Reviewing the Systematic Reviews in OSA Surgery

Victor Certal, Naoya Nishino, Macario Camacho and Robson Capasso

Otolaryngology -- Head and Neck Surgery published online 23 October 2013
DOI: 10.1177/0194599813509959

AMERICAN ACADEMY OF OTOLARYNGOLOGY–HEAD AND NECK SURGERY FOUNDATION

Otolaryngology–Head and Neck Surgery
XX(X) 1–13

DOI: 10.1177/0194599813509959
http://otojournal.org
SAGE

Reviewing the Systematic Reviews in OSA Surgery

Victor Certal, MD[1,2], Naoya Nishino, MD[3], Macario Camacho, MD[4], and Robson Capasso, MD[4]

No sponsorships or competing interests have been disclosed for this article.

Abstract

Objective. There is an extensive amount of literature on surgeries as treatment for obstructive sleep apnea syndrome on adults. Previous systematic reviews have been performed to summarize the outcomes for sleep surgeries, with conflicting results. The objective of this study was to critically evaluate these systematic reviews to provide an overview of their quality, strengths, and conclusions.

Data Sources. MEDLINE, Scopus, and the Cochrane Collaboration databases were searched from inception to April 2013.

Review Methods. An overview of systematic reviews was undertaken. Studies included in this review are the systematic reviews whose primary objective was to evaluate the outcomes of sleep apnea surgery on adults. The methodological quality of the studies was analyzed with AMSTAR checklist, and the quality of evidence was evaluated using the GRADE assessment tool. Primary outcome measures assessed the effect of surgery on snoring, sleepiness, and the apnea-hypopnea index.

Results. A total of 11 studies were included in this study, and the pooled overview includes 378 studies. The systematic reviews were mostly graded as low quality using the GRADE tool and low to moderate according to the AMSTAR checklist. Outcome for apnea-hypopnea index demonstrated substantial variation leading to conflicting results. Despite a high amount of heterogeneity, outcomes for sleepiness and snoring demonstrated significant improvement across included reviews.

Conclusions. Although obstructive sleep apnea surgery is associated with improved outcomes in most studies, the level and quality of evidence reviews requires improvement.

Keywords

adult sleep apnea, systematic review, obstructive sleep apnea syndrome, sleep surgery

Received June 24, 2013; revised September 20, 2013; accepted October 2, 2013.

Introduction

Obstructive sleep apnea (OSA) is a condition characterized by recurring episodes of increase in upper airway resistance and decreased airflow due to upper airway collapse or narrowing during sleep.[1] Common symptoms include loud snoring, breathing interruptions, excessive daytime sleepiness, and cognitive impairment. Its association with an increased risk of cardiovascular complications is well described.[2]

The primary course of treatment for OSA is therapy with continuous positive airway pressure (CPAP) devices. However, CPAP compliance is often low, and in cases where effectiveness is not optimal, surgery is often utilized as an alternative after medical management has failed.[3] Nasal surgeries may be used to improve CPAP compliance in patients with nasal obstruction.[4]

Several procedures may be used to increase upper airway patency or to decrease its collapsibility, anywhere from the nasal cavity down to the hypopharynx, including the neck and the maxillomandibular complex. These procedures are often combined as part of a multiple procedures instead of an isolated procedure.[5]

Although there is extensive literature regarding these surgeries and a considerable number of reviews comparing the different procedures and techniques, a systematic overview of these reviews is lacking. This article provides an overview of these systematic reviews and provides a comparison of their findings. The objective was to identify the existing published reviews on OSA surgery, evaluate their quality, and discuss the strength of the various conclusions.

[1]Department of Otorhinolaryngology, Hospital Sao Sebastiao, Sta Maria da Feira, Portugal
[2]CINTESIS–Center for Research in Health Technologies and Information Systems, University of Porto Porto, Portugal
[3]Kansai Medical University, Osaka, Japan
[4]Sleep Surgery Division, Department of Otolaryngology-Head & Neck Surgery, Stanford University School of Medicine, Stanford, California, USA

Corresponding Author:
Victor Certal, Department of Otolaryngology-Head & Neck Surgery, Stanford University School of Medicine, 801 Welch Road, Stanford, CA 94305-5739, USA.
Email: victorcertal@gmail.com

Methods

Selection Criteria

For proper identification of studies eligible for the analysis, the study selection criteria were defined before data collection. Only systematic reviews whose primary objective was to evaluate the efficacy of surgery to treat OSA on adults were selected. Reviews that predominantly evaluated other treatment modalities (eg, oral appliances, positive airway pressure devices, positional therapy, or adjunctive treatment modalities) or focused on pediatric population were excluded. Authors sought to identify and include only those papers that met standard criteria for systematic reviews. This required that the review had to have used and described adequate methods to search for, appraise, and describe the included studies. No restrictions in the inclusion criteria of the identified reviews were applied regarding participants, details of the interventions, or outcomes of interest.

Search Strategy

Our primary method to identify potentially eligible studies was by utilizing a computerized literature search in the MEDLINE database from inception to April 2013, without any restriction on the language of publication. The following search keywords and MeSH terms were used: *sleep apnea, surgery, nose, palate, tongue, tongue base, hypopharynx, hyoid, epiglottis, genioglossus, radiofrequency, advancement, glossectomy, suspension*, and *stabilization*; the following terms were used for systematic reviews: *systematic review* and *meta-analysis*. Literature searches were also undertaken, using the same search keywords, in the following databases: the Cochrane collaboration and SCOPUS database. To minimize the risk of missing relevant reviews, a manual search of key journals and of the reference lists of reviews captured by the initial searches was performed.

