Premiere Publications from
The Triological Society

Read all three of our prestigious publications, each offering high-quality content to keep you informed with the latest developments in the field.

**The Laryngoscope**

*Founded in 1896*

Editor-in-Chief: Michael G. Stewart, MD, MPH

The leading source for information in head and neck disorders.

[Laryngoscope.com](https://laryngoscope.com)

**Laryngoscope: Investigative Otolaryngology**

*Open Access*

Editor-in-Chief: D. Bradley Welling, MD, PhD, FACS

Rapid dissemination of the science and practice of otolaryngology-head and neck surgery.

[InvestigativeOto.com](https://investigativeoto.com)

**ENTtoday**

A publication of the Triological Society

Editor-in-Chief: Alexander Chiu, MD

Must-have timely information that Otolaryngologist-head and neck surgeons can use in daily practice.

[Enttoday.org](https://enttoday.org)
Multilevel Palate and Tongue Base Surgical Treatment of Obstructive Sleep Apnea: A Systematic Review and Meta-analysis

Graeme B. Mulholland, MD, FRCSC; Caroline C. Jeffery, MD, FRCSC; Hedyeh Ziai, MD; Varinder Hans, BSc; Hadi Seikaly, MD, FRCSC; Kenny P. Pang, FRCSEd, FRCSI(OTO); Brian W. Rotenberg, MD, FRCSC

Objective: To evaluate multilevel palate and tongue base surgery as a method of treatment of obstructive sleep apnea by comparing the pre- and postoperative apnea-hypopnea index.

Methods: We conducted a systematic review. MEDLINE and Embase databases were searched in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines for conducting systematic reviews. Two authors screened all articles and performed methodological quality assessment. Relevant articles where reviewed in detail. Standard inclusion criteria were applied for article selection. Relevant data were extracted and summarized, a difference of means random-effects model was performed. Our primary outcome measure was change in apnea-hypopnea index pre-/postsurgical treatment.

Results: Of 1,172 studies identified from January 2006 to March 2017, 46 studies met inclusion criteria and were included in the systematic review. This included 11 surgical subgroups and 1,806 patients. Methodological quality and risk of bias assessments were completed. There was strong male predominance 86.8 (standard deviation [SD] = 10.3%), and the average age was 46.8 (SD = 4.0) years. All studies included overweight to obese patients (average body mass index = 29.1 [SD = 3.5]). The average preoperative apnea-hypopnea index was 39.0 (SD = 15.4), and the average postoperative apnea-hypopnea index decreased to 18.3 (SD = 7.5). Meta-analysis data yielded a decrease in apnea-hypopnea index of −23.67 with a 95% confidence interval of −27.27 to −20.06.

Conclusions: Non-maxillomandibular advancement, multilevel surgical procedures for obstructive sleep apnea demonstrate significant improvements in reduction of apnea-hypopnea index following surgery in addition to improvement in many other sleep-specific outcomes. Future research should include larger, higher-level studies that compare surgical treatments and identify factors associated with outcomes.

Key Words: Obstructive sleep apnea, surgical treatment of obstructive sleep apnea, sleep medicine, sleep apnea, quality of life.

Level of Evidence: NA

INTRODUCTION
Obstructive sleep apnea (OSA) is a common condition affecting between 3% and 7% of the world's population.1 It is characterized by partial or complete collapse of the upper airway during sleep.2 OSA has been connected with numerous serious health consequences including neurocognitive impairment, hypertension, stroke, motor vehicle accidents, abdominal aortic aneurysm enlargement, and insulin resistance.1,3–5 Continuous positive airway pressure (CPAP) is currently the first-line treatment for OSA. However, CPAP compliance is incomplete and ranges between 30% and 80%, calling into question the appropriateness of considering CPAP as a gold standard of therapy.6

Surgery has evolved over time from initially that of single-stage, single-level procedures to that of the current concept of multilevel surgery to address all areas of obstruction observed in OSA. Numerous procedures have been developed and applied with varying levels of success. Not every patient presenting with OSA is appropriate for soft-tissue surgery, but surgery does have a relevant role in well-selected patients with the right anatomically modifiable segments.7 The only other commonly employed procedure is that of maxillomandibular advancement (MMA). MMA does show the greatest consistency in improvement of polysomnography (PSG) parameters achieving surgical...
success according to the Sher criteria (postoperative apnea-hypopnea index [AHI] <20 and >50% reduction of preoperative AHI) in just over 85% of patients. However, MMA is a far more invasive means of treatment that many patients are hesitant to undergo, and many other reported outcome measures do not improve as robustly as AHI.

