JAMA | Preliminary Communication

Effect of Multilevel Upper Airway Surgery vs Medical Management on the Apnea-Hypopnea Index and Patient-Reported Daytime Sleepiness Among Patients With Moderate or Severe Obstructive Sleep Apnea
The SAMS Randomized Clinical Trial

Stuart MacKay, MD; A. Simon Carney, DM; Peter G. Catcheside, PhD; Ching Li Chai-Coetzer, MD, PhD; Michael Chia, MD; Peter A. Cistulli, MD, PhD; John-Charles Hodge, MD; Andrew Jones, MD; Billingsley Kaambwa, PhD; Richard Lewis, MD; Eng H. Ooi, MD, PhD; Alison J. Pinczel, PhD; Nigel McArdle, MD; Guy Rees, MD; Bhajan Singh, MD, PhD; Nicholas Stow, MD; Edward M. Weaver, MD, MPH; Richard J. Woodman, PhD; Charmaine M. Woods, PhD; Aeneas Yeo, MD; R. Doug McEvoy, MD

IMPORTANCE Many adults with obstructive sleep apnea (OSA) use device treatments inadequately and remain untreated.

OBJECTIVE To determine whether combined palatal and tongue surgery to enlarge or stabilize the upper airway is an effective treatment for patients with OSA when conventional device treatment failed.

DESIGN, SETTING, AND PARTICIPANTS Multicenter, parallel-group, open-label randomized clinical trial of upper airway surgery vs ongoing medical management. Adults with symptomatic moderate or severe OSA in whom conventional treatments had failed were enrolled between November 2014 and October 2017, with follow-up until August 2018.

INTERVENTIONS Multilevel surgery (modified uvulopalatopharyngoplasty and minimally invasive tongue volume reduction; n = 51) or ongoing medical management (eg, advice on sleep positioning, weight loss; n = 51).

MAIN OUTCOMES AND MEASURES Primary outcome measures were the apnea-hypopnea index (AHI; ie, the number of apnea and hypopnea events/h; 15-30 indicates moderate and >30 indicates severe OSA) and the Epworth Sleepiness Scale (ESS; range, 0-24; >10 indicates pathological sleepiness). Baseline-adjusted differences between groups at 6 months were assessed. Minimal clinically important differences are 15 events per hour for AHI and 2 units for ESS.

RESULTS Among 102 participants who were randomized (mean [SD] age, 44.6 [12.8] years; 18 [18%] women), 91 (89%) completed the trial. The mean AHI was 47.9 at baseline and 20.8 at 6 months for the surgery group and 45.3 at baseline and 34.5 at 6 months for the medical management group (mean baseline-adjusted between-group difference at 6 mo, −17.6 events/h [95% CI, −26.8 to −8.4]; $P < .001$). The mean ESS was 12.4 at baseline and 5.3 at 6 months in the surgery group and 11.1 at baseline and 10.5 at 6 months in the medical management group (mean baseline-adjusted between-group difference at 6 mo, −6.7 [95% CI, −8.2 to −5.2]; $P < .001$). Two participants (4%) in the surgery group had serious adverse events (1 had a myocardial infarction on postoperative day 5 and 1 was hospitalized for observation following hematemesis of old blood).

CONCLUSIONS AND RELEVANCE In this preliminary study of adults with moderate or severe OSA in whom conventional therapy had failed, combined palatal and tongue surgery, compared with medical management, reduced the number of apnea and hypopnea events and patient-reported sleepiness at 6 months. Further research is needed to confirm these findings in additional populations and to understand clinical utility, long-term efficacy, and safety of multilevel upper airway surgery for treatment of patients with OSA.

TRIAL REGISTRATION Australian New Zealand Clinical Trials Registry: ACTRN12614000338662

Editorial page 1161

Author Audio Interview

Supplemental content

CME Quiz at jamacmelookup.com

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Stuart MacKay, MD, Illawarra ENT Head and Neck Clinic, Suites 1&2, 8-10 Victoria St, Wollongong, NSW 2500, Australia (sgmackay@ozemail.com.au).

