Impact of endoscopic sinus surgery on Eustachian tube dysfunction in patients with chronic rhinosinusitis: systematic review and meta-analysis*

O. Rotimi, N. Mohamed, K. Steele, P. Bowles
Brighton and Sussex University Hospital Trust, Eastern Road, BN2 5BE, Brighton, United Kingdom

Abstract

Background: Eustachian tube dysfunction (ETD) has been associated with inflammatory conditions (1). Many studies have identified a high prevalence of ETD in patients with chronic rhinosinusitis (CRS) (2). However, there is a paucity of higher-level evidence assessing the impact of endoscopic sinus surgery (ESS) on patients with ETD concurrent disease.

Methods: Systematic review and meta-analysis of non-randomised studies on the impact of ESS on ETD in patients with CRS, based on the eustachian tube dysfunction questionnaire (ETDQ-7) scores. PRISMA guidelines were followed according to a priori study protocol (PROSPERO Registration number: CRD42021245677). A random-effects model was employed.

Results: 21 results were obtained using our search strategy. Four studies met our inclusion criteria. 501 patients were identified in the included studies. The prevalence of ETD in CRS patients in our review was 55.1%. Pooled estimates showed a statistically significant reduction in ETDQ-7 scores.

Conclusions: The evidence to date suggests there is a high prevalence of concurrent ETD in CRS patients, the symptoms of which improved following ESS for CRS in this patient group. However, the current evidence base is comprised of uncontrolled case series. High-quality, randomised controlled studies with long-term follow-up are lacking.

Key words: endoscopic sinus surgery, Eustachian tube dysfunction, rhinosinusitis

Introduction

Chronic rhinosinusitis (CRS) is defined based on three criteria according to European Rhinological Society. These criteria include symptoms of nasal obstruction or rhinorrhea lasting more than 12 weeks in duration with endoscopic or radiological evidence of nasal polyps; or other mucosal changes within the osteomeatal complex and/or mucopurulent discharge from the middle meatus (3). Eustachian tube dysfunction [ETD] is thought to be associated with sinonasal conditions such as rhinosinusitis (4). Given previous studies have estimated the prevalence of ETD in the general population to be 0.9% (5) and in CRS patients to be as high as 48.5% (6), this represents a common and significant health issue. ETD has been associated with significant discomfort and reduced quality of life. Various studies have investigated the impact of endoscopic sinus surgery on quality of life in individuals with concurrent ETD and rhinosinusitis. However, there is a paucity of studies investigating this at a higher level of evidence (7-9).

Objectives

Primary objective: to review the available literature on the effect of Endoscopic sinus surgery ETD symptoms, in CRS patients with concurrent ETD. Secondary objective: to review the prevalence of ETD in CRS.

Materials and methods

Data sources and search strategy

This study was produced according to Preferred Reporting Item
Pre-operative controls

Impact on patient reported outcomes of sino-nasal Endoscopic sinus surgery

Adult patients with concurrent eustachian tube dysfunction and chronic rhinosinusitis

Study selection and eligibility

Population

Intervention

Control

Outcome

Adult patients with concurrent eustachian tube dysfunction and chronic rhinosinusitis

Endoscopic sinus surgery

Pre-operative controls

Impact on patient reported outcomes of sino-nasal symptoms [ETDQ-7 and/or SNOT-22 scores]

An electronic database (Healthcare Database Advanced Search (HDAS)) was used to search through seven databases comprised of MEDLINE, CINAHL, EMBASE, BNI, AMED, EMCARE and PubMed between January 2005 and March 2022. The search terms are detailed in APPENDIX 1. Other sources included in our initial literature search were ‘Cochrane library’ and ‘NICE Evidence’ for the treatment of ETD. All the references listed by each study were interrogated, to assess for key studies not included in our results.

Patient reported outcome measures

Eustachian Tube Dysfunction Questionnaire (ETDQ-7) is a validated disease specific instrument for dysfunction of the Eustachian tube. It consists of seven questions with responses on a seven-item Likert scale ranging from 1 (no problem) to 7 (severe problem). Total scores range from 7 to 49 with a total item score ≥ 14.5 or mean item score ≥ 2.1 suggestive of significant eustachian tube dysfunction.

Study selection and eligibility

Study abstracts were independently screened for by two authors, to identify suitable studies for full-text review. In cases of disputes during the screening process, a third independent author scrutinised the study. Full-text articles were reviewed after screening using the same process to determine inclusion or exclusion of the study. Once each author had reviewed the articles, consensus for study inclusion was reached through discussion.