Few published systematic reviews regarding multilevel surgery for OSA exist. Three such reviews were identified. The first by Caples et al. looked at all types of surgery for treatment of OSA. No meta-analysis was performed, as the information gathered was deemed too heterogenous. A review by Kezirian and Goldberg in 2006 examined hypopharyngeal surgery. It provided a thorough and evidence-based presentation of the literature up to the time it was published. This review did not require that both the retropalatal and hypopharyngeal airway be addressed in a single stage. Its focus was on the hypopharyngeal airway, and therefore, studies utilizing staged procedures were included. The final relevant review by Lin et al. was published in 2008 and examined multilevel surgery for treatment of OSA. The inclusion criteria were not specific to AHI, but rather AHI and the Respiratory Disturbance Index, adding an additional degree of clinical diversity in terms of determining true treatment effect. Given the major shift in surgical treatment paradigm that occurred with the acceptance of the concept of lateral wall collapse in the mid-2000s, we chose to include only those studies published after 2005. Our study was aimed at providing a current up-to-date summary of modern multilevel surgery for OSA, while performing a meta-analysis on the current gold standard marker of OSA treatment, the AHI.

The relative paucity of published systematic reviews places emphasis squarely on the need for thorough evaluation of the success of soft-tissue surgery. Much of what exists is historical as opposed to modern, and is also confounded by including MMA in the analyzed procedure lists. The objective of this study was, therefore, to perform an evidence-based systematic review and meta-analysis of multilevel palate and tongue base surgery outcomes, focusing on the change in AHI pre- and postoperatively for treatment of OSA, but while excluding MMA.

METHODS

Search Strategy

We utilized the Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist and recommendations for this systematic review. The search was performed on March 30, 2017. Databases searched included PubMed/MEDLINE and Embase. The search strategy used for each database utilized is outlined as follows:

1. (Uvulopalatopharyngoplasty OR Tracheostomy Tracheotomy OR Mandibular advancement OR Genioglossus advancement OR Maxillomandibular advancement OR Laser assisted uvuloplasty OR Midline glossectomy OR Aryepiglottic fold reduction OR Epiglottectomy OR Nasal surgery OR Cautery-assisted uvulopalatoplasty OR Cautery-assisted palatal stiffening operation OR Epiglottopexy Epiglottoplasty OR Radiofrequency ablation of the soft palate OR Rapid maxillary expansion Surgically-assisted rapid palatal expansion OR Radiofrequency ablation of the tongue OR Hypoglossal nerve stimulator implant OR Mandibular distraction OR Palatal implant OR Z-palatoplasty OR Tongue base coblation OR Submucosal minimally invasive lingual excision OR Expansion sphincter pharyngoplasty OR Transoral Robotic Surgery OR Lateral pharyngoplasty OR Tonsillectomy OR Tongue stabilization Tongue Suspension OR Transpalatal Advancement Pharyngoplasty).af.
2. multilevel.af.
3. 1 OR 2.
4. apnea-hypopnea index.af.
5. obstructive sleep apnea.af.
6. 3 and 5.
7. 4 and 6.

Two authors (G.B.M. and V.H.) independently reviewed the selected studies with the following PICOS (patients, intervention, comparison, outcomes, and study design) format.

Inclusion Criteria

Inclusion criteria were 1) Studies published from January 2006 to March 2017, this date range being chosen so as to minimize bias from historical data. 2) Patients: adult patients (>18 years old) with OSA. 3) Intervention: multilevel surgery (surgery addressing at least both the retropalatal and hypopharyngeal levels of obstruction). 4) Comparison: presurgical AHI data. 5) Outcome: mean difference of pre- and postsurgical AHI. 6) Study design: case series, case control, cohort and randomized trials. 7) English language or translation available.

Exclusion Criteria

Exclusion criteria were 1) Studies not addressing retropalatal and hypopharyngeal obstruction surgically during same procedure (single stage). 2) Studies including MMA, wherein the MMA data could not be extracted, no individual patient data, and/or no data for procedure of interest. 3) Studies with no quantitative data. 4) Patients with hypventilation syndrome or central sleep apnea. 5) Studies containing patients whose body mass index (BMI) changed by >10% between polysomnograms.

STUDY EXTRACTION, CATEGORIZATION, AND ANALYSIS

We analyzed the articles that met criteria for variables, including:

1. Preoperative characteristics of patient populations, such as BMI (calculated as weight in kilograms divided by the square of height in meters) and preoperative sleep study results such as the reported AHI and lowest oxygen saturation (LSAT) during sleep (as per definitions by the American Academy of Sleep...
Medicine Manual for the Scoring of Sleep and Associated Events.14

2. Postoperative sleep study results, such as AHI and LSAT.

3. Intervention performed, operative data, complications, and surgical success defined by the Sher criteria (i.e., a reduction in AHI ≥50% and an AHI of <20), and surgical cure was defined as postoperative AHI of less than five events per hour.8

