Peer Review File

Article information: https://dx.doi.org/10.21037/ajo-23-1

Reviewer A：

This paper is a systematic review of facial nerve dehiscence in the setting of cholesteatoma and related other findings. The authors have done a tremendous amount of work; however the manuscript is too long (>5000words) for its relative novelty and necessitates major revision.

Comments and Author's Reply: We would like to thank the reviewer for their feedback on our manuscript and for taking the time to address areas that require revision.

1: It is not necessary to discuss all negative and postive findings, only that is relevant to the novel aspect of the paper. 2: The findings are in keeping with what is to be expected in advanced cholesteatoma, therefore the authors are requested to focus on the contribution to the literature and how this changes clinical practice

Comments and Author's Reply: We have edited the manuscript to reduce the word count by 599 words, including the parts on historical and embryologic studies on facial canal dehiscence. We have also edited the discussion section to make it more succinct, rather than discuss all the findings from the included studies. We believe this provides a better overview of the literature that remains engaging and informative to the reader, without being too lengthy.

Changes in Text: the removed sections are highlighted in the tracked changes version of the document.

The contribution of our manuscript to the literature is in estimating the prevalence of FCD in cholesteatomas, as well as in specific subgroups that can be considered risk factors for encountering dehiscence. We also review the literature on other surgical and radiographic findings that are associated with FCD, which surgeons should look for to estimate the chance that they will encounter dehiscence intraoperatively. The findings of the paper in that FCD is more prevalent in revision cases and cholesteatomatous disease of greater duration can influence pre-operative discussions and counselling and surgical planning, such as in considering a canal-wall up (CWU) or canal-wall down (CWD) procedure. Surgeons may opt to exteriorize the disease with a CWD in patients with recurrent disease or more extensive disease with FCD or other cautionary findings.

The below paragraph has been added prior to the conclusion.

"Overall, this systematic review and meta-analysis highlights that FCD is a common intraoperative finding in cholesteatoma surgery, including upto one-third of revision surgeries. Given, the potential challenges in determining FCD on preoperative imaging, it is important to consider other clinical risk factors and radiographic findings in the patient workup, discussed in our review, to estimate the likelihood of encountering dehiscence. Furthermore, this can guide preoperative patient counselling about surgical risk and assist with surgical planning, with the option to consider a canal wall down procedure in recurrent cholesteatoma or extensive disease with FCD or other associated cautionary findings."

3: Lastly, could the authors please confirm if there is a true signicance in the difference between the dehiscence rates in primary and revision surgery. Looking at the confidence intervals, it appears that it is going to be unlikely to be significant.

Comments and Author's Reply: We believe the higher dehiscence rate in revision surgery (33.54%, 95% CI: 27.30-39.78) compared to primary surgery (24.47%, 95% CI: 21.27-27.66) is statistically significant. Whilst there is some overlap in the CI between the two groups, this is minor, and does not imply statistical non-significance.

When determining significance between two groups, non-overlapping confidence intervals suggest statistical significance, whilst overlapping confidence intervals does not suggest that the difference is not statistically significant (1-4). A test of group difference was significant (p=0.01) and a meta-analysis of studies that allowed comparisons between primary and revision surgeries revealed an odds ratio of 1.67 (95% CI: 1.23-2.27). This favours higher FCD incidence in revision surgery, and the confidence intervals for the odds ratio does not cross the line of no-effect, i.e. OR = 1.0. The test for overall effect was also significant in this meta-analysis, p=0.001.

Reference

1. Cornell Statistical Consulting Unit. Overlapping confidence intervals and statistical significance. 2020. Accessed from https://cscu.cornell.edu/wp-content/uploads/73_ci.pdf

2. Austin, P., Hux, J. A brief note on overlapping confidence intervals. J Vasc Surg. 2002. 36(1):194-195.

3. Knol, M., Pestman, W., Grobbee, D. The (mis)use of overlap of confidence intervals to assess effect modification. Eur J Epidemiol. 2011; 26(4): 253-254.

4. Parasurama, P. Why Overlapping Confidence Intervals mean Nothing about Statistical Significance. Towards Data Science. 2017. Accessed from https://towardsdatascience.com/why-overlapping-confidence-intervals-mean-nothing-about-statistical-significance-48360559900a

Changes in text: nil

Reviewer B

This is a well written paper and excellent example of detailed methodology and rigorous scientific metanalysis.