Table 1. PICO framework for systematic review and meta-analysis.

<table>
<thead>
<tr>
<th>PICO framework</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Adult patients with concurrent eustachian tube dysfunction and chronic rhinosinusitis</td>
</tr>
<tr>
<td>Intervention</td>
<td>Endoscopic sinus surgery</td>
</tr>
<tr>
<td>Control</td>
<td>Pre-operative controls</td>
</tr>
<tr>
<td>Outcome</td>
<td>Impact on patient reported outcomes of sino-nasal symptoms [ETDQ-7 and/or SNOT-22 scores]</td>
</tr>
</tbody>
</table>

Experimental studies, including prospective case series, randomised control trials and non-randomised trials, assessing the impact of endoscopic sinus surgery on the quality of life of patients with concurrent ETD and CRS were included in the systematic review. Further inclusion criteria include studies assessing adult patients (age > 18) utilising validated patient reported outcome measures including Eustachian Tube Dysfunction Questionnaire (ETDQ-7). Descriptive studies or reviews of practice were excluded, as were studies published in a language other than English as unfortunately, resources were not available to thoroughly evaluate these articles.

Quality assessment

Quality assessment of the included articles was performed using the ‘National Heart, Lung and Blood Institute Quality assessment tools’ (NHLBI QA tool). This is a multi-item instrument designed to assess the methodological quality and internal validity of various study types including case series. There is a different set of questions, between 8 and 12, for each type of study. The most appropriate QA tool was selected based on the study design of the article being assessed. Two authors independently used the tools which provide a rating of ‘good’, ‘fair’ or ‘poor’. Although there is no numerical output using this tool, the authors felt it was appropriate for quality assessment because detailed guidance on each question item is provided. In the event of disputes of rating, consensus was achieved through discussion with a third author.

Data extraction

The following data was extracted into an Excel spreadsheet (Microsoft Excel, 2021, Version 16.54, Microsoft Corporation) from each study: country, publication date, publication journal, recruitment period, sample size, demographic characteristics, endoscopic intervention, study effect size, patient reported outcome measures and post-operative follow-up.

Data synthesis and meta-analysis

Studies were pooled using the random-effects model. We used paired-data meta-analysis to compare the change in outcome measure Eustachian Tube Dysfunction Questionnaire (ETDQ-7) in each group. Standardised mean difference was calculated. Subgroup analysis was not performed in line with the a priori protocol.

Statistical analysis of the included studies was performed using the Cochrane Collaboration Review Manager (Revman version 5.4.1 software). Study characteristics were summarised using descriptive statistics. Heterogeneity of study outcomes was assessed using Chi-squared test and reported using the I² statistic (p < 0.05). I² of 25%, 50% and 75% corresponded to low, moderate and high heterogeneity among studies respectively. A forest
plot was produced for quantitative group comparisons.

Risk of bias assessment
Funnel plots were utilised to screen for the likelihood of significant publication bias.

Ethics approval
Ethical approval was not sought as there was no involvement of human or animal participants in this study. It was assumed ethical approval was sought for all human participants of the included studies.

Results
Search results
24 studies were identified via our database search strategy. 6 studies underwent full-text article review following removal of duplicates and exclusion of ineligible studies. Following full-text review, two further studies were excluded. 4 studies were eligible and included for qualitative synthesis and meta-analysis (Figure 1). These included studies were published between 2019 and 2020 with a maximum follow up of 9 months.

Out of the 20 excluded studies, two papers were case reports and were excluded on the basis they provided no objective evidence of the impact of endoscopic sinus surgery \(^{17,18}\). One retrospective case series was excluded because the outcome was purely observational data with no use of validated patient reported outcome measures \(^{19}\). One study was excluded from meta-analysis after full text review and quality assessment because the outcome data of one study was not relevant for this meta-analysis \(^{5}\). The data provided was not in a format that allowed reliable comparisons to be made. One study was excluded after full text review because the data was not in a format to allow useful comparisons to be made \(^{20}\). Two studies were excluded because they were systematic reviews \(^{8,21}\). One study could not be retrieved for full text review and therefore was excluded \(^{22}\). The remaining 12 studies were duplicates.

Assessment of quality and bias
On the basis of the NHLBI quality assessment tool, two authors (KS and NM) assigned a rating of ‘good’, ‘fair’ and ‘poor’ as well as a judgement whether they felt it was suitable for inclusion. In the case of disagreement, further review and discussion of the paper was undertaken until agreement reached. Overall, four of the studies included in the systematic review were given a rating of ‘good’ and one study was rated ‘fair’ (Figure 3).