4. Any factors (e.g., preoperative BMI or AHI) that were associated with outcomes.

5. Improvement in excessive daytime somnolence (EDS), measured by the Epworth Sleepiness Scale (ESS).

6. Pooling of collected parameters (age, BMI, pre-/postoperative AHI, surgical success, pre-/postoperative LSAT, and pre-/postoperative ESS) was performed by weighted averages:

\[
\frac{\left(\frac{\text{number of patients in study A}}{\text{number of patients in study A}}\times\text{mean of collected parameter (i.e., BMI, AHI, surgical success etc) of study A}\right) + \left(\frac{\text{number of patients in study B}}{\text{number of patients in study B}}\times\text{mean of collected parameter (i.e., BMI, AHI, surgical success etc) of study B}\right) + \ldots + \left(\frac{\text{number of patients in study Z}}{\text{number of patients in study Z}}\times\text{mean of collected parameter (i.e., BMI, AHI, surgical success etc) of study Z}\right) + \ldots + \left(\text{number of patients in study Z}\right)}{\text{number of patients in study A} + \text{number of patients in study B} + \ldots + \text{number of patients in study Z}}
\]

7. Methodologic quality assessment for indvidual studies was performed by means of the Mixed Methods Appraisal Tool.15

8. Risk of bias assessment for individual studies was performed by means of the methodological index for non-randomized studies (MINORS) appraisal tool.16

9. The meta-analysis utilized the difference in means by way of a random-effects model as the principle summary measure. Review Manager (RevMan) 5.3 by the Cochrane Collaboration (London, United Kingdom) was utilized to perform the meta-analysis.

10. The \(I^2\) statistic was used to examine heterogeneity across studies for the random-effects model. Levels of inconsistency guidelines were used to assign levels: low inconsistency >25%, moderate inconsistency >50%, and high inconsistency >75%.17

11. Risk of publication bias across studies was evaluated by funnel plots for individual surgical subgroups and all studies combined together.

Authors G.B.M. and V.H. determined by consensus the studies to include in this analysis.

RESULTS

The flowchart of studies identified, screened, and selected for review are summarized in Figure 1. A total of 46 studies were included for analysis and included 1,806 patients. Overall, the data published in these studies
were heterogeneous. Utilizing the Mixed Methods Assessment Tool, the majority of the studies (30 of 46) were quantitative descriptive studies (case series) (see Supporting Table 1 in the online version of this article).

In addition to this, a risk of bias assessment was performed via the MINORS assessment tool (see Supporting Table 2 in the online version of this article). The patient population had a strong male predominance (standard deviation [SD] = 86.8, 10.3%), and the average age was 46.8 years (SD = 4.0 years). All studies included overweight patients (average BMI = 29.1, SD = 3.5), and preoperative AHI on average was severe (average = 39.0, SD = 15.4). The overall surgical success weighted by study was 58.5%. A subgroup summary displaying weighted average values for variables collected including average length of follow-up is presented in Table I.

We identified 11 different interventions to address hypopharyngeal obstruction: trans-oral robotic surgery for the base of tongue (TORS BOT), tongue base suspension (TBS), genioglossal advancement (GA), genioglossus advancement and hyoid suspension (GAHM), hyoid suspension (HS), radiofrequency tongue base reduction (RFTBR), submucosal lingualplasty (SMLR, SMILE or SML), lingual tonsillectomy (LT), coblator tongue base resection and coblation endoscopic lingual lightening (CTBR and CELL), midline glossectomy (MG) and hyoid suspension and radiofrequency tongue base reduction (HS & RFTBR).

Supplemental Table 3 contains pertinent data extracted from individual studies organized by surgical subgroups.

For retropalatal obstruction we identified six different modalities of treatment: uvulopalatopharyngoplasty (UPPP), lateral pharyngoplasty (LPP), Z-palatopharyngoplasty (ZPPP), expansion sphincteroplasty (ESP), relocation pharyngoplasty (RP), and temperature controlled radiofrequency of soft palate (TCRF).

Seven studies examining 216 patients were identified that utilized TORS BOT to address the hypopharynx. These studies applied varying soft-palate techniques. On average, sleep apnea was classified as severe preoperatively. Four of five studies reporting ESS showed statistically significant improvement. From the random-effects model comparing difference in mean preoperative to postoperative AHI, a decrease in AHI favoring surgical treatment was found with a value of $-28.01$, with a 95% confidence interval (CI) of $-33.83$ to $-22.19$ and an $I^2$ of 73% (Fig. 2A).