Study Quality Assessment and Data Abstraction

The titles and abstracts of the retrieved studies were independently screened for relevance by 2 reviewers (VC and NN). As currently recommended for overviews of reviews,[6] the reviewers evaluated the quality of included studies at 2 levels: the methodological quality (AMSTAR) and the quality of the evidence (GRADE). For both assessments, the criteria were applied independently. All disagreements were resolved by consensus.

The AMSTAR checklist is validated as a means to assess the methodological quality of systematic reviews and was used in overview of reviews to determine if the potentially eligible reviews met minimum requirements on the basis of quality.[7]

The GRADE approach for appraisal of the quality of the evidence suggests the initial separate consideration of 4 categories of reasons for rating down the quality of evidence and 3 categories for rating up, with a yes/no decision regarding rating up or down in each case. The GRADE approach results in an assessment of the quality of a body of evidence as high, moderate, low, or very low.

The primary outcome measures included apnea-hypopnea index (AHI), Epworth Sleepiness Scale (ESS), and snoring improvement. Secondary outcome measures included any other outcome measures evaluated. To reduce heterogeneity among findings, whenever it was feasible, authors grouped results by the surgical approach type: single procedure or multiple simultaneous procedures (occasionally called multilevel surgery). In the case of maxillomandibular advancement (MMA), authors presented results separately, as it is skeletal surgery and quite commonly it is preceded by previous upper airway surgeries for OSA.

Results

All database searches were performed in March 2013. A flow chart of the process of study identification is shown in **Figure 1**.

Two hundred fifty-three articles were identified using the search strategy and sources listed previously. After the titles and abstracts were screened for relevance, 220 articles were excluded. The remaining 33 articles were retrieved for more detailed full-text evaluation, and 23 were excluded for the following reasons: 7 were not surgical reviews,[8-14] 10 focused on pediatric OSA,[15-19] and 6 were not actually systematic reviews and did not meet minimum methodological inclusion criteria.[20-25] One study[26] was included after a manual search of references of the included studies. A total of 11 reviews[1,3,5,26-33] were included in the final overview.

Description of Included Reviews

The included reviews contained a total of 378 primary studies. Four systematic reviews analyzed only single procedures, 1 review analyzed only multiple procedures, 1 review presented results for single and multiples procedures, 3 reviews presented results for single and/or multiple procedures together with MMA surgery, and 2 reviews analyzed only MMA surgery results. Substantial heterogeneity was found in methodological inclusion criteria, but most of the included reviews did not have any study design limitation.

Methodological and Evidence Quality of Included Reviews

The main characteristics of the included studies are presented in **Table 1**. The systematic reviews were mostly graded as low quality (**Tables 2 and 3**). The Cochrane review[1] scored higher, partly because of standardized reporting and greater word allowances, relative to peer-reviewed journals, to allow for greater detail, which permitted us to more completely evaluate methods and risk of biases. The other non-Cochrane reviews were low-quality reviews mostly based on nonrandomized studies. All but one[28] of the included reviews failed to perform any kind of publication bias evaluation. Publication bias is a systematic underestimate or an overestimate of the underlying beneficial or harmful effect due to the selective publication of studies. In the included systematic reviews, it is unclear how the authors managed the extent to which they were uncertain about the magnitude of the effect due to selective

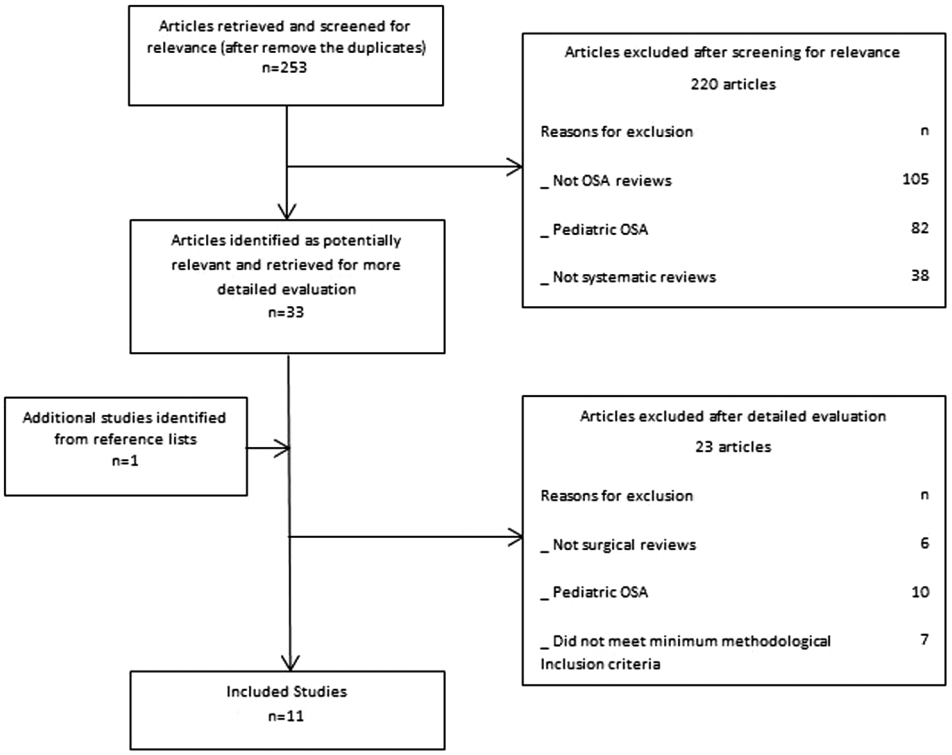

Figure 1. Flow diagram of study identification and selection. OSA indicates obstructive sleep apnea.

publication of studies. For example, Franklin et al[30] used different inclusion criteria for searching primary studies, according to the outcomes of interest (only randomized controlled trials [RCTs] for evaluation of efficacy and minor study design limitation for the evaluation of potential harm of surgery). The combination of these factors limit the ability to assess the external validity of the studies included in the reviews.

