reliability of your synthesis? How might cultural or systemic factors (e.g., availability of surgeons, postoperative care, antibiotic stewardship policies) affect the applicability of your results?

Response: *Thank you for the valuable comments and insightful suggestions. We have added that "Additionally, the heterogeneity in study design and population characteristics, as well as cultural and systemic factors such as differences in surgical availability, postoperative care, and antibiotic stewardship policies, may influence the comparability and applicability of our results." to our limitation.*

> **Comment:** Define "frailty" explicitly in the Methods (e.g., which indices were used across studies) --> Do you think it is justified to conclude that "frailty significantly modifies risks and benefits," or should this be framed more cautiously? Could you expand on which frailty indices were applied and whether their differences could affect the findings?

Response: *Thank you for these valuable suggestions. In the Methods section, we have clarified the definition of frailty and noted that its assessment was not uniform across the included studies. " Frailty was defined as a declined physiological reserve and increased vulnerability to stressors, leading to adverse health outcomes (Doody et al., 2023)." In the result session, we documented the frailty indices applied by each study (e.g., Modified Frailty Index [mFI-5], Clinical Frailty Scale [CFS]) and reported their corresponding thresholds as defined by the original authors.*

To illustrate frailty stratification, Ashbrook et al. (2024) compared overall postoperative complication rates between NOM and OM groups among frail and non-frail patients. Accordingly, four analytical dimensions were examined: (1) overall management comparison (NOM vs OM), (2) frailty stratification (frail vs non-frail), (3) additional analysis of conditional probabilities [Appendix B], P(OM | Frailty) vs P(NOM | Frailty), and (4) P(OM | non-Frailty) vs P(NOM | non-Frailty). Complication rates were substantially higher in frail patients (57.1%) compared with non-frail patients (28.8%), indicating that frailty strongly influences postoperative vulnerability regardless of management type. Conditional probabilities further demonstrated that operative management predominated in both groups (P(OM | Frailty) = 66.3%; P(OM | non-Frailty) = 69.8%), but complication rates rose markedly with frailty.

While the data suggest that frailty modifies the association between management type and complication risk, consistent with an effect-modification pattern. That said, we have framed this conclusion cautiously, acknowledging the heterogeneity of frailty definitions across studies and the lack of standardized measurement. We therefore interpret frailty as a potential effect modifier, recognizing that confirmatory evidence will require future studies using uniform frailty assessments.

> **Comment:** Add NOS quality assessment results in the text. How did study quality influence your interpretation of the findings?

Response: *Thank you for addressing this point, we have made changes. In the result session, we stated "The methodological quality of the included observational studies, assessed using the Newcastle–Ottawa Scale (NOS), ranged from 6 to 8.5 out of 9, indicating an overall low risk of bias across the domains of selection, comparability, and outcome assessment. In the discussion session, we further mentioned "Since our review included four observational studies with low risk NOS scores, the findings can be interpreted with moderate confidence. The consistency of low bias risk across studies strengthens the reliability of the summarized evidence despite the limited sample size.*

Apart from the advised studies, we have carefully scrutinized and verified the data throughout the manuscript and decided to make the following changes. We have inspected all the data once again, discussed extensively, and made the following revisions:
Key updates include:
- **Ashbrook et al., 2024:** Revised total case number and complication percentages (now accurately reflecting 9,065 total cases, with corrected NOM vs. OM and frailty subgroup data). Based on the availability of data, we also computed the rates of operative versus non-operative management among frail patients—P(OM | Frailty) and P(NOM | Frailty)—to allow further analysis of complications within the frailty subgroup in this study. We now have "...".
- **Lunardi et al., 2024:** Updated case count and complication rate (1,209 cases; 10.7%).
- **Meier et al., 2023:** Removed the previously listed number of complication cases, as the reported figure represented the overall population rather than the number of patients with complications. Post-treatment complication data were not reported.
- **Chehab et al., 2021:** We removed the original complication data. The reasons are as follows: although this study compared complications between NOM and OM, the NOM group included patients who later required surgery due to failed conservative treatment. Therefore, it effectively compared OM with OM′, which is not ideal for assessing NOM vs. OM complications and could be misleading. However, we retained *Clostridioides difficile* infection as a complication, since it was explicitly mentioned in the text and is a legitimate complication relevant to both OM and NOM. These changes do not substantially affect our results or interpretations. We appreciate this opportunity to fine-tune our paper to ensure it accurately reflects the overall evidence.