Three TORS BOT studies utilized soft-palate techniques other than UPPP. These studies employed ZPPP, ESP, and LPP. The success rates for these studies were: 67%, 75%, and 63%, respectively. In the study by Friedman et al., where ZPPP and TORS BOT were performed, there was significant improvement in LSAT from 78.5 (SD = 7.4) to 86.5 (SD = 6.3) ($P < .01$) and daytime somnolence by ESS measures of 14.4 (SD = 4.5) to 5.4 (SD = 3.1) ($P < .001$). The mean volume of lingual tissue excised was 2.28 g (SD = 0.43 g). Another study by Vicini et al. compared ESP to UPPP with TORS BOT surgery. Each group contained 12 patients. When comparing results from the UPPP group to the ESP group, the AHI changed from 38.4 (SD = 19.7) to 19.8 (SD = 3.4)...

### Table I.

Weighted Average Summary of Combined Study Characteristics.

<table>
<thead>
<tr>
<th>Technique</th>
<th>No. of Patients</th>
<th>No. of Studies</th>
<th>Sex (% M)</th>
<th>Age (Year)</th>
<th>BMI (kg/m²)</th>
<th>AHI</th>
<th>LSAT</th>
<th>ESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP and CTBR</td>
<td>133</td>
<td>4</td>
<td>87.3</td>
<td>45.1</td>
<td>28.3</td>
<td>39.5</td>
<td>18.1</td>
<td>73.2</td>
</tr>
<tr>
<td>SP and MG</td>
<td>34</td>
<td>1</td>
<td>100</td>
<td>36.7</td>
<td>30.9</td>
<td>56.0</td>
<td>19.0</td>
<td>62.1</td>
</tr>
<tr>
<td>SP and SMLR</td>
<td>149</td>
<td>4</td>
<td>346</td>
<td>1</td>
<td>80.5</td>
<td>48.0</td>
<td>29.1</td>
<td>51.3</td>
</tr>
<tr>
<td>SP and TBS</td>
<td>34</td>
<td>2</td>
<td>119</td>
<td>1</td>
<td>86.4</td>
<td>47.9</td>
<td>30.1</td>
<td>51.1</td>
</tr>
<tr>
<td>SP and GA</td>
<td>49</td>
<td>3</td>
<td>118</td>
<td>1</td>
<td>82.4</td>
<td>47.0</td>
<td>30.4</td>
<td>51.1</td>
</tr>
<tr>
<td>SP and GAHM</td>
<td>216</td>
<td>5</td>
<td>118</td>
<td>1</td>
<td>86.7</td>
<td>47.9</td>
<td>30.2</td>
<td>51.1</td>
</tr>
<tr>
<td>SP and HG</td>
<td>164</td>
<td>6</td>
<td>1</td>
<td>84.5</td>
<td>45.5</td>
<td>29.4</td>
<td>31.1</td>
<td>21.3</td>
</tr>
<tr>
<td>SP and MG</td>
<td>95</td>
<td>2</td>
<td>92.0</td>
<td>1</td>
<td>90.0</td>
<td>45.0</td>
<td>28.6</td>
<td>36.1</td>
</tr>
<tr>
<td>SP and HS</td>
<td>472</td>
<td>8</td>
<td>1</td>
<td>87.7</td>
<td>49.4</td>
<td>27.7</td>
<td>31.1</td>
<td>21.3</td>
</tr>
<tr>
<td>SP and RFTBR</td>
<td>30</td>
<td>1</td>
<td>84.8</td>
<td>46.8</td>
<td>29.1</td>
<td>39.0</td>
<td>18.3</td>
<td>76.5</td>
</tr>
<tr>
<td>All</td>
<td>1,806</td>
<td>46</td>
<td>86.8</td>
<td>46.8</td>
<td>28.3</td>
<td>39.0</td>
<td>18.3</td>
<td>76.5</td>
</tr>
</tbody>
</table>

Values from pooled data are presented as weighted averages. Blank fields indicate no reported values.

% M = percent male; AHI = apnea-hypopnea index; BMI = body mass index (kg/m²); CTBR = coblator tongue base resection; ESS = Epworth Sleepiness Scale; GA = genioglossal advancement; GAHM = genioglossus advancement and hyoid suspension; HS = hyoid suspension; HG = glossectomy; HGHM = genioglossus advancement and hyoid suspension and tongue base resection; HS & RFTBR = hyoid suspension and radiofrequency tongue base reduction; CTBR and CELL = coblator tongue base resection and coblation endoscopic lingual lightening; TCRF = temperature controlled radiofrequency of soft palate; TORS BOT = transoral robotic surgery for the base of tongue.
(P = .023) and 38.5 (SD = 14.4) to 9.9 (SD = 8.6) (P = .002), respectively. Comparing the changes in excessive daytime somnolence pre- and postoperatively between the UPPP and the ESP group are as follows. The ESS for the UPPP group started as 13.8 (SD = 4.0) and changed to 7.6 (SD = 4.4) postoperatively, whereas for the ESP group, ESS preoperatively was 12.0 (SD = 4.9) and postoperatively was 4.4 (SD = 4.1), no information on statistical significance for these changes was reported. Overall surgical success of the UPPP and the ESP group were 33.3% and 75%, respectively.19 The third study made use of LPP, and identified patients with tonsils grade II or larger had a greater chance of successful surgery (P = .002).20

Coblation was used for open resection of the tongue base and lingual tonsils in four reports. Soft-palate techniques included UPPP in two series, and RP and ZPPP each in one study. Sleep apnea was severe in all studies, and statistically significant improvement in EDS was seen in all groups. Surgical success results were not reported for one study.21 In the random-effects model comparing difference in mean preoperative to postoperative AHI, a decrease in AHI favoring surgical treatment was found, with a value of −21.30, a 95% CI of −29.75 to −12.85, and an I² of 76% (Fig. 2B).