All studies satisfied 4 or more of the 11 items in the AMSTAR checklist (**Table 3**) for the assessment of methodological quality and were therefore considered low to moderate quality. The major methodological limitations include the lack of assessment of the scientific quality of the included studies and the lack of publication bias evaluation. Also, 6 included reviews based their comprehensive literature searching on only 1 database, which could have led to a source-selection bias

Outcomes

Findings are summarized in **Table 4**. All the included reviews presented substantial variation in the outcome definition and in the presentation of results. We therefore did not report any pooled result or perform statistical analysis other than that already performed by the authors and presented individual main results and range of effect for each outcome according to the outcome definition in each review.

Apnea-Hypopnea Index

The results regarding the impact on AHI were reported in all the included reviews. Heterogeneity of this impact was evident; even categorizations such as single-level, multiples procedures, and MMA showed significant differences among reviews. Most of the reviews[5,26,27,29,31,33] presented their results as overall success (defined by AHI <20 events/hr and 50% reduction from baseline, pre surgical state), but some reviews[1,3,28,32] reported AHI reduction or did not provide any numeric data to support surgical success.[30] Elshaug et al[29] used different definitions of surgical success, thereby showing a substantial degree of heterogeneity according to the variation of the success criteria (phase I procedures or soft tissue procedures: from 13% to 55%; phase II procedures or skeletal procedures: from 43% to 86%) and suggested that all surgical articles should report ''objective cure'' rates with success based on AHI outcomes of ≤ 5 and/or ≤ 10. Lin et al[5] showed an average success rate (excluding MMA) of 59.2% and advocated multiple procedures as the standard procedure in sleep surgery because it produced superior results.

Sleepiness

The results regarding the impact on Epworth Sleepiness Scale or other psychometric validated tool such as the Stanford Sleepiness Scale were not reported in all the included reviews. Two reviews[26,29] did not provide any information about this outcome, and substantial variation was found in the presentation of results. However, despite this heterogeneity, only 1 review[30] found no significant reduction in ESS outcome, and 1 subgroup (LAUP subgroup) of Sundaram et al's review[1] also failed to show any effect on sleepiness.

Table 1. General characteristics of included systematic reviews.

Author	Aim	Search database	Type of surgery	Number of studies	Total of participants	Study design
Caples et al 2010[27]	Provide a basis for standards of practice recommendations on surgical modifications of the upper airway for treatment of OSA	Medline Embase Current Contents Cochrane	Single procedure (LAUP, UPP, UPPP, TCRAFTA, palatal implant; MMA) Multiples procedures	Total: 79 studies Single procedure: 47 studies Multiples procedures: 32 studies	3913 participants	Randomized and nonrandomized controlled trials, cohort studies, and case series (sample size $>$ 1)
Choi et al 2012[28]	Efficacy of Pillar implants on OSA and snoring	Medline LILACS Scopus Cochrane	Single procedure: Pillar implants	7 studies	174 participants	No study design limitation No randomized trials were identified
Elshaug et al 2006[29]	Highlight how surgical "success" rates decrease when contemporary, evidence-based criteria of effectiveness are applied	Medline	Single or multiples procedures (UPPP, genioplasty, MMA, GA, LAUP, TCRFTA, hyoid suspension)	Phase I procedures: 14 studies Phase II procedures: 4 studies	Phase I procedures:347 patients Phase II procedures: 38 patients	No study design limitation
Franklin et al 2009[30]	Efficacy and adverse effects of surgery for snoring and OSA	PubMed Cochrane	Only single procedure (LAUP, UPP, UPPP, TCRAFTA)	Efficacy: 4 studies Adverse effects: 45 studies	Efficacy: 145 participants Adverse effects: 6016 participants (3130 enrolled in 1 study)	Efficacy: Only RCTs of medium and high quality Adverse effects: randomized, controlled studies and observational studies of medium and high quality
Holty and Guilleminault 2010[3]	Clinical efficacy and safety of MMA in treating OSA	Medline	MMA as single procedure (33% of participants) MMA with previous or concurrent phase-I surgery (67% of participants)	22 studies	627 participants	No study design limitation No randomized trials were identified
Kezirian et al 2006[31]	Review of hypopharyngeal surgery outcomes in OSA	Medline	Single procedure: GA, mortised genioplasty, tongue radiofrequency, midline glossectomy, hyoid suspension, tongue stabilization Multiples procedures	36 studies	973 participants	No study design limitation 94% were case series

(continued)

Table 1. (continued)