Whilst Borelli et al. \(^{5}\) presented a study that was rated as ‘good,’ as per the NHLBI QA tool, the results were not relevant to our research question and therefore had to be excluded from our
Table 2. National Heart, Lung and Blood Institute Quality assessment tools (NHLBI QA tool).

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of study/assessment tool</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q7</th>
<th>Q8</th>
<th>Q9</th>
<th>Q10</th>
<th>Q11</th>
<th>Q12</th>
<th>Quality rating:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Borelli</td>
<td>Before-after; no control group</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Un known</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Good</td>
</tr>
<tr>
<td>Bowles</td>
<td>Before-after; no control group</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Un known</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Good</td>
</tr>
<tr>
<td>Higgins</td>
<td>Case-control</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N/A</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Chang</td>
<td>Case series</td>
<td>Y</td>
<td>Y</td>
<td>Not reported</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Fair</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wu</td>
<td>Before-after; no control group</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Good</td>
</tr>
</tbody>
</table>

Questions

Quality Assessment for Case Series

1. Was the study question or objective clearly stated?
2. Was the study population clearly and fully described, including a case definition?
3. Were the cases consecutive?
4. Were the subjects comparable?
5. Was the intervention clearly described?
6. Were the outcome measures clearly defined, valid, reliable and implemented consistently across all study participants?
7. Was the length of follow-up adequate?
8. Were the statistical methods well-described?
9. Were the results well-described?

Quality Assessment for Before-After (Pre-post) studies

1. Was the study question or objective clearly stated?
2. Were eligibility/selection criteria for the study population prespecified and clearly described?
3. Were the participants in the study representative of those who would be eligible for the test/service/intervention in the general or clinical population of interest?
4. Were all eligible participants that met the prespecified entry criteria enrolled?
5. Was the sample size sufficiently large to provide confidence in the findings?
6. Was the test/service/intervention clearly described and delivered consistently across the study population?
7. Were the outcome measures prespecified, clearly defined, valid reliable, and assessed consistently across all study participants?
8. Were the people assessing the outcomes blinded to the participants’ exposures/interventions?
9. Was the loss to follow-up after baseline 20% or less? Were those lost to follow-up accounted for in the analysis?
10. Did the statistical methods examine changes in outcome measures from before to after the intervention? Were statistical tests done provided p values for the pre-to-post changes?
11. Were outcome measures of interest taken multiple times before the intervention and multiple times after the intervention (i.e., did they use an interrupted time-series design)?
12. If the intervention was conducted at a group level (e.g., a whole hospital, a community, etc.) did the statistical analysis taken into account the use of individual-level data to determine effects at the group level?

Quality Assessment for Case-control studies

1. Was the research question or objective in this paper clearly stated and appropriate?
2. Was the study population clearly specified and defined?
3. Did the authors include a sample size justification?
4. Were controls selected or recruited from the same or similar population that gave rise to the cases (including the same timeframe)?
5. Were the definitions, inclusion and exclusion criteria, algorithms or processes used to identify or select cases and control valid, reliable and implemented consistently across all study participants?
6. Were the cases clearly defined and differentiated from controls?
7. If less than 100% of eligible cases and/or controls were selected for the study, were the cases and/or controls randomly selected from those eligible?
8. Was there use of concurrent controls?
9. Were the investigators able to confirm that the exposure/risk occurred prior to the development of the condition or event that defined a participant as a case?
10. Were the measures of exposure/risk clearly defined, valid, reliable and implemented consistently across all study participants?
11. Were the assessors of exposure/risk blinded to the case or control status of participants?
12. Were key potential confounding variables measured and adjusted statistically in the analyses? If matching was used, did the investigators account for matching during study analysis?

Quality Assessment for Systematic reviews and Meta-analyses

1. Is the review based on a focused question that is adequately formulated and described?
2. Were eligibility criteria for the included and excluded studies predefined and specified?
3. Did the literature search strategy use a comprehensive systematic
4. Were titles, abstracts, and full-text dually and independently reviewed for inclusion and exclusion to minimise bias?
5. Was the quality of each included study rated independently by two or more reviewers using a standard method to appraise its internal validity?
6. Were the included studies listed along with important characteristics and results of each study?
7. Was publication bias assessed?
8. Was heterogeneity assessed?