One study reported results of UPPP combined with coblation lingual tonsillectomy. The authors identified two groups: patents with palatal tonsils removed concurrently and patients who had previous palatal tonsillectomy. They found concurrent removal of tonsils during the same procedure significantly increased the likelihood of success of the procedure compared to patients who had previous tonsillectomy. The pre- and postoperative AHI changed from 37.2 (SD = 17.8) to 12.8 (SD = 12.5), and the overall success rate was 84.4% in the concurrent palatal tonsillectomy group. In patients who previously underwent palatine tonsillectomy, the AHI changed from 30.9 (SD = 21.1) to 23.1 (SD = 19.7), with overall success of 30.8%. BMI was not reported for these subgroups.22
A single study examined midline glossectomy in combination with UPPP and included 58 patients, of which 34 patients were followed for 5 years. Regarding surgical success, cure of OSA (defined by AHI <5) occurred in 27 of 34 patients and success in an additional six of 34 patients. At 5 years there was some relapse of OSA; seven of 34 retained cure, and 17 of 34 maintained surgical success. These findings describe a 97% success rate at 6 months and a 70.6% success rate at 5 years for the group of patients follow by PSG at these time points. \(^{23}\)

For the submucosal lingualplasty group, patients were obese and sleep apnea reported by the AHI was severe. Success ranged from 45.5% to 100%, all achieving statistically significant decreases in AHI. Statistically significant improvement in LSAT varied; however, EDS consistently decreased. The difference in mean preoperative to postoperative AHI from the random-effects model showed a decrease in AHI favoring surgical treatment, with a value of −26.79 with a 95% CI of −37.89 to −15.69 and an \(I^2\) of 95% (Fig. 2C).

Sorrenti et al. published a study with 100% surgical success. They operated on 10 patients from January 2001 to March 2005. Preoperative and postoperative values included: AHI = 54.7 (SD = 11.5) to 9.4 (SD = 5.4), LSAT = 77% (SD = 6.2%) to 90.7% (SD = 3%), and ESS = 14.3 (range, 11–20) to 5.3 (range, 4–9), and followed patients for an average of 8.2 months (range, 6–19 months). \(^{24}\)

Suture tongue base stabilization studies totaled seven, with 346 patients combined. Sleep apnea was classified as severe in five of seven studies and moderate in two of seven studies. LSAT and ESS where collected in three of seven studies for both parameters, showing statistically significant improvement where reported. In the random-effects model comparing the difference in mean preoperative to postoperative AHI, a decrease in AHI favoring surgical treatment was found with a value of...
Study or Subgroup & Treatment & Mean & SD & Total & Control & Mean & SD & Total & Weight & Mean Difference & IV, Random, 95% CI & Mean Difference & IV, Random, 95% CI
--- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---
Eun 2009a & 20.5 & 20.9 & 62 & 23.1 & 5.6 & 62 & 12.4% & -2.60 [-7.99, 2.78] & & & 
Fernandez 2015 & 15.4 & 7.6 & 29 & 32.5 & 9 & 29 & 12.9% & -17.10 [-21.12, -13.08] & & & 
Friedman 2012 & 5 & 22.5 & 24 & 54.7 & 26.6 & 24 & 8.4% & -49.70 [-63.64, -35.76] & & & 
Lin 2010 & 23.4 & 24.7 & 43 & 51.5 & 25.4 & 43 & 10.0% & -28.10 [-38.69, -17.51] & & & 
Wang 2013 & 24 & 24 & 36 & 58.9 & 20.5 & 36 & 10.1% & -35.80 [-46.11, -25.49] & & & 
Total (95% CI) & 382 & 382 & 100.0% & -18.09 [-24.68, -11.50] & & & & & & &

Heterogeneity: Tau^2 = 83.62, Chi^2 = 90.51, df = 6 (P < 0.00001); I^2 = 91%
Test for overall effect: Z = 5.38 (P < 0.00001)

Fig. 4. Difference of means (change in pre-/postoperative apnea-hypopnea index) random-effects model meta-analysis for soft palate and RFTBR. Mean difference: the difference in mean apnea-hypopnea index from preoperatively to postoperatively. CI = confidence interval; IV = independent variable; RFTBR = radiofrequency tongue base reduction; SD = standard deviation [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

−26.97 with a 95% CI of −33.52 to −20.42 and an I^2 of 88% (Fig. 3A). The study by Tuncel et al. was not included in meta-analysis due to missing data.