Author	Aim	Search database	Type of surgery	Number of studies	Total of participants	Study design
Li et al 2011[32]	Analyze the efficacy of nasal surgery in the treatment of OSA	Medline	Single procedure: nasal surgery	13 studies	474 participants	Prospective/retrospective clinical trial, controlled, and randomized
Lin et al 2008[5]	Efficacy of multiples procedures for OSA patients	PubMed Cochrane Database Medline	Multiples procedures	49 studies	1978 participants	No study design limitation
Pirklbauer et al 2011[33]	Review the published data and evaluate the effectiveness of MMA for OSA	PubMed	MMA	39 studies	1213 patients	No study design limitation
Sher et al 1996[26]	Efficacy of surgery for OSA patients	Medline	Single procedure: nasal procedures, UPPP, tongue procedures Multiples procedures: GA with hyoid myotomy and suspension MMA	54 studies	1707 participants	No study design limitation
Sundaram S. et al 2013[1]	Assess the effects of any type of surgery for the treatment of the symptoms of OSA	Cochrane Central Medline EMBASE CINAHL	Single procedure: TCRFTA, LAUP, UPPP, palatal implants	12 studies	709 participants	Only RCT were included

Abbreviations: GA, genioglossus advancement; UPP, uvulopalatoplasty; UPPP, uvulopalatopharyngoplasty; LAUP, laser-assisted uvulopalatoplasty; TCRFTA, temperature-controlled radiofrequency tissue ablation; MMA, maxillomandibular advancement; RCT, randomized controlled trial; OSA, obstructive sleep apnea.

Table 2. Quality assessment of evidence GRADE.[a]

Author	Limitations	Inconsistency	Indirectness	Imprecision	Publication bias	Quality of evidence (GRADE)
Caples et al 2010[27]	Serious	No serious inconsistency	No serious indirectness	Serious	—	Low
Choi et al 2012[28]	Very serious	No serious inconsistency	No serious indirectness	No serious imprecision	Undetected	Low
Elshaug et al 2006[29]	Very serious	No serious inconsistency	No serious indirectness	Very serious	—	Very low
Franklin et al 2009[30]	Very serious	Serious	No serious indirectness	Serious	—	Very low
Holty and Guilleminault 2010[3]	Very serious	No serious inconsistency	No serious indirectness	No serious imprecision	—	Low
Kezirian et al 2006[31]	Very serious	Serious	Serious	No serious imprecision	—	Very low
Li at al 2011[32]	Very serious	No serious inconsistency	No serious indirectness	No serious imprecision	—	Low
Lin et al 2008[5]	Very serious	No serious inconsistency	No serious indirectness	No serious imprecision	—	Low
Pirklbauer et al 2011[33]	Very serious	Serious	No serious indirectness	No serious imprecision	—	Very low
Sher et al 1996[26]	Very serious	Serious	No serious indirectness	Serious	—	Very low
Sundaram et al 2013[1]	Serious	No serious inconsistency	No serious indirectness	No serious imprecision	—	Moderate

[a]GRADE Working Group grades of evidence: high quality, further research is very unlikely to change our confidence in the estimate of effect; moderate quality, further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate; low quality, further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate; very low quality, we are very uncertain about the estimate. The downgrading in the grades of evidence (decrease quality of evidence) is based on the assessment of 5 factors: limitations in design, inconsistency in results, indirectness of evidence, imprecision of results, and publication bias. Upgrading of evidence (increase quality of evidence) may occur based on the assessment of 3 factors: the magnitude of effect, influence of all residual confounding, and the dose-response gradient. Rating factors that influence quality of evidence within each category were summarized in 3 classes (no serious, serious, and very serious) according to GRADE guidelines. Rating limitation category should assess if the studies had limitations in design or execution (risk of bias); rating inconsistency category should assess if the results were consistent across studies; rating indirectness category should assess if the evidence answered directly the study question; and rating indirectness category should assess if the results were precise enough. For each category, the objective is to evaluate if any problem may have been serious enough to downgrade the quality of evidence for this outcome.

Table 3. Methodological quality of the included reviews (AMSTAR checklist).

AMSTAR checklist	Caples et al 2010[27]	Choi et al 2012[28]	Elshaug et al 2007[29]	Franklin et al 2009[30]	Holty and Guilleminault 2010[3]	Kezirian et al 2006[31]	Li et al 2011[32]	Lin et al 2008[5]	Pirklbauer et al 2011[33]	Sher et al 1996[26]	Sundaram et al 2011[1]
1. Was an a priori design provided?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2. Was there duplicate study selection and data extraction?	Yes	Yes	Partially	Yes	Yes	Partially	Yes	Yes	Partially	Yes	Yes
3. Was a comprehensive literature search performed?	Yes	Yes	No	Yes	No	No	No	Yes	No	No	Yes
4. Was the status of publication (ie, grey literature) used as an inclusion criterion?	No	No	No	No	No	No	No	No	No	No	No
5. Was a list of studies (included and excluded) provided?	No	No	No	Yes	Yes	No	No	No	No	No	Yes
6. Were the characteristics of the included studies provided?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
7. Was the scientific quality of the included studies assessed and documented?	No	No	No	No	No	No	No	No	No	No	Yes
8. Was the scientific quality of the included studies used appropriately in formulating conclusions?	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
9. Were the methods used to combine the findings of studies appropriate?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
10. Was the likelihood of publication bias assessed?	No	Yes	No	No	No	No	No	No	No	No	No
11. Was the conflict of interest stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes

Table 4. Summary of findings.