JAMA. 2020;324(12):1168-1179. doi:10.1001/jama.2020.14265
Published online September 4, 2020.

Adult obstructive sleep apnea (OSA) is associated with numerous adverse effects if left untreated, including daytime sleepiness, reduced quality of life, increased cardiovascular morbidity and mortality, and an increased risk of motor vehicle crashes.[1-3] A narrow or unstable upper airway[4] predisposes to episodes of complete or partial airflow reduction (ie, obstructive apnea or hypopnea) during sleep when pharyngeal dilator muscle tone falls. Continuous positive airway pressure (CPAP) delivered by mask, or jaw advancement with an oral appliance, can alleviate obstruction; however, many patients either refuse these therapies, use them inconsistently, or (as in the case of oral appliances) achieve suboptimal efficacy, leaving at least 50% of patients without effective treatment[5,6] and exposed to the risk of serious long-term morbidity. Surgical treatments aim to enlarge and stabilize the upper airway and may provide an option for patients in whom conventional medical treatment had failed.[7] However, to date, there have been few randomized trials evaluating the benefits and complications of these surgical procedures compared with controls,[8-11] and most have tested surgical treatment at a single level of the upper airway in highly select patients.[8,9,11]

The Sleep Apnea Multilevel Surgery (SAMS) trial was a randomized clinical trial of a standardized surgical procedure to relieve both retropalatal and retrolingual obstruction in adult patients who had moderate or severe symptomatic OSA for whom conventional device treatment had failed. The primary hypothesis was that surgery would be more effective than ongoing medical management in improving OSA, as measured by a decrease in the frequency of sleep apneas and hypopneas and a reduction in subjective daytime sleepiness.

Key Points

Question Is multilevel airway surgery effective in adults with moderate or severe obstructive sleep apnea (OSA) who cannot tolerate or adhere to device use?

Findings In this randomized clinical trial of 102 adults, modified uvulopalatopharyngoplasty and radiofrequency in saline tongue reduction, compared with ongoing medical management, significantly improved the apnea-hypopnea index (mean baseline-adjusted between-group difference, −17.6 events per hour of sleep) and patient-reported sleepiness (mean baseline-adjusted between-group difference in Epworth Sleepiness Scale, −6.7; range, 0-24; higher score indicates greater sleepiness) at 6 months.

Meaning In this preliminary study of adults with moderate or severe OSA who failed conventional treatment, combined palatal and tongue surgery improved patient-reported sleepiness and polysomnographic measures of OSA severity at 6 months, although further research is needed to confirm these findings in additional populations and to evaluate the safety, clinical utility, and long-term efficacy of multilevel upper airway surgery for treatment of individuals with OSA.

Methods

Ethical Approval and Informed Consent

Each of the participating sites had human research ethics approval to conduct the study. An independent data and safety monitoring board regularly reviewed data on serious and nonserious adverse events and study quality. Written informed consent was obtained from each participant.

Study Design and Patients

This was a multicenter, parallel-group, open-label, randomized clinical trial of upper airway surgery vs ongoing medical management conducted across 6 Australian academic centers. The study reporting conforms to CONSORT guidelines for randomized trials. The full protocol and statistical analysis plan for this trial are available in Supplement 1. The rationale and design of the study has been published previously.[12]

Eligible adults were aged 18 to 70 years with moderate or severe OSA (defined as apnea-hypopnea index [AHI] of 15-30 and >30 events/h of sleep), body mass index less than 38, and Epworth Sleepiness Scale (ESS) greater than 8 (range, 0-24; higher scores indicate greater sleepiness) in whom medically supervised attempts to use CPAP and, when deemed appropriate, a mandibular advancement device failed or were refused. Patients were excluded if they had significant medical or psychiatric comorbidities, were judged to be a high anesthetic risk, were pregnant, or had specific anatomical contraindications to the intended surgery (eg, severe palatal scarring from previous surgery or severe retrognathia). A full description of the inclusion and exclusion criteria is provided in eTable 1 in Supplement 2.