One study obtained polysomnography at the 3-year follow-up on all 54 patients included. The preoperative and postoperative values were as follows: AHI = 52.8 (SD = 14.9) to 14.1 (SD = 23.5) (P < .001), LSAT = 76.2% (SD = 12.4%) to 82.2% (SD = 11.2%) (P < .05), and ESS = 12.2 (SD = 3.3) to 8.2 (SD = 6.1) (P = .002). The authors performed a logistic regression analysis identifying the odds ratio for preoperative BMI as the only statistically significant variable affecting surgical success (P < .01). From this, a BMI <35 was calculated to be a positive predictor of success (P = .02).25

The GA and UPPP subgroup totaled three studies and included 49 patients. Of these studies, one by Emara et al. described a modified GA technique where in which they utilized intraoperative x-ray to plan their axial osteotomies 5 mm below the roots of the incisors and 5 mm above the inferior border of the mandible. Providing and average advancement of 11.8 mm (SD = 2.6 mm) and allowing capture of the entire anterior insertion of the genioglossus on the mandible. The reported preoperative data were AHI = 40.7 (SD = 17.4), LSAT = 78.9% (SD = 12.6%), and ESS = 14.2 (SD = 2.3), and with postoperative data were AHI = 15.4 (SD = 10.7) (P < .001), LSAT = 87.2 (SD = 11.1) (P < .01), and ESS = 8.3 (SD = 3.9) (P < .01).26 In the random-effects model (Fig. 3B) comparing the difference in mean preoperative to postoperative AHI, a decrease in AHI favoring surgical treatment was found with a value of −25.86 with a 95% CI of −32.98 to −18.73 and an I^2 of 0%.

All genioglossus advancement and hyoid suspension studies were retrospective case series with the exception of the prospective cohort study conducted by Kezirian et al.27 Patients were obese and had severe sleep apnea, with the exception of one study demonstrating moderate sleep apnea based on AHI values. The majority of the studies showed significant decreases in daytime somnolence. ZPPP was employed in place of UPPP in one of the seven studies. The difference in the mean random-effects model of preoperative to postoperative AHI showed a decrease in AHI favoring surgical treatment with a value of −28.00 with a 95% CI of −41.75 to −14.25 and an I^2 of 82% (Fig. 3C). The study by Yin et al. was excluded from the meta-analysis due to missing data.28

In the hyoid suspension group, patients included ranged from overweight to morbidly obese. Severity of OSA ranged from moderate to severe. A large retrospective case series by Tschopp et al. that included 144 patients performed multiple linear and logistic regression over the 12-month follow-up period. They found that preoperative BMI and tonsillectomy had a highly significant positive influence on postoperative AHI. Patients having a tonsillectomy as part of the procedure had a 58% (P < .001) chance of success versus a 19% chance of success without a tonsillectomy.29 The meta-analysis could not be performed for this subgroup because it is represented by a single study.

Two studies (retrospective case controlled trial and prospective cohort study) also by Tschopp et al. examined UPPP with HS and RFTBR. OSA was severe in both studies. The first study’s pre-/postoperative values included AHI = 38.9 (SD = 20) to 20.7 (SD = 20.6) (P < .00001), LSAT = 81.0% (SD = 8.8%) to 84.3% (SD = 8.3%) (P < .05), and ESS = 9.4 (SD = 5.3) to 7.2 (SD = 4.4) (P < .05).29 The second study included 50 patients with pre/postoperative values of AHI = 32.9 (SD = 14) to 17 (SD = 13.5) and ESS = 11.8 (SD = 4.5) to 7.6 (SD = 4.4). Verse et al. described two subgroups: patients having a tonsillectomy with procedure and those having performed remotely. Success for these groups was 71.9% and 22.2%, respectively.30

The RFTBR and soft-palate surgery (UPPP or ZPPP) group included a total of 472 patients. Preoperative AHI ranged from moderate to severe, and surgical success ranged from 35.5% to 66.7%. For the random-effects model comparing the difference in mean preoperative to postoperative AHI, a decrease in AHI favoring surgical treatment was found with a value of −16.24 with a 95% CI of −24.62 to −7.85 and an I^2 of 95% (Fig. 4).