Author	Effect on AHI	Effect on ESS	Effect on snoring	Other findings
Caples et al 2010[27]	Single procedure: Reduction of AHI variable (range, 21%-33% AHI reduction) Multiples procedures: Overall mean success range (40%-78.3%) defined by AHI $<$ 20 and 50% reduction. MMA: overall reduction in AHI of 87% with a mean postoperative AHI of 7.7	Single procedure: Improvement in sleepiness (no measure stated) Multiples procedures: 61% of patients were deemed successfully treated MMA: Reduction from 17.8 to 4.7	Single procedure: significant reduction Multiples procedures: NR MMA: NR	Improvements in blood pressure 6 months after MMA Data on nonrespiratory sleep outcomes, quality of life, and cardiovascular/systemic interactions are sparse
Choi et al 2012[28]	Single procedure (Pillar implants): substantial degree of heterogeneity (range, −17% to 50% AHI reduction) Significant overall standardized mean difference of −0.378 (CCI, −0.619 to −0.138, P = .002)	Substantial degree of heterogeneity (range, 5%-43% ESS reduction)	Range from 25% to 52% snoring reduction on visual analogue scale	Extrusion rate of 9.3%
Elshaug et al 2006[29]	Single procedure: Overall mean success range (48.8% to 60.8%) defined by AHI $<$ 20 and 50% reduction Multiples procedures: Overall mean success of 55% defined by AHI $<$ 20 and 50% reduction MMA: Overall mean success of 86% defined by AHI $<$ 20 and 50% reduction	NR	NR	Authors proposed that all surgical audits should report "objective cure" rates with success based on AHI outcomes of $\leq$ 5 and/or $\leq$ 10 Substantial degree of heterogeneity according to the definition of success Phase I procedures (overall success): from 13% to 55% Phase II procedures (overall success): from 43% to 86%.
Franklin et al 2009[30]	Single procedure: Conflicting results (reduced in one trial after laser-assisted uvulopalatoplasty but not in another trial)	No significant effect	Significant effect	No effect on quality of life (through the Calgary Sleep Apnea Quality of Life Index) Persistent side effects occurred after UPPP and UPP in about half the patients and difficulty in swallowing, globus sensation, and voice changes were especially common
Holty and Guilleminault 2010[3]	MMA: significant reduction (mean AHI decreased from 63.9/h to 9.5/h)	MMA: Significant reduction (mean ESS decreased from 13.2 to 5.1)	NR	Significant improvement in all domains of the functional outcomes of sleep questionnaire

(continued)

Table 4. (continued)

Author	Effect on AHI	Effect on ESS	Effect on snoring	Other findings
				Reduction in reported symptoms of depression, irritability, morning headaches, memory loss, and impaired concentration. Statistically and clinically relevant improvements in blood pressure Younger age, lower preoperative weight and AHI, and greater degree of maxillary advancement were predictive of increased surgical success Low complication rate
Kezirian et al 2006[31]	Single procedure: Overall mean success range (35% to 62%) defined by AHI < 20 and 50% reduction. Multiples procedures[a]: Overall mean success range (22% to 77%) defined by AHI < 20 and 50% reduction	Significant improvement	NR	Significant improvement in quality of life Several factors such as the body mass index apnea-hypopnea index, Friedman stage, and SNB angle on lateral cephalogram have been associated with surgical outcomes
Li et al 2011[32]	Single procedure (nasal procedures): nonsignificant reduction (mean AHI decreased from 35.2/h to 33.5/h)	Mean ESS scores decreased significantly from 10.6 to 7.1	Significant in bed partner's snoring VAS	
Lin et al 2008[5]	Multiples procedures: overall redefined success[b] rate was 66.4%; the redefined success[b] rate (excluding MMA) was 59.2% The weighted average percentage change of AHI showed a significant 60.3% improvement after multiple procedures	The weighted average percentage change improved by 43.0% postoperatively	65.1% decrease in bed partner's snoring VAS	8.8% increase in quality of life The percentage changes of weighted average on mO2 and BMI revealed no significant difference after surgery The recalculated success rates with a meta-analysis using the commonly agreed upon criteria (postoperative AHI of 50% or more and an AHI of less than 20) were 56.5% for mild/moderate disease and 69.3% for severe disease, respectively
Pirklbauer et al 2011[33]	MMA: Mean success rates ranged from 65% to 100%, depending on the criteria defined by the investigators	Reduced after MMA in all studies	NR	The Functional Outcomes of Sleep Questionnaire showed improvement in quality of life in patients with OSA treated with MMA

(continued)

Table 4. (continued)

Author	Effect on AHI	Effect on ESS	Effect on snoring	Other findings
				Long-term success rates of 80% during an observation period of 24 months
Sher et al 1996[26]	Single procedure: nasal procedures: no significant differences between baseline and postoperative values in AI UPPP: overall mean success rate defined by 50% decrease in AI: range from 18.2% to 90% Tongue procedures: overall mean success rate defined by 50% decrease in AI: range from 41.7% to 77% Multiples procedures: overall mean success rate defined by 50% decrease in AI: range from 67% to 79.2% MMA: overall success rate defined by 50% decrease in AI: range from 77.8% to 97.8%	NR	NR	The site of pharyngeal narrowing or collapse, although identified by different and unvalidated methods, has a marked effect on the probability of success of uvulopalatopharyngoplasty Patients who achieve a favorable response with uvulopalatopharyngoplasty tend to have less severe obstructive sleep apnea than those who do not
Sundaram et al 2013[1]	Single procedure: CRFTA[c]: no significant difference LAUP[c]: 21% reduction in mean AHI UPPP[c]: no significant differences were observed between the 2 groups Palatal implants[c]: mean change in AHI was significantly greater when compared with placebo	TCRFTA[c]: mean difference in change: −1.1; significant greater reduction in symptoms in favor of TCRFTA compared with placebo LAUP[c]: no significant difference UPPP[c]: a significant difference was observed in favor of surgery over conservative management Palatal implants[c]: ESS scores improved more with implants than placebo	NR	TCRFTA[c]: quality of life: there was a significantly greater improvement on the Functional Outcomes of Sleep Questionnaire (FOSQ) in favor of TCRFTA LAUP[c]: there was no significant difference between LAUP and control in Calgary Sleep Apnoea Quality of Life Index (SAQLI) scores Palatal implants[c]: surgery led to better quality of life when compared with placebo (SF-36–Short Form 36 Health Survey)