After eligibility was confirmed, participants were randomly assigned to receive either surgery or ongoing medical management at an independent central location. A minimization program (MinimPy 0.3)[13] incorporating biased-coin minimization (base probability, 0.7; marginal balance; 1:1 allocation ratio) was used to ensure balance between variables that had the potential to affect outcomes, namely, study site, sex, age (<50, ≥50 years), AHI (<50, ≥50 events/h), and body mass index (<28, ≥28). Enrollment began in August 2014, but, because of slow recruitment, the protocol was amended in January 2016 after reaching 30% of the recruitment target (31 of 102 planned participants) to relax the body mass index exclusion threshold from greater than 35 to greater than 38 and lower the AHI inclusion requirement from greater than 20 to greater than 15 events per hour.

Interventions

The surgery intervention (eFigure in Supplement 2) consisted of a modified uvulopalatopharyngoplasty to widen and stabilize the velopharynx and 7 to 9 submucosal insertions of a radiofrequency-in-saline wand to reduce tongue volume as previously described.[14] A training workshop was conducted to standardize the surgical techniques among the 7 participating surgeons.[12] Ongoing medical management consisted of a range of evidenced-based treatments as appropriate (eg, weight loss, alcohol reduction, sleep posture modification, medical management of nasal obstruction) and assistance with retrial of CPAP or mandibular advancement device therapies if participants were willing.

Assessments

The follow-up assessment was scheduled for 6 months after the surgical intervention for the surgery group and for 6 months after the date of baseline assessments for the ongoing medical management group. All participants were assessed by a physician certified in sleep medicine at baseline and at 1, 3, and 6 months. Participants in the surgery group were also assessed by a surgeon before surgery; at the time of surgery; and 1, 3, and 6 months after surgery (eTable 2 in Supplement 2). Validated questionnaires measured patient- and partner-reported outcomes at baseline and each follow-up. Demographic information, medical information, OSA treatments used, and body mass index were measured for all enrolled participants at baseline and at each follow-up visit. Information on race was collected because severity of OSA can vary by race, and it was defined by participants using an open-ended questionnaire. Friedman stage was determined at baseline by direct visualization of tongue position, tonsil size, and measurement of body mass index. Stages I through IV denote the presence of anatomical features previously considered to range from the most to least favorable for upper airway soft tissue surgery: stage I indicates low tongue position with moderate or severe tonsil enlargement; stage II, low tongue/no or minor tonsil enlargement or high tongue position/moderate or severe tonsil enlargement; stage III, high tongue position/no or minor tonsil enlargement; and stage IV, body mass index greater than 40 or skeletal deformities (eg, micrognathia or midface hypoplasia).[15] The exclusion criterion for this study precluded the recruitment of patients with Friedman stage IV. All adverse events and postsurgery symptoms, regardless of their severity, were systematically recorded at each of the follow-up appointments.[12] Serious adverse events were based on the guidelines adopted by the International Conference on Harmonization of Good Clinical Practice.[16]

Polysomnography, multiple sleep latency tests, and blood pressure measurements occurred at baseline and 6 months. The sleep tests were scored according to the American Academy of Sleep Medicine 2007 manual[17] at a central sleep laboratory (Adelaide Institute for Sleep Health) by 2 experienced sleep technicians who were blinded to patient treatment allocation. For the AHI, the alternate criteria was used with hypopnea defined as 50% reduction in airflow for at least 10 seconds and associated with either a 3% oxyhemoglobin desaturation or a cortical arousal.[17] The multiple sleep latency test consisted of 5 daytime 20-minute sleep opportunities spaced 2 hours apart from which mean sleep latency was calculated as the mean time to sleep onset after lights out.[18] Office and 24-hour ambulatory blood pressure measurements were obtained at baseline and again at 6 months. It was not possible to blind the research staff assisting with the other evaluations. Further details on patient assessments are provided in Supplement 2.