The only randomized controlled trial included in our analysis examined temperature-controlled radiofrequency treatment of the soft palate and tongue base and compared this to sham TCRF control and CPAP. Woodson et al. included patients with mild to moderate OSA.31 The
primary outcome measures were 1/slowest reaction time (1/ms), Functional Outcomes of Sleep Questionnaire responses, and Symptoms of Nocturnal Obstruction and Related Event Questionnaire responses. The authors report pre-/postoperative parameters as \( \text{AHI} = 21.3 \) (SD = 11.1) to 16.8 (SD = 13.8), \( \text{LSAT} = 86.3\% \) (SD = 7.6\%) to 85.7\% (SD = 4.6\%), both not statistically significant, and \( \text{ESS} = 11.9 \) (SD = 4.6) to 9.8 (SD = 3.9) (\( P = .005 \)).

All studies with sufficient data (33 of 46) from all surgical subgroups were combined into a random-effects model. The overall difference in AHI favors treatment with a value of \(-23.67\) and 95\% CI of \(-27.27\) to \(-20.06\). However, the studies compared were very heterogeneous having an \( I^2 \) of 93\%.

A funnel plot examining overall risk of publication bias of all studies together is displayed in Supporting Figure 1 in the online version of this article.

**DISCUSSION**

CPAP has long been considered the standard in treating OSA, but recent data suggest that CPAP compliance is so poor that most patients no longer use the therapy after 1 year postprescription. Hence, surgery for OSA has gained a larger footprint on the management spectrum, with a growing body of data to support use. We performed a systematic review and meta-analysis comprising 46 individual studies for single-stage multilevel palate and tongue base surgery for OSA. Variables such as BMI, pre-/postoperative AHI, and surgical success were pooled by weighted averages creating a detailed summary of the data. The 11 surgical subgroups included 1,802 patients. Our meta-analysis data comparing the difference in means of pre- to postoperative AHI demonstrated an effect favoring surgery, with a decrease in AHI across studies by \(-23.67\) (95\% CI: \(-27.27\) to \(-20.06\)). The overall mean preoperative AHI was 39.0 (±15.4), and average postoperative AHI decreased to 18.3 (±7.5). By a qualitative measure applying the Sher criteria to the overall mean difference of preoperative AHI (39.0) and postoperative AHI (15.4), generally speaking, surgical treatment as a whole could be considered successful as it reduces the AHI by 50\% and to a final value of \(<20\).

Overall, multilevel treatment of OSA with surgery shows a significant improvement in AHI. The largest overall improvement in single subgroup from the meta-analysis was seen in the TORS BOT subgroup, with a decrease of the AHI of \(-28.01\) (95\% CI: \(-33.83\) to \(-22.19\)). There is a moderate degree of heterogeneity among the studies in this group (\( I^2 = 73\% \)), suggesting differences between how each individual study has been reported and differences in surgical technique between studies. All surgical subgroups, where data were available to run the meta-analyses, showed a net decrease in AHI of \(\geq 20\), with the exception studies utilizing radio-frequency ablation of the tongue base. The soft-palate procedure (SP), HS and RFTBR, and SP and RFTBR groups showed decreases in AHI of \(-16.57\) (95\% CI: \(-21.11\) to \(-12.04\)) and \(-18.09\) (95\% CI: \(-24.68\) to \(-11.50\)), respectively.

The results of the meta-analysis may hint at superior and inferior surgical treatment modalities based on an overall significant decrease in AHI. However, these predictions are cautious at best, as direct comparisons between different surgical subgroups were not possible due to the substantial heterogeneity. When looking at heterogeneity across the 11 surgical subgroups, the results of \( I^2 \) are highly inconsistent (\( I^2 > 75\% \)) with the exception of the TORS BOT, the GA, and the RFTBR groups (73\%, 0\%, and 0\%, respectively). This means that the results from each study, even when the same or similar surgical techniques were employed, yielded results that were different beyond the probability of this occurring by chance alone.

The Mixed Methods Assessment Tool was used analyze methodological quality in individual studies. The one randomized controlled trial included met all four methodological quality criteria, making it the strongest study that we looked at. Unfortunately, there were no other randomized controlled trials in the literature to include. The second strongest group of studies, the non-randomized trial group, which included cohort and case-control studies, included 15 studies. The majority of these studies had one major deficiency (9 of 15). These deficiencies include patients in comparison groups that were different from one another (4 of 9), they did not have complete patient data (4 of 9), and one study had bias in patient selection. These deficiencies help explain why such significant heterogeneity was observed in the meta-analysis. The final grouping described as quantitative descriptive studies included case series. Case series made up close to two-thirds of the studies included (30 of 46); this form of uncontrolled, nonrandomized data, without exception, leads to significant heterogeneity. The methodological quality was good overall for these quantitative descriptive studies, with 25 of 30 studies meeting all four criteria of methodological quality assessment. The five studies that had deficiencies were in the area of incomplete patient data.