Abbreviations: UPP, uvulopalatoplasty; UPPP, uvulopalatopharyngoplasty; LAUP, laser-assisted uvulopalatoplasty; TCRFTA, temperature-controlled radiofrequency tissue ablation; MMA, maxillomandibular advancement; RCT, randomized controlled trial; OSA, obstructive sleep apnea; AHI, apnea-hypopnea index; ESS, Epworth Sleepiness Scale; BMI, body mass index; NR, not reported.

[a]The patient populations, on average, were overweight or obese.

[b]The definition of success used by the authors of the various papers reviewed was not consistent. Linet al[5] redefined the success rate to be consistent with the commonly agreed upon criteria, namely, "a reduction in AHI of 50% or more and an AHI of less than 20."

[c]Compared with conservative management or placebo.

Snoring

Five reviews[5,27,28,30,32] contained data regarding improvement in snoring. Although there was considerable heterogeneity in the criteria for snoring improvement, all the reviews reported significant reduction.

Other Outcomes

Seven reviews[1,3,5,27,30,31,33] contained data regarding the quality of life. Only 1 review[30] found no significant improvement in quality of life, and 1 subgroup (LAUP subgroup) of Sundaram et al's review[1] also failed to show any effect on this outcome. In 1 review, MMA surgery showed positive results in several outcomes such as improvement in blood pressure and reduction of symptoms of depression, irritability, morning headaches, memory loss, and impaired concentration.[3]

Discussion

The reviews present conflicting results regarding not only methods and indications of OSA surgery, but also definition of a successful surgery.

The different inclusion/exclusion criteria used in different reviews may have influenced the respective results and conclusions. Franklin et al[30] only included RCTs for the evaluation of the efficacy of surgical treatment, leading to the exclusion of most of the published surgical literature, thus limiting included studies to minimally invasive treatments, which by nature are less comprehensive treatments for OSA and some are not meant for OSA surgical procedures. These facts not only weaken their conclusions but also strongly affect the external validity of the results. There is debate about the feasibility of RCTs for surgical interventions[34] and the superiority of RCTs over non-RCTs or observational designs.[35] In reality, experimental and observational studies contribute to complementary evidence, as it is the research question that dictates the appropriate study design. Solomon et al[36] performed a systematic review of treatment evaluation questions and concluded that only 40% of treatment questions involving surgical procedures could have been evaluated by an RCT. It is important to recognize the value of evidence from non-RCTs that evaluate OSA surgical interventions when the conduct of RCTs is impractical or unethical, which could apply to the sleep-surgery field. Relying exclusively on RCTs may have led to a set of reports that were not representative of all the relevant literature that would have been identified through a search with no study design restriction.

Half of the included reviews[3,26,29,31-33] based their research on only 1 database, which could have led to a source-selection bias. A search of MEDLINE alone is not considered adequate.[6] A systematic review showed that only 30% to 80% of all known published randomized trials were identifiable using MEDLINE.[37] Going beyond MEDLINE is important not only to ensure that as many relevant studies as possible are identified but also to minimize selection bias for those that are found.

The findings of this overview support that sleep apnea surgeries are associated with improved outcomes, but the level and quantity of evidence reviews remain low because of the following factors: difficulty in standardizing procedures (the surgical approaches seem to vary according to the upper airway anatomy), experience of the study center with all surgical procedures[38,39] (few surgical centers have the experience to perform all available techniques), and lack of long-term follow-ups.

The commonly accepted criteria for ''success'' is a 50% or more reduction in AHI from baseline preoperative values, with an AHI of less than 20 with an associated improvement in the condition of the patient. Others maintain that ''success'' should be defined as a postoperative AHI of less than 5 or 10. When applying the latter definition of success to multiple procedures, there is a significant decrease in the percentage of overall surgical success rates; however, it is the authors' opinion that this should not be used as an argument to rule out the utility and value of surgical treatment. The goal of surgery is not solely to cure OSA but to control symptoms and minimize ongoing multisystem damage for those who cannot or will not accept CPAP therapy. For instance, most of the reviews indicate that sleepiness, snoring, and even quality of life improve with many various surgical treatments, even if OSA is not cured.