Study Outcomes

The primary study outcomes were the baseline-adjusted differences between groups in the AHI and ESS[19] at 6 months. The AHI reports the number of apnea and hypopnea events per hour of sleep, with 0 to 5 signifying no OSA; 5 to 14, mild OSA; 15-30, moderate OSA; and greater than 30, severe OSA (a change of ≥15 is considered the minimal clinically important difference).[2] The ESS evaluates patient self-reported sleepiness (range, 0-24; higher scores indicate greater severity; >10 signifies pathological sleepiness; change of 2 indicates the minimal clinically important difference).[20]

Secondary outcomes included the following polysomnography measures of OSA severity and sleep quality: the percentage of participants with AHI less than 10, the apnea index, 3% and 4% oxygen desaturation indices, the lowest oxygen saturation in sleep, percentage of sleep time with oxygen saturation <90%, AHI and lowest oxygen saturation in supine and nonsupine sleep positions, arousal index, and percentages of total sleep time spent in sleep stages nonrapid eye movement stage 1 (N1) through N3 and in rapid eye movement (REM) sleep (mean available age-appropriate normal values: lowest oxygen saturation, 90.5% [95% CI, 89.3%-91.7%]; arousal index, 12.5 [95% CI, 10.7-14.2]; N1 sleep, 8.0% of sleep time [95% CI, 6.9%-9.2%]; N2 sleep, 52.2% of sleep time [95% CI, 50.6%-53.8%]; N3 sleep, 20.4% of sleep time [95% CI, 18.5%-22.2%]; REM sleep, 19.3% of sleep time [95% CI, 18.2%-20.3%])[21]; there are no agreed minimal clinically important difference values for these parameters. Other secondary outcomes were objective daytime sleepiness, measured with the multiple sleep latency test and expressed as mean sleep latency in minutes (values below 10 minutes signify an increased propensity to fall asleep; the minimal clinically important difference is 1 minute[2]); partner-reported patient snoring, measured with the Snoring Severity Scale[22] (range, 0-9; higher scores indicate more intense snoring; no agreed minimal clinically important difference); sleep-specific quality of life, measured with the Functional Outcomes of Sleep Questionnaire-30[23] (range, 5-20; higher scores indicate better functional status; ≥17.9 is considered normal; a score change of 1 is considered to be the minimal clinically important difference[24]); generic health-related quality of life[25] and self-rated health status,[26] measured with the EuroQol Group 5-dimension 5-level (EQ-5D-5L) instrument (utility index range, −0.594 to 1[25]; higher scores indicate better quality of life; score change ranging from 0.03-0.52 are considered the minimal clinically important differences in other diseases there is no minimal clinically important difference established for OSA[27]) and EQ visual analogue scale (range, 0 [worst possible health] to 100 [best possible health]; score changes ranging from 3.5-10.1 considered to be the minimal clinically important differences in other diseases[28,29]); the Glasgow Benefit Inventory,[30] which was completed by participants randomized to the surgery group 6 months after the surgical procedure (change in quality of life following surgery is graded from −100 [poorest outcome] to 100 [best outcome], with 0 indicating no change; no reported minimal clinically important difference); and office and ambulatory 24-hour, daytime, and nighttime blood pressure (minimal clinically important difference for 24-hour ambulatory blood pressure is 1 mm Hg[2]). Some other outcomes were collected to provide mechanistic insights (eg, anthropometric measures, adherence to other OSA therapies). Data on cost, imaging, and nasendoscopic findings were collected but are not reported in this article. Serious adverse event rates were

scored as the percentage of patients with 1 or more serious adverse events.