Table 3. Table of study details.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Study type</td>
<td>Prospective observational (Pre-Post study)</td>
<td>Prospective observational (Pre-Post study)</td>
<td>Case-control series</td>
<td>Retrospective case series</td>
</tr>
<tr>
<td>Baseline sample size</td>
<td>57</td>
<td>82</td>
<td>60</td>
<td>302</td>
</tr>
<tr>
<td>Target population</td>
<td>Adult participants undergoing ESS for CRS refractory to maximal medical therapy according to EPOS guidelines with no pre-existing otological disease</td>
<td>Adult participants with CRS undergoing ESS following failed medical therapy according to Internal Consensus Statement on Allergy and Rhinology</td>
<td>Adult participants with significant eustachian tube symptoms, defined as ETDQ-7 &gt; 2.1, despite up to 8 weeks of maximum medical therapy and no pre-existing otological disease</td>
<td>Adult participants undergoing ESS for CRS or RARS according to Internal Consensus Statement on Allergy and Rhinology with no pre-existing otological disease</td>
</tr>
<tr>
<td>Primary outcome measures</td>
<td>Correlation between SNOT-22 and ETDQ-7 score pre-and post-operatively</td>
<td>Differences in pre-and post-operative ETDQ-7 score, SNOT-22 total score</td>
<td>Differences in pre-and post-operative ETDQ-7 score, SNOT-22 total score</td>
<td>Differences in pre-and post-operative ETDQ-7 score, SNOT-22 total score and SNOT-22 ear 1 score</td>
</tr>
<tr>
<td>Maximum post-intervention follow-up</td>
<td>9 months</td>
<td>6 months</td>
<td>2 months</td>
<td>6 months</td>
</tr>
</tbody>
</table>

Table 4. Summary of participant characteristics from included studies.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Median Age</td>
<td>Undisclosed</td>
<td>Undisclosed</td>
<td>49.6</td>
</tr>
<tr>
<td>Diagnoses</td>
<td>Chronic rhinosinusitis with polyps, Chronic rhinosinusitis without polyps</td>
<td>Chronic rhinosinusitis</td>
<td>Chronic rhinosinusitis with polyps, Chronic rhinosinusitis without polyps, Recurrent acute rhinosinusitis</td>
</tr>
<tr>
<td>Co-morbidities</td>
<td>Asthma</td>
<td>Undisclosed</td>
<td>Cystic fibrosis, Sarcoïdosis, Granulomatosis with polyangiitis, Asthma, Allergic rhinitis, Gastroesophageal reflux disease, Tobacco smoking</td>
</tr>
<tr>
<td>Pre-op ETDQ-7</td>
<td>20.60 ± 10.34</td>
<td>15.80 ± 8.80</td>
<td>3.45 ± 1.06</td>
</tr>
<tr>
<td>Post-op ETDQ-7</td>
<td>11.40 ± 6.15</td>
<td>12.70 ± 6.80</td>
<td>2.16 ± 1.21</td>
</tr>
</tbody>
</table>

Review. Chang et al. was rated “fair” and included in the meta-analysis. The limitations of this study being that it is retrospective and had complete data for approximately 40% of their patients. The remaining three studies were all rated good by the two authors and included in the meta-analysis (Table 2).

Study characteristics
The total number of participants from the four included studies was 501 (M : F, 0.9 : 1). Baseline sample size from each included study ranged from 57 to 302 participants. Three of the four included studies were performed in United States (US) with one study being conducted in the United Kingdom (UK). Characteristics of the studies are summarised in Table 3. Participant characteristics are summarised in Table 4. 50% of the included studies (n=2) included completed questionnaires from all participants pre-operatively and post-operatively,
including at follow-up review at 2 months and 6 months \(^\text{[9,23]}\). These studies excluded patients with incomplete questionnaires, and therefore these studies exclude those who did not attend follow-up reviews or respond to questionnaires. For the remaining two studies, follow-up figures ranged from 38% to 91%. Bowles et al. \(^\text{[7]}\) had a follow-up rate of 91% at 3 months and 38% at 9 months; Wu et al. \(^\text{[6]}\) had a follow up rate of 64.6% at 3 to 6 months.

The total number of patients with clinically significant pre-operative ETD in CRS patients according to ETDQ-7 scores in the included studies (n=4) was 334. The prevalence of ETD in CRS patients in our systematic review was 55.1% \((n=334/501)\). 96.4% \((n=322/334)\) of patients with concurrent ETD and CRS had statistically significantly lower ETDQ-7 scores following ESS for CRS.