In addition, the surgeries can also lead to higher CPAP compliance by allowing the patient to reduce the amount of pressure required to operate the device, resulting in less discomfort.[40] Another major complication of OSA is its association with cardiovascular disease,[41,42] and the AHI, which is commonly thought to be the more reliable parameter in sleep medicine, is commonly used to determine OSA severity and by consequence this association. However, great variation exists whenever this metric is used.[43,44]

The variation in AHI scoring and results can be attributed to a number of factors such as intra-individual variability of a single-night recording, as well as scoring variability between PSG technologists and sleep laboratories due to different scoring criteria and/or disparity in interpretation of the rules.[43] In addition, the AHI depends largely on hypopnea scoring, which is not entirely correlated with cardiovascular disease risk. By varying the criteria for defining hypopnea, Punjabi et al[45] showed that hypopneas associated with desaturations are independently associated with cardiovascular disease, but no association was observed between cardiovascular disease and hypopneas associated with arousals. As one of the purposes of surgery in OSA should be the improvement of cardiovascular disease risk, the oxygen desaturation index (ODI) might provide a more consistent metric for appraisal of surgical treatment.[46]

Limitations and Research Perspectives

The limitations of the overview are due to the heterogeneity in the presentation of the results and methodological limitations in the reviews themselves. Most of the reviews were based on case series, and the AMSTAR results reflect a lack of thoroughness in the methodology of the investigations.

Future reviews would benefit from consistency in AHI scoring (eg, hypopnea definition), inclusion of ODI outcomes, standardized definitions of surgical success, and a greater focus on improvement in clinical outcomes. Also, the evidence presented here suggests that further research should be performed using a more robust methodological approach to create systematic reviews and perform primary studies with higher quality.

Conclusion

This overview of reviews suggests that the majority of the literature reviews targeting surgical treatment for obstructive sleep apnea present deficient quality assessment of evidence, with conflicting results not only on metrics, methods, and indications, but also in the definition of a successful OSA surgery. The level and quality of the evidence provided in the reviews remain low, and it is unclear how these reviews reflect OSA surgery literature overall

Acknowledgment

We gratefully acknowledge the support provided by Helder Silva, MD, from the Hospital Sao Sebastiao.

Author Contributions

Victor Certal, drafting, data analysis; **Naoya Nishino**, drafting, data analysis; **Macario Camacho**, drafting, data analysis; **Robson Capasso**, revising, data analysis, final approval.

Disclosures

Competing interests: None.

Sponsorships: None.

Funding source: None.

References

1. Sundaram S, Bridgman SA, Lim J, Lasserson TJ. Surgery for obstructive sleep apnoea. *Cochrane Database Syst Rev*. 2005: CD001004.
2. Marin JM, Carrizo SJ, Vicente E, Agusti AG. Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: an observational study. *Lancet*. 2005;365: 1046-1053.
3. Holty JE, Guilleminault C. Maxillomandibular advancement for the treatment of obstructive sleep apnea: a systematic review and meta-analysis. *Sleep Med Rev*. 2010;14:287-297.
4. Powell NB, Zonato AI, Weaver EM, et al. Radiofrequency treatment of turbinate hypertrophy in subjects using continuous positive airway pressure: a randomized, double-blind, placebo-controlled clinical pilot trial. *Laryngoscope*. 2001;111:1783-1790.
5. Lin HC, Friedman M, Chang HW, Gurpinar B. The efficacy of multiple procedures of the upper airway in adults with obstructive sleep apnea/hypopnea syndrome. *Laryngoscope*. 2008; 118:902-908.
6. Higgins JPT, Green S, Cochrane Collaboration. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester, England; Hoboken, NJ: Wiley-Blackwell; 2008.
7. Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol*. 2009;62:1013-1020.
8. Alsufyani NA, Al-Saleh MA, Major PW. CBCT assessment of upper airway changes and treatment outcomes of obstructive sleep apnoea: a systematic review. *Sleep Breath*. 2013;17: 911-923.
9. Georgalas C, Garas G, Hadjihannas E, Oostra A. Assessment of obstruction level and selection of patients for obstructive sleep apnoea surgery: an evidence-based approach. *J Laryngol Otol*. 2010;124:1-9.
10. Moyer CA, Sonnad SS, Garetz SL, Helman JI, Chervin RD. Quality of life in obstructive sleep apnea: a systematic review of the literature. *Sleep Med*. 2001;2:477-491.
11. Vasu TS, Grewal R, Doghramji K. Obstructive sleep apnea syndrome and perioperative complications: a systematic review of the literature. *J Clin Sleep Med*. 2012;8:199-207.
12. Panossian LA, Veasey SC. Daytime sleepiness in obesity: mechanisms beyond obstructive sleep apnea—a review. *Sleep*. 2012;35:605-615.
13. Rocke D, Sharp S, Wiener D, Puscas L, Lee WT. Effectiveness of a postoperative disposition protocol for sleep apnea surgery. *Am J Otolaryngol*. 2013;34:273-277.
14. Thomasouli MA, Brady EM, Davies MJ, et al. The impact of diet and lifestyle management strategies for obstructive sleep apnoea in adults: a systematic review and meta-analysis of randomised controlled trials. *Sleep Breath*. 2013;17:925-935.
15. Marcus CL, Brooks LJ, Draper KA, et al. Diagnosis and management of childhood obstructive sleep apnea syndrome. *Pediatrics*. 2012;130:e714-755.
16. Nolan J, Brietzke SE. Systematic review of pediatric tonsil size and polysomnogram-measured obstructive sleep apnea severity. *Otolaryngol Head Neck Surg*. 2011;144:844-850.
17. Valrie CR, Bromberg MH, Palermo T, Schanberg LE. A systematic review of sleep in pediatric pain populations. *J Dev Behav Pediatr*. 2013;34:120-128.
18. Khan AM. Adenotonsillectomy outcomes in treatment of obstructive sleep apnea and children: a multicenter retrospective study. *Am J Respir Crit Care Med*. 2011;183:826.
19. Friedman M, Wilson M, Lin HC, Chang HW. Updated systematic review of tonsillectomy and adenoidectomy for treatment of pediatric obstructive sleep apnea/hypopnea syndrome. *Otolaryngol Head Neck Surg*. 2009;140:800-808.
20. Aurora RN, Casey KR, Kristo D, et al. Practice parameters for the surgical modifications of the upper airway for obstructive sleep apnea in adults. *Sleep*. 2010;33:1408-1413.
21. Li KK. Surgical therapy for adult obstructive sleep apnea. *Sleep Med Rev*. 2005;9:201-209.
22. Ephros HD, Madani M, Yalamanchili SC. Surgical treatment of snoring & obstructive sleep apnoea. *Indian J Med Res*. 2010;131:267-276.
23. Kao YH, Shnayder Y, Lee KC. The efficacy of anatomically based multiple procedures for obstructive sleep apnea. *Otolaryngol Head Neck Surg*. 2003;129:327-335.
24. Holty JE, Guilleminault C. Surgical options for the treatment of obstructive sleep apnea. *Med Clin North Am*. 2010;94: 479-515.

25. Blumen M, Crampette L, Fischler M, et al. Surgical treatment of obstructive sleep apnea syndrome. *Rev Mal Respir*. 2010; 27(suppl 3):S157-165.
26. Sher AE, Schechtman KB, Piccirillo JF. The efficacy of surgical modifications of the upper airway in adults with obstructive sleep apnea syndrome. *Sleep*. 1996;19:156-177.
27. Caples SM, Rowley JA, Prinsell JR, et al. Surgical modifications of the upper airway for obstructive sleep apnea in adults: a systematic review and meta-analysis. *Sleep*. 2010;33:1396-1407.
28. Choi JH, Kim SN, Cho JH. Efficacy of the Pillar implant in the treatment of snoring and mild-to-moderate obstructive sleep apnea: a meta-analysis. *Laryngoscope*. 2013;123:269-276.
29. Elshaug AG, Moss JR, Southcott AM, Hiller JE. Redefining success in airway surgery for obstructive sleep apnea: a meta analysis and synthesis of the evidence. *Sleep*. 2007;30:461-467.
30. Franklin KA, Anttila H, Axelsson S, et al. Effects and side-effects of surgery for snoring and obstructive sleep apnea—a systematic review. *Sleep*. 2009;32:27-36.
31. Kezirian EJ, Goldberg AN. Hypopharyngeal surgery in obstructive sleep apnea: an evidence-based medicine review. *Arch Otolaryngol Head Neck Surg*. 2006;132:206-213.
32. Li HY, Wang PC, Chen YP, Lee LA, Fang TJ, Lin HC. Critical appraisal and meta-analysis of nasal surgery for obstructive sleep apnea. *Am J Rhinol*. 2011;25:45-49.
33. Pirklbauer K, Russmueller G, Stiebellehner L, et al. Maxillomandibular advancement for treatment of obstructive sleep apnea syndrome: a systematic review. *J Oral Maxillofac Surg*. 2011;69:e165-176.
34. Fu WK. Randomized controlled trials in surgery: an essential component of scientific progress (*Spine* 1999; 24:2553-5). *Spine*. 2000;25:2002-2004.
35. Concato J. Observational versus experimental studies: what's the evidence for a hierarchy? *NeuroRx*. 2004;1:341-347.
36. Solomon MJ, McLeod RS. Should we be performing more randomized controlled trials evaluating surgical operations? *Surgery*. 1995;118:459-467.
37. Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ*. 1994;309:1286-1291.
38. Shen T, Shimahara E, Cheng J, Capasso R. Sleep medicine clinical and surgical training during otolaryngology residency: a national survey of otolaryngology residency programs. *Otolaryngol Head Neck Surg*. 2011;145:1043-1048.
39. Kezirian EJ, Hussey HM, Brietzke SE, et al. Hypopharyngeal surgery in obstructive sleep apnea: practice patterns, perceptions, and attitudes. *Otolaryngol Head Neck Surg*. 2012;147: 964-971.
40. Poirier J, George C, Rotenberg B. The effect of nasal surgery on nasal continuous positive airway pressure compliance [published online April 10, 2013]. *Laryngoscope*. doi:10.1002/lary. 24131.
41. Bixler EO, Vgontzas AN, Lin HM, et al. Association of hypertension and sleep-disordered breathing. *Arch Intern Med*. 2000;160:2289-2295.
42. Shahar E, Whitney CW, Redline S, et al. Sleep-disordered breathing and cardiovascular disease: cross-sectional results of the Sleep Heart Health Study. *Am J Respir Crit Care Med*. 2001;163:19-25.
43. Aarab G, Lobbezoo F, Hamburger HL, Naeije M. Variability in the apnea-hypopnea index and its consequences for diagnosis and therapy evaluation. *Respiration*. 2009;77:32-37.
44. Newell J, Mairesse O, Verbanck P, Neu D. Is a one-night stay in the lab really enough to conclude? First-night effect and night-to-night variability in polysomnographic recordings among different clinical population samples. *Psychiatry Res*. 2012;200:795-801.
45. Punjabi NM, Newman AB, Young TB, Resnick HE, Sanders MH. Sleep-disordered breathing and cardiovascular disease: an outcome-based definition of hypopneas. *Am J Respir Crit Care Med*. 2008;177:1150-1155.
46. Oeverland B, Skatvedt O, Kvaerner KJ, Akre H. Pulseoximetry: sufficient to diagnose severe sleep apnea. *Sleep Med*. 2002;3:133-138.