Statistical Analysis

Based on data from a preliminary study,[12,14] a sample size of 51 participants per group was estimated to provide at least 80% power to detect superior reductions in AHI and ESS with surgery compared with medical management as demonstrated by the lower limit of the 95% CI for the difference in mean values at 6 months being at least 20 events per hour for the AHI and 3 units for the ESS.[12] These differences are greater than the minimal clinically important differences of 15 events per hour for the AHI and 2 units for the ESS to account for the cost and potential morbidity of surgery, a concept termed *sufficiently important difference*.[31]

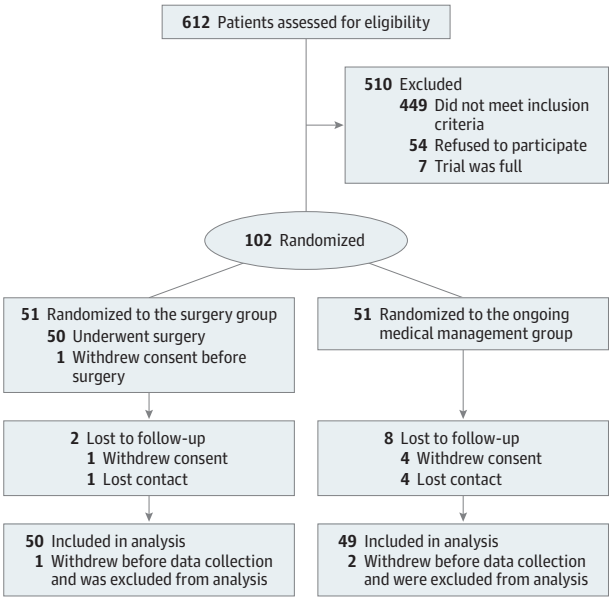
The sample size was based on the ESS primary outcome because it required a larger sample than the AHI primary outcome. The a priori sufficiently important[31] superiority margin was set to 3 in ESS change between the groups at 6 months, and an SD of 5 was estimated based on the preliminary study.[12,14] For the sample size estimation, the *t* test was used at a 2-sided α level of .05 for significance and power of 80% (β = 0.20) and provided an adjustment in case of a nonnormal distribution, yielding the sample size of 102.[12] The use of a mixed-effects model to account for missing data further increased statistical power compared with that estimated with the *t* test.

Descriptive data are presented using mean and SD for normally distributed continuous data and frequency and percentage for categorical data. Within-group changes and between-group differences are presented with 95% CIs. Participants were analyzed according to their randomization group. All participants with data at baseline were included in the analysis, with missing data assumed to be missing at random and any potential bias in estimated effects accounted for by the use of mixed-effects models. For all continuous outcome variables, the effect of treatment was assessed on the basis of the difference in the 6-month follow-up values between groups after adjustment for baseline values. Differences were obtained using linear mixed-effects models with a Poisson link distribution function used for those variables for which the outcome could be interpreted as a count and for which the distribution of either the level 1 or level 2 residuals were nonnormally distributed. The relevant outcome was the dependent variable and fixed-effects terms for the intervention group, visit (baseline or 6 months), and an interaction between group and visit, which was interpreted as the intervention effect. The participant ID was included as a random intercept term. Binary variables were analyzed with the Fisher exact test. Serious adverse event rate between-group differences and 95% CIs were calculated with the Wilson procedure without correction for continuity. All data were analyzed using Stata, version 15.0 (StataCorp).

Effect estimates are reported as mean (95% CI). Hypothesis testing was 2-sided and $P < .05$ was considered statistically significant. Corrections for multiple comparisons were not applied to the 2 primary outcomes because they were considered independent and fall into distinct families of equally important outcomes. Because of the potential for type I error due to multiple comparisons, findings for analyses of secondary end points should be interpreted as exploratory. Although it was not prespecified in the trial protocol or statistical analysis plan (Supplement 1), it was considered important that both primary outcomes (ie, nighttime respiratory disturbance [AHI] and self-reported sleepiness [ESS]) would need to decrease significantly to demonstrate clinical efficacy.