The MINORS appraisal tool was used to perform risk of bias assessment. Risk of bias assessment was performed for case series, cohort, and case-control trials. The scoring scheme for noncomparative studies was out of a total of 16 and 24 for comparative studies with a comparison group. Overall the results were heterogeneous. The results for the noncomparative studies ranged between 6 and 12 out of 16 and 14, to 20 out of 24 for the comparative studies. The greatest areas in deficiency indicating a high risk of bias were in the domains of unblinded assessment of study outcome and prospective calculation of study size. The outcome of individual studies varied based on the proposed objective set for each study. Surgical success as defined by the Sher criteria was not uniformly used as the primary outcome in most included studies. Overall, the MINORS appraisal tools indicates a high risk of bias for studies included in the meta-analysis.

However, our study had several limitations. First, we chose to focus on single-stage, multilevel procedures addressing the tongue base and palate, even though nasal surgery in combination with either palate or tongue base
surgery is considered to be multilevel surgery and has been shown to improve the likelihood of surgical success. This was because nasal surgery has traditionally been considered an adjunct to OSA surgery; future studies will likely incorporate nasal surgery into the scope of multilevel surgery.32 Furthermore, modalities such as tracheostomy, MMA, and hypoglossal nerve stimulation have been established as successful treatment modalities for OSA in numerous published studies. The goal of our review was to limit our analysis to the benefit of single-stage, multilevel, soft-tissue surgery of the palate and base of tongue, with the exclusion of other treatment modalities. Next, a decision had to be made in regard to organization of surgical subgroups. The tongue base/hypopharyngeal procedures are significantly different from each other, making it necessary to separate these procedure types. The vast majority of multilevel procedures utilized UPPP (46 of 55). Current evidence in the literature suggests similar results when comparing different soft-palate procedures.33 For these reasons, the 11 subgroups were created based on the treatment of the tongue base. Successful treatment as accepted in the literature is based on objective polysomnography outcomes and may differ from what patients deem successful treatment. Patients often strongly value quality of life and improvement in daytime somnolence.34 However, these measures are not consistently reported in all studies. Recently, Pang and Rotenberg published a comprehensive and patient-centered method in evaluating successful treatment of OSA. The SLEEP GOAL criteria synthesizes numerous domains of OSA outcome measures including polysomnographic, morbidity outcomes (such as systolic blood pressure), and quality-of-life measures.35

A funnel plot has been included to examine risk of publication bias across all studies (see Supporting Figure 1 in the online version of this article). The included studies are centered around the overall treatment effect of decrease in AHI by 23.67 with multilevel surgery. There appears to be bias in small studies published that demonstrate smaller treatment effects. This is compared to the majority of larger studies that show greater treatment effects. Overall, there appears to be a publication bias against publishing large studies showing poor treatment effect.

Future research will be necessary to improve the quality of evidence. In particular, high-quality, randomized controlled trials are necessary to address issues of selection bias and improve confidence regarding the true treatment effect size of each modality of treatment. Presently a number of randomized multi-institutional and multi-disciplinary trials are underway.36 An excellent example of the future direction for research is the only randomized controlled trial included in our review. Woodson et al. made a comparison of radiofrequency therapy of the soft palate and base of tongue with a sham placebo treatment and CPAP.31 The study randomized 60 patients to receive TCRF and RFTBR or sham placebo. CPAP was used as the comparison group. The focus of this study was on patient outcomes rather than objective clinical (polysomnographic) outcomes. The surgical treatment did not lower the AHI effectively or substantially, but did show significant improvement in patient daytime somnolence, quality of life, and vigilance. This is supported by the lack of change in these same areas when examining the placebo group.

Furthermore, the other major area of future research should surround finding correlation of patient quality-of-life outcomes with objective PSG outcomes.37 We know there is correlation between the significant morbidity of OSA and PSG Values,38 however, patient-desired outcomes, such as sleep quality and improved daytime somnolence, are often independent of objective PSG data.39 Surgery appears to offer an advantage in achieving patient-desired outcomes when compared to CPAP.31,34 The choice between the benefits of surgical management of OSA versus objective improvement in PSG criteria via CPAP need to be weighed carefully when selecting the most appropriate treatment option for OSA. This is ultimately a case-by-case selection and should be based on thorough clinical evaluation.

CONCLUSION

From our systematic review and meta-analysis we have demonstrated that there is a substantial benefit favoring treatment of OSA with multilevel surgery. Comparison among different surgical subgroups cannot be performed in a meaningful way at the present time due to the significant degree of heterogeneity between current studies available. In addition, multilevel surgery procedures for OSA appear to have improvements in many outcomes, including respiratory physiology during sleep, daytime somnolence, and quality of life. Future research should include larger, higher-level studies that compare surgical treatments and identify factors associated with outcomes.

Acknowledgments

The authors acknowledge statistician Meghan Sebastianski, PhD, and the Alberta SPOR SUPPORT Unit at the University of Alberta for assistance with and funding for the meta-analysis.

BIBLIOGRAPHY