Once again, we sincerely thank you for your efforts in helping us improve. Your valuable comments and encouragement have enabled us to produce a stronger paper, and we look forward to hearing positive news from you.

Yours faithfully,
Group 9 PPCR team

Reviewer 4

To the Reviewer 4: Dear Reviewer 4, thank you for your valuable comments and insightful suggestions. We have provided detailed, point-by-point responses to each of your comments below. In addition, we have revised the content of our paper based on your advice and clarified certain sections where necessary. We sincerely appreciate your time and consideration. The changes are as follows, and we have highlighted your original comments in bold, followed by our replies. Please kindly check.

> **Comment:** Mortality was not addressed. Mortality is an important outcome to be considered. Previous studies reported a modest reduction in short-term morbidity, however, this benefit was offset by a small but statistically significant increase in short-term mortality (absolute increase of approximately 1.8–2%). Among the included studies, Meier, 2023 reported that in patients 65 years and older, nonoperative

management was associated a 1.82% increase in mortality (95% CI, 1.49-2.15). Lunardi, 2024 reports a 2% increase in mortality (95% CI 2%-3%). Chehab, 2021 - patients undergoing appendectomy after failed NOP had significantly higher rates of mortality (4% vs. 2%; $p = 0.019$).

Response: *Thank you for addressing this concern. Mortality is not our primary outcome of interest, or an isolated outcome of interest in our study or in the initial study designs, primarily due to its ambiguous nature. Elderly patients may die from causes unrelated to the disease course or its treatment, particularly among frail populations. We believe that focusing on mortality as a single outcome could introduce bias; therefore, it was not selected as a primary outcome. Instead, we aimed to explore outcomes measured within 3- , 6-months, and up to 1 year in our predefined analysis.*
During data extraction, we observed that some papers mentioned mortality but did not specify whether it referred to all-cause or cause-specific mortality. Even Ashbrook et al. (2024) and Lunardi et al. (2024) referred to in-hospital mortality, which is relevant but still lacked definitional clarity (all cause mortality vs treatment-related or cause-specific mortality?). Deaths resulting from comorbidities such as dementia or unrelated events (e.g., falls) could confound the interpretation of mortality outcomes, as they may not reflect treatment effects. Therefore, including overall mortality without distinguishing cause-specific mortality could lead to misleading conclusions when comparing NOM and OM in elderly or frail populations. That said, we believe that future studies, particularly analytical observational studies or randomized controlled trials, should clearly define both mortality and treatment failure.

Comment: In methods the authors declare "Studies were excluded if they focused exclusively on complicated appendicitis (defined by perforation, abscess, or peritonitis)". However, in results they say "*Most studies were excluded based on the inclusion of complicated appendicitis, early in-hospital death, or incomplete data*." I believe studies that include both complicated and uncomplicated appendicitis should have been included if outcomes of both groups are reported separately. Also, the exclusion of early in-hospital death could lead to an important bias. I wonder if the inclusion and exclusion criteria are not too strict for a scoping review, as only 4 studies were included in the final analysis. In a scoping review, the inclusion and exclusion criteria are generally broader, as the aim is to map existing evidence and identify gaps. The exclusion criteria should not be so strict that they eliminate potentially informative evidence.

Response: *Thank you for these helpful comments. We have revised the Methods to clarify eligibility for mixed populations, to ensure early in-hospital deaths are treated as outcomes (not exclusion criteria), and to explain why a focused scope was prespecified despite conducting a scoping review. We also corrected the in-text citation format in the first paragraph of Methods and in the Outcomes subsection.*
Changes made/ Addition: (Methods → Eligibility). "Studies enrolling both uncomplicated and complicated cases were eligible provided that outcomes for the uncomplicated subgroup were reported separately; when subgroup separation was not feasible, studies were excluded due to non-extractable data. Early in-hospital deaths, when reported, were included in the outcome assessment and were not used as an exclusion criterion." "Although scoping reviews often adopt broader criteria, we prespecified focused eligibility to enable a clinically relevant comparison between antibiotic-only NOM and appendectomy in older adults with uncomplicated appendicitis."

Comment: The authors state that "Across all studies, NOM was compared to OM", however one of the included studies (Lunardi, 2024) did not perform this comparison. The study compared patients according to age (<65 and ≥ 65 y; table 1), and also compared non-operative success and failure).

Response: *Thank you for this observation. We agree that Lunardi et al. (2024) did not perform a direct comparison between NOM and OM. Instead, the study compared outcomes by age group (< 65 and ≥ 65 years) and analyzed differences between successful and failed NOM cases. Our data extraction only focused on the ≥ 65 years patients clusters. Those did not contain ≥ 65 were excluded.*

Comment: Citation is incomplete in the 3rd paragraph of the introduction.

Response: *Thank you for addressing this point. It has been updated.*

Comment: Review citation format in the 1st paragraph of methods.

Response: *Thank you for addressing this point. It has been updated.*

Comment: Review format citation in outcomes (Methods)

Response: *Thank you for addressing this point. It has been updated.*

Comment: Abbreviations (NOM, OM, NR…) should be explained in a footnotes beneath the tables

Response: *Thank you for addressing this point, we have added to the explanation of the abbreviations to the latest version.*

Comment: Table 3: most columns have all cells filled with (NR), discuss the need for including this table. In addition, the absolute lack of information regarding antibiotic therapy is an important limitation.

Response: *Thank you for noticing a lack of information regarding antibiotic therapy would be an important limitation and that should be included in the discussion. We very much agreed so and therefore, in the discussion, the last paragraph, we mentioned ". This lack of standardization makes comparisons between studies difficult and may mask important differences in treatment efficacy. "*

Comment: The outcome *readmission* was included in Table 5, however there is no information regarding this outcome (I suggest excluding it from the table, maybe including in the footnote an observation that there is no information regarding readmission). As explained above, the lack of information may be related to the exclusion criteria being too strict.

Response: *Thank you for proposing these questions. We noted two main points here.*
First, regarding the secondary outcome about readmission: this is included in our study design, and therefore its results should be clearly displayed, regardless of whether they are positive or not. We believe that negative results are also meaningful results.
There are several advantages to listing missing data in studies. One is that, as a review paper, this study can directly demonstrate to what extent data are missing in current research. The second is that the missing columns themselves represent meaningful design aspects, such as timeframe, number of cases, and percentages. We envisage that future studies, both RCTs and observational studies, will pay more attention to data collection, and the missing columns we identified in our table can serve as a reference framework. By doing so, future studies can produce more standardized results, making comparisons between NOM vs. OM, or even among NOM studies alone, more feasible. Currently, there is still a lack of complete information in the existing literature.
Regarding the second point about whether our stricter selection criteria resulted in fewer included papers, we would like to respond with three clarifications. First, the use of antibiotics (NOM) as a first-line treatment is still in the experimental stage, while surgery (OM) remains the gold standard in the field; therefore, there are naturally fewer studies available. Second, the onset of appendicitis in elderly patients is relatively limited, which also reduces the number of relevant studies. Third, conducting trials of antibiotics as the first-line treatment can be ethically challenging, especially when surgery remains an available and proven option. As a result, clinical cases using antibiotics alone are limited in nature. Considering these three factors, along with our study objectives, we believe our selection criteria are justified.

Apart from the advised studies, we have carefully scrutinized and verified the data throughout the manuscript and decided to make the following changes. We have inspected all the data once again, discussed extensively, and made the following revisions:

Key updates include:

- **Ashbrook et al., 2024:** Revised total case number and complication percentages (now accurately reflecting 9,065 total cases, with corrected NOM vs. OM and frailty subgroup data). Based on the availability of data, we also computed the rates of operative versus non-operative management among frail patients—P(OM | Frailty) and P(NOM | Frailty)—to allow further analysis of complications within the frailty subgroup in this study. We now have "...".
- **Lunardi et al., 2024:** Updated case count and complication rate (1,209 cases; 10.7%).
- **Meier et al.,** 2023: Removed the previously listed number of complication cases, as the reported figure represented the overall population rather than the number of patients with complications. Post-treatment complication data were not reported.
- **Chehab et al., 2021:** We removed the original complication data. The reasons are as follows: although this study compared complications between NOM and OM, the NOM group included patients who later required surgery due to failed conservative treatment. Therefore, it effectively compared OM with OM′, which is not ideal for assessing NOM vs. OM complications and could be misleading. However, we retained *Clostridioides difficile* infection as a complication, since it was explicitly mentioned in the text and is a legitimate complication relevant to both OM and NOM. These changes do not substantially affect our results or interpretations. We appreciate this opportunity to fine-tune our paper to ensure it accurately reflects the overall evidence.

Once again, we sincerely thank you for your efforts in helping us improve. Your valuable comments and encouragement have enabled us to produce a stronger paper, and we look forward to hearing positive news from you.

Yours faithfully,
Group 9 PPCR team