Post hoc sensitivity analyses were conducted to explore whether there was significant heterogeneity of treatment effect among the 7 surgeons by including a group × surgeon interaction term in the models for the 2 primary outcomes (AHI and ESS). Five of the 6 sites had only 1 assessing surgeon and the other site had 2 assessing surgeons. To assess whether there was any clustering effect of participants within surgeon, an additional random intercept for surgeon was used in the mixed-effects models. In addition, because inclusion criteria were altered when enrollment was approximately 30% complete to change AHI from greater than 20 to greater than 15 and body mass index from less than or equal to 35 to less than or equal to 38, a post hoc sensitivity analysis was conducted to determine whether results were different among participants who met the original inclusion criteria for the study.

Figure 1. Recruitment, Randomization, and Flow of Participants in a Study of the Effect of Upper Airway Surgery vs Ongoing Medical Management on Patients With Obstructive Sleep Apnea



Results

Study Overview

A total of 612 participants were screened for eligibility (**Figure 1**): 102 met the eligibility criteria and were enrolled in the study between November 2014 and October 2017, with the last follow-up visit conducted in August 2018. Eleven participants did not complete the study (3 in the surgery group and 8 in the ongoing medical management group). For the 2 primary outcomes, there was missing data for the ESS for

Table 1. Baseline Characteristics of Randomized Participants in a Study of the Effect of Upper Airway Surgery vs Ongoing Medical Management on Patients With Obstructive Sleep Apnea (OSA)

Characteristic	Surgery group (n = 51)	Medical management group (n = 51)
Demographics		
Age, mean (SD), y	42.7 (12.8)	46.4 (12.6)
Sex, No. (%)		
Men	41 (80)	43 (84)
Women	10 (20)	8 (16)
Race, No. (%)		
White	44 (86.3)	46 (90.2)
Other[a]	7 (13.7)	5 (9.8)
Apnea-hypopnea index, mean (SD)[b]	47.9 (23.1)	45.3 (23.9)
Epworth Sleepiness Scale, mean (SD)[c]	12.4 (3.6)	11.1 (4.7)
Anatomy		
Friedman stage, No. (%)[d]		
I (best)	14 (27)	10 (20)
II	23 (45)	21 (41)
III (worst)	14 (27)	20 (39)
Body mass index, mean (SD)		
Men	30.1 (4.0)	30.0 (3.6)
Women	33.3 (2.8)	26.6 (2.9)
Previous OSA treatment, No. (%)		
Tried CPAP	38 (75)	37 (73)
Refused CPAP	13 (25)	14 (27)
Tried mandibular advancement device	16 (31)	12 (24)
Comorbidities, No. (%)[e]		
Respiratory diseases	12 (23.5)	6 (11.8)
Cardiometabolic diseases	10 (19.6)	20 (39.2)
Other chronic disorders	10 (19.6)	7 (13.7)

Abbreviations: CPAP, continuous positive airway pressure.

[a] Race "other" was self-reported by 12 participants as Asian (n = 3), Australian Aboriginal (n = 2), Indian (n = 2), Hispanic/Latino (n = 1), Seychelles (n = 1), Sri Lankan (n = 1), Greek (n = 1), and mixed (n = 1).

[b] The apnea-hypopnea index reports the number of apnea and hypopnea events per hour of sleep (0-5 indicates no OSA; 5-14, mild OSA; 15-30, moderate OSA; and >30, severe OSA; a change of ≥15 is considered the minimal clinically important difference[2]). To take account of the cost and potential morbidity of the surgical intervention, a sufficiently important difference value of 20 was adopted.

[c] Epworth Sleepiness Scale[19] evaluates sleepiness (range, 0-24; higher scores indicate greater severity; score >10 is considered to signify pathological sleepiness; a change of 2 is considered the minimal clinically important difference[20]). To take account of the cost and potential morbidity of the surgical intervention, the sufficiently important difference of 3 was adopted.

[d] Friedman stage incorporates palate size and tongue position, with stage I being most favorable and stage IV least favorable (Friedman stage IV definition includes body mass index >40, which was an exclusion criterion for this study).