Comments and Author's Reply: we would like to thank the reviewer for their positive feedback of our manuscript

Minor issues only:

- Line 446 – suggest "a high risk"

Comments, Author's Reply, and Changes in Text: we have made the change to Line 446 so that the sentence reads "with respect to patient selection where most studies had **a** high risk of bias."

Editorial Comments

1. A clear description of the systematic review's eligibility criteria allows readers to judge the applicability of findings. Eligibility criteria should include language, study type etc. Please add it in the Abstract.

Reply and Changes in Text: We have included "*The search was conducted on 25th October 2021. The selection criteria included studies published in the English literature between 1981-2021* that reported FCD incidence diagnosed intraoperatively during middle ear surgery for cholesteatomatous disease" in the abstract

2. The Abstract should report the date of the last search. This informs readers of the recency of the search. In addition, specify the timeframe of the searching.

Reply and Changes in Text: We have included "***The search was completed on 25th October 2021.*** *The selection criteria included studies published in the English literature between 1981-2021* that reported FCD incidence diagnosed intraoperatively for middle ear surgery for cholesteatomatous disease" in the abstract

3. The authors could specify the risk of bias tool or approach used (JBI Checklist for Prevalence Studies) in the Abstract.

Reply and Changes in Text: We have included "*The JBI checklist for prevalence studies was used for quality assessment of included articles*" in the abstract

4. To my knowledge, the PubMed search engine includes MEDLINE data. Please revise the statement in line 44.

Reply: PubMed does access the Medline database as well as additional sources. However, the user interface and search strategies are different in Pubmed and Medline, with the Medline interface preferable for carrying out a more comprehensive, structured, and systematic search. Hence, we used both databases, and this approach is seen in many systematic reviews. As we can see from the results, 437 results were obtained from Medline compared to 63 from Pubmed.

Changes in Text: In section 2.1, we have included "A literature review was performed on 25th October 2021 using the following databases, including studies from their earliest date of cataloguing: *PubMed (which includes MEDLINE data),* MEDLINE, Embase, and Cochrane Library.

5. Lines 128-129, "The search was limited to the English language and articles published within the last forty years", please report the exact timeframe instead of the vague reporting.

Reply and Changes in Text: We have revised the sentence to "The search was limited to the *English language and articles published between 1981-2021*."

6. Please describe when to apply random-effect model or fixed-effects model (with reference), and why the authors used the random-effect model.

Reply: A fixed-effects model is used when the studies included in the meta-analysis are similar enough, i.e. with low heterogeneity, that the measurement of the common parameter (i.e. dehiscence of the facial canal) is not affected by other parameters. However, high heterogeneity is seen with the I^2, Cochran Q, and tau^2 statistics, and from the discussion we believe this is in part due to different proportions of adult and paediatric patients, primary and revision surgeries, and different types of surgeries being included in the study cohort (e.g. tympanoplasty with or without mastoid surgery, atticotomy, etc). Given the presence of these moderator variables, we cannot apply a fixed-effects model as we believe the prevalence of FCD within each study will be influenced by these factors. A random-effects model was therefore justifiably used.

Changes in Text: In section 3.2 we have added "*Given the overall heterogeneity, presence of moderator variables, and inter-study variability in patient selection, a random-effects model was used in the meta-analysis*"

References

1) Tufanaru, C., Munn, Z., Stephenson, M., et al. Fixed or random effects meta-analysis? Common methodological issues in systematic reviews of effectiveness. International Journal of Evidence-Based Healthcare. 2015, 13(3): 196-207. DOI: 10.1097/XEB.0000000000000065

7. Authors should report the methods used to collect data from included studies, to enable readers to assess the potential for errors in the data presented. This includes how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators.

Reply: Two authors (SA, GB) independently performed the study selection and data extraction. To ensure maximal numbers of studies were included in the systematic review, we included all studies deemed eligible by a single author in the first stage of screening. A full-text manuscript review was then performed as part of final screening to include articles in the meta-analysis, with senior consultant input sought when required. The data was extracted by both authors and no discrepancies found. We did not contact the authors of the individual studies.

Changes in text:

In section 2.2 we have revised the paragraph to "*Titles and abstracts were first independently assessed by the authors (SA, GB) to screen for eligible studies by applying the inclusion and exclusion criteria described. To maximise inclusivity in the early stage of the systematic review, we included all studies deemed eligible by at least one author. A full-text manuscript of screened articles was then conducted to determine final eligibility for inclusion in the meta-analysis. If disagreements arose, the input of a senior colleague (VS) was sought until consensus was reached*."

In section 2.4 we have included "*The following data were extracted independently by the authors (SA, GB)*"

8. The authors should cite here the relevant studies in Tables 1 and 2. This is essential for those who wish to double-check these results. Please check through the manuscript.

Reply: All the 27 included studies were referenced in the manuscript as references 3, 4, 7-31

Changes in Text: We have added the citation for each article in Table 1 and 2 as requested

9. It's suggested that the authors briefly present the demographics characteristics and prevalence of FCD, despite readers can acquire this from Tables 1 and 2.

Reply: The demographic characteristics were not available for each of the 27 studies and hence was not included in the main manuscript. This was especially an issue in studies of mixed cholesteatoma and non-cholesteatoma datasets, where the authors

did not present demographic data for each pathology separately. Age data for the cholesteatoma patients was only mentioned in 13/27 studies, and gender data only in 17/27 studies. Because of the missing data we believe it's more appropriate to list the demographics for each individual study in Table 2, rather than combine the data that is available as it may not represent the entire cohort from the 27 studies.

Changes in Text: Nil changes made due to above reason

10. The authors should mention they conducted a hand search of reference lists to obtain additional articles in the Methods, since this was indicated in the flow diagram.

Changes in Text: We have included in section 2.1 "*The reference lists of all included articles were searched by the authors* to identify further articles that met the inclusion criteria. The Google Scholar database was utilised to supplement the literature review."

11. Why didn't the authors collect and retrieve the kinds of surgical approaches of the included reports? If possible, please summarize in Table 2.

Reply: This was not possible because most studies did not provide data of details on each surgical approach. In Table 2, we have listed the inclusion criteria for each study and if possible, provided the type of surgeries that were performed on the cohort. Most studies only said "primary or revision surgery for cholesteatoma," without saying how many had mastoid surgery, how many had CWU or CWD procedures, etc. This was discussed as a limitation of the meta-analysis, and we recognized inter-study variability in patient selection and surgeries performed as a likely contributor to the study heterogeneity.

Changes in Text: Nil changes made due to the above reasons

12. The two figures in Figure 3 should be clearly indicated as Figure 3a and Figure 3b. Also, please change the statement in lines 226-227 "The pooled prevalence of FCD was 24.67% (95% CI: 21.51-27.84) with the representative Forest plot shown in Figure 3".

Changes in Text:
1. Figure 3 has been changed to include the sub-labels Figure 3a and 3b as requested
2. The statement has been changed to "*The pooled prevalence of FCD was 24.67% … with the representative Forest plot shown in Figure 3B*."

13. It's recommended to present the figure results about the LOO sensitivity analysis of the pooled prevalence of FCD between adult and paediatric patients.

Changes in Text: We provide the results of the LOO sensitivity analysis in Supplement 5 as shown below.

Omitted Study	Odds ratio	Test for overall effect
None	1.83 (0.96-3.47)	Z=1.84 (p=0.07)
Arias Marzan, 2019 (7)	1.78 (0.90-3.55)	Z=1.65 (p=0.10)
Bizakis, 2006 (9)	2.25 (1.30-3.91)	Z=2.90 (p=0.004)
Gulotta, Visconti, 2020 (14)	1.54 (0.91-2.92)	Z=1.32 (p=0.19)
Gulustan, 2014 (16)	2.05 (0.92-4.55)	Z=1.77 (p=0.08)
Magliulo, 2011 (20)	1.57 (0.82-3.01)	Z=1.35, (p=0.18)
Sahin, 2019 (4)	1.97 (0.92-4.21)	Z=1.76, (p=0.08)
Shinnabe, October 2013 (27)	1.73 (0.82-3.67)	Z=1.43, (p=0.15)

Supplement 5: Leave-one-out sensitivity analysis performed for calculating the odds ratio of FCD in adult patients compared to paediatric patients. The odds ratio is calculated using the Mantel-Haenszel test with a random-effects model and 95% confidence intervals. On omission of Bizakis et al., the test for overall effect is significant and the confidence interval of the summary effect does not cross the line-of-no-effect.