Post-operative ETDQ-7 scores
For the total comparisons \((n=4)\), there were significant reductions in ETDQ-7 scores post-operatively (Hedges’ \(g\) -0.93, 95% CI=-1.32 to -0.54, \(p=0.00001\), \(I^2=80\%\)). Leave-one-out analysis did not identify any single study that resulted in a non-significant result if omitted from the random-effects model. Omitting Chang et al. \(^\text{[9]}\) has a significant effect on the reduction of ETDQ-7 values, increasing Hedges’ \(g\)-value from -0.93 to -0.83. As indicated by the \(I^2\) values, there is evidence of significant heterogeneity in the outcomes between the included studies. Interestingly, omission of Wu et al. \(^\text{[6]}\) results in 0% heterogeneity with a significant change in the pooled estimate from -0.93 \((p=0.00001)\) to -1.14 \((p=0.00001)\). The funnel plot suggests there is evidence of reporting bias.

Discussion
Synopsis of findings
According to our systematic review and meta-analysis, of the existing evidence base, endoscopic sinus surgery is effective in improving the reported outcomes (significant reduction in ETDQ-7 scores) of CRS patients with ETD. We identified a pooled estimate of 0.93-point reduction \((p=0.00001)\) in the ETDQ-7 scores of the study participants. These findings are in line with the effect sizes reported in the literature \((6-4,23)\), suggesting a significant reduction in ETD symptoms in CRS patients undergoing endoscopic sinus surgery.

Our review has highlighted the prevalence of ETD in patients with CRS is higher than the figures quoted in prior studies. Our study found the prevalence to be 55.1% whereas in other studies this is found to be 47.6% and 48.5%, respectively \((6,9)\). This is likely to reflect the higher sensitivity of ETDQ-7 for identifying ETD when compared to SNOT-22 ear symptom subdomain scores suggesting that the prevalence of ETD in CRS may be under-reported in studies that do not use ETDQ-7. There is a strong correlation between the ETDQ-7 and SNOT-22 ear subdomain questions during assessment of patients with ETD \((9)\). ETDQ-7 has three questions, with no counterpart in the SNOT-22, that appear more specific for ETD \((9)\). However, to our knowledge, there are no studies that compare the sensitivity of SNOT-22 ear subdomain and ETDQ-7 in assessment of patients with ETD.

All four studies assess pre- and post-operative differences in patient reported outcome measures for eustachian tube dysfunction (ETDQ-7 and SNOT-22), however the aims of these studies differ. Two of the included studies were focused on elucidating factors associated with favourable ETDQ-7 response in patients undergoing ESS \((9,23)\) whilst the other two studies are focused on the prevalence and severity of ETD symptoms in patients with CRS undergoing ESS \((6,7)\).

Study heterogeneity
There is high degree of heterogeneity in the outcomes of the included studies of our systematic review. This may be explained by several factors including heterogenous definitions for significant ETD and improvement in symptoms. Significant ETD was defined by Higgins et al. \((23)\) as a mean item ETDQ-7 score > 2.1 whilst the other three studies used the total ETDQ-7 score. However, validation studies have identified a mean-item ETDQ-7 score > 2.1 was equivalent to a total ETDQ-7 score > 14.5 \((13)\).
Figure 3. Funnel plot. Funnel plot showing study outcomes are not symmetrical around the standard mean difference suggesting reporting bias is present.

Even amongst the studies that utilised total ETDQ-7 scores, there were differences in the cut-offs to define clinically significant or severe ETD. Wu et al. (6) and Chang et al. (9) used a total ETDQ-7 score > 14.5 as clinically significant because this cut-off has been proven to have 100% sensitivity and specificity for ETD. Bowles et al. (7) used a total ETDQ-7 score > 13.5 as their cut-off but no detailed explanation for this specific value was provided in their paper. With respect to clinically relevant improvement in ETD symptoms, two of the included studies aimed to measure this using the Minimal Clinically Important Difference (MCID) measure. In the studies specifically focused on the impact of ESS on ETDQ-7 scores, there are differences in definition for which figures constitute a significant MCID. In Higgins et al. study (23), a MCID > 0.5 was significant whilst Change et al. (9) used a MCID > 3.5 as significant. Bowles et al. (7) did not provide any description of significant improvement in symptoms/severity based on MCID.

Another factor to consider is the variation in study design and exclusion criteria in the analysed studies (n=4). With respect to study design, two studies stratified participants based on their subtype of CRS whilst the remaining two studies analysed their cohort as a single group. Bowles et al. (7) analysed patients based on the presence of CRS with polyps (CRSwNP) or without polyps (CRSsNP). Chang et al. (9) analysed their data using the same subtypes in addition to recurrent acute rhinosinusitis (RARS). Bowles et al. (7), found a difference in the mean pre-op ETDQ-7 scores between CRSwNP and CRSsNP patients respectively. However, these were not statistically significant (24.34 ± 9.2 vs 18.11 ± 10.3, p=0.6101). Chang et al. (9) showed similar findings that support no difference between CRSwNP, CRSsNP and RARS. With respect to exclusion criteria, three studies discussed variables such as incomplete questionnaires, history of ear surgery, presence of middle ear pathology or presence of sinonasal disease apart from sinusitis (6,9,21). However, Wu et al. (6) did not provide detailed information on their exclusion criteria. Interestingly, only one study studies explicitly detailed temporomandibular joint dysfunction (TMJD) as a confounding factor in their methodology (23). Patients with TMJD can have comparable ETDQ-7 scores even in the absence of eustachian tube dysfunction (24).

The pre-operative burden of disease is another important consideration that may account for heterogeneity. Omission of Wu et al. (6) data during sensitivity analysis results in 0% heterogeneity and maintenance of the statistically significant pooled estimate. Comparison of the studies show Bowles et al. (7) and Chang et al. (9) had much higher mean pre-operative total ETDQ-7 scores compared to the other studies, suggesting their patient cohort may have had a greater burden of disease. The cohort in Wu et al. (6) study had mean pre-op ETDQ-7 score is just above the threshold of clinically significant disease. The authors hypothesise as the burden of ETD symptoms according to ETDQ-7 increase, the greater the post-operative benefit reported by patients. Studies assessing the impact of balloon tuboplasty on ETD symptoms in patients with CRS have reported similar findings about high pre-operative disease which corroborate this hypothesis (25,26).

The lack of long-term follow-up, and the relatively high proportion of patients lost to follow-up in the included studies imply the risk of reporting bias.

Implications and challenges
ETD and CRS are common therefore expanding our understanding of these conditions and the relationship between them may help improve quality of life outcomes for this patient group. ETD has been divided into three subtypes (27). These recommendations were made because of the differences in patient characteristics and presenting complaints suggesting different underlying aetiologies at play in ETD (27,28). CRS has been associated with dilatory ETD, which is thought to be due to the relation between dynamic dysfunction (muscular failure) or functional obstruction of the eustachian tube orifice or lumen and postsaline space inflammation (27). The improvements in ETDQ-7 scores from our meta-analysis further support the association between CRS and dilatory ETD.

Strengths and limitations
Our search strategy and methodology were developed in line with the PRISMA guidelines for systematic reviews (10). However, there is a lack of high-quality studies, with the evidence base to date comprised largely of uncontrolled case series, within incomplete and/or short-term follow-up. Heterogeneity of the existing studies limits outcome comparison.
Conclusion
There is a high degree of heterogeneity in outcomes of studies evaluating the impact of endoscopic sinus surgery on eustachian tube dysfunction symptoms due to variation in study design and varied exclusion criteria amongst the studies. However, the evidence suggests that endoscopic sinus surgery has a significant impact reducing the severity of concurrent eustachian tube dysfunction in patients suffering with chronic rhinosinusitis. Further research is required in the form of well-designed randomised controlled studies, with long-term follow-up.

Acknowledgments
None.

Funding
None.

Authorship contribution
OR conceptualised the study and developed the protocol. OR and PB reviewed the abstracts, papers and extracted the data. All authors (OR, KS, NM, PB) contributed to the analysis and interpretation of the results and drafting of the manuscript. All authors approved the manuscript for publication. PB is the Guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Ethics approval and consent to participate
Not applicable

Conflict of interest
None to declare.

References
APPENDIX 1. Search terms.

Search terms utilised in Healthcare Database Advanced search using seven databases including MEDLINE, CINAHL, EMBASE, BNI, AMED, EMCARE and PubMed.

Search terms included (endoscopic sinus surgery) AND (chronic rhinosinusitis) AND (eustachian tube dysfunction), (endoscopic sinus surgery) AND (rhinosinusitis) AND (eustachian tube dysfunction), (endoscopic sinus*) AND (rhino*) AND (eustachian tube dysfunction), (Balloon tuboplasty) AND (chronic rhinosinusitis) AND (eustachian tube dysfunction).