[e] Comorbidities were classified as respiratory diseases (asthma and chronic obstructive pulmonary disease), cardiometabolic diseases (diabetes, coronary heart disease, atrial fibrillation, hypertension, stroke), and chronic diseases (arthritis and chronic pain syndrome).

2 participants in the medical management group and 1 participant in the surgery group at both baseline and 6 months. There was missing data for the AHI for 2 participants in the medical management group and for 1 in the surgery group, each of whom withdrew before baseline measurements. Baseline characteristics of trial participants (**Table 1**) showed that the 2 groups were well matched in terms of demographic and clinical characteristics. Participants were predominantly middle-aged men with overweight or obesity and severe OSA. Friedman stages were reasonably evenly distributed within each group. Further details on the study participants and the surgical intervention are provided in eTable 3 to eTable 6 in Supplement 2.

Study Outcomes

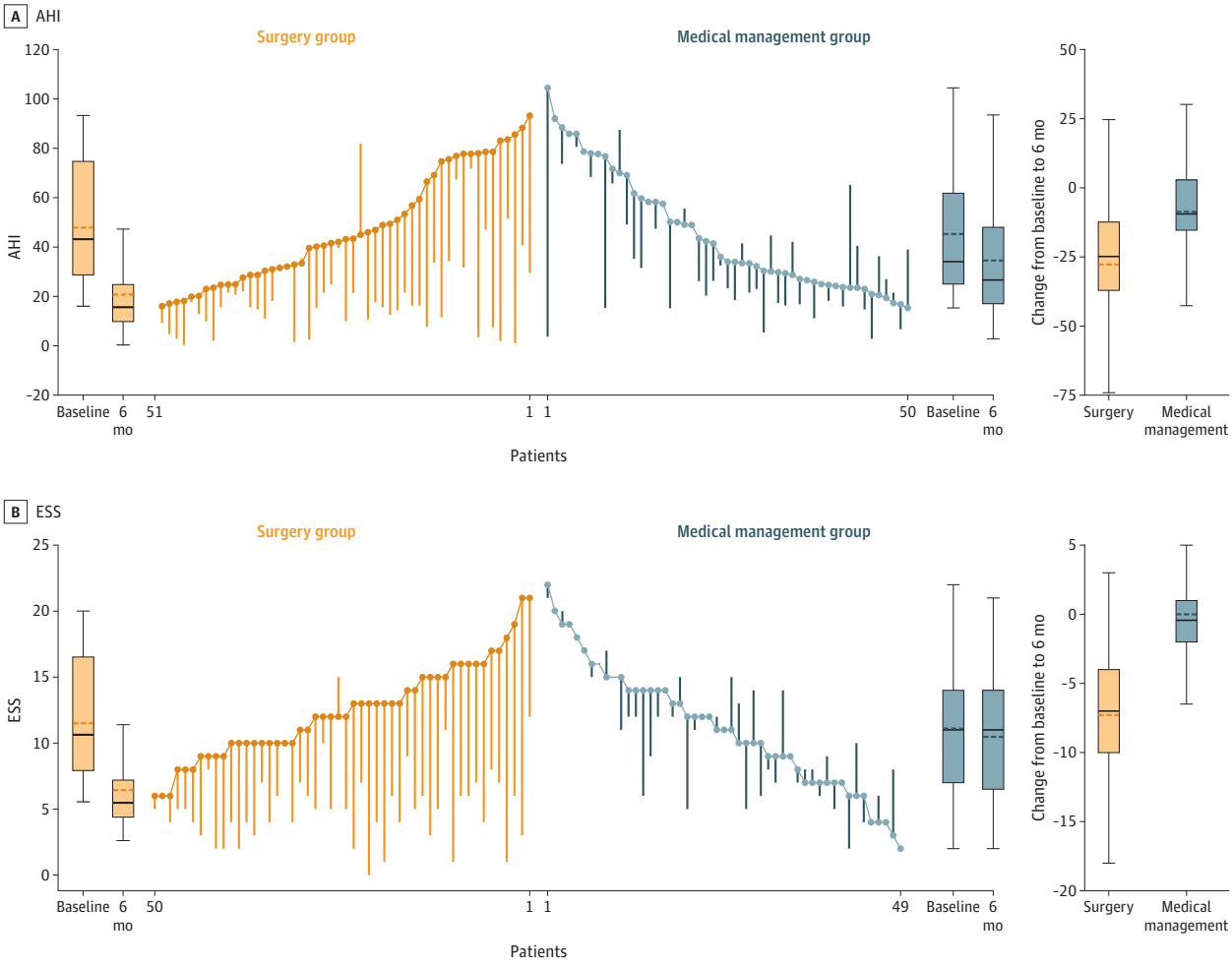
Primary Outcomes

There was a statistically significantly greater improvement from baseline to 6 months in the AHI in the surgery group (47.9 vs 20.8) than in the ongoing medical management group (45.3 vs 34.5) (mean baseline-adjusted between-group difference, −17.6 events/h of sleep [95% CI, −26.8 to −8.4]; $P < .001$) and in the ESS in the surgery group (12.4 vs 5.3) compared with the ongoing medical management group (11.1 vs 10.5) (mean baseline-adjusted between-group difference, −6.7 [95% CI, −8.2 to −5.2]; $P < .001$) (**Figure 2** and **Table 2**).

Secondary Outcomes

Although only a minority of participants achieved complete or near-complete OSA resolution (ie, AHI <10), the percentage of participants who achieved resolution was significantly higher in the surgery group than the ongoing medical management group (26% vs 8%; mean baseline-adjusted between-group difference, 18% [95% CI, 4%-32%]; $P = .01$) (Table 2). Results of analyses of other AHI cutoff values are shown in eTable 7 in Supplement 2. Surgery had a similar favorable significant effect on supine AHI (mean baseline-adjusted between-group difference, −18.7 events/h [95% CI, −31.1 to −6.3]; $P = .003$) and nonsupine AHI (mean baseline-adjusted between-group difference, −18.4 events/h [95% CI, −29.5 to −7.3]; $P = .001$) (eTable 7 in Supplement 2). Compared with participants in the ongoing medical management group, participants in the surgery group had significantly improved sleep quality in terms of lower frequency of arousals (mean baseline-adjusted between-group difference, −11.2 events/h of sleep [95% CI, −14.9 to −7.5]; $P < .001$), less light sleep (N1 sleep: mean baseline-adjusted between-group difference, −3.9% of sleep time [95% CI, −5.5% to −2.3%]; $P < .001$), and improved sleep oxygenation (3% oxygen desaturation index: baseline-adjusted between-group difference, −13.5 events/h of sleep [95% CI, −20.0 to −7.1]; $P < .001$) (Table 2 and eTable 7 in Supplement 2), although several of these parameters remained abnormal at study completion. Multiple sleep latency test findings were within normal limits at baseline and showed no significant between-group difference in mean sleep onset latency at 6 months (mean baseline-adjusted between-group difference, 0.6 min [95% CI, −1.3 to 2.5]; $P = .54$). At 6 months, there were statistically significant mean between-group differences favoring surgery for partner-reported snoring (Snoring Severity Index: −4.0 [95% CI, −4.9 to −3.1]; $P < .001$), sleep-specific quality of life (Functional Outcomes of Sleep Questionnaire score: 3.4 [95%

Figure 2. Apnea-Hypopnea Index (AHI) and Epworth Sleepiness Scale (ESS) Outcomes in a Study of the Effect of Upper Airway Surgery vs Ongoing Medical Management on Patients With Obstructive Sleep Apnea (OSA)



The ends of the boxes in the boxplots are located at the first and third quartiles, with the black line in the middle illustrating the median. The dashed line signifies the mean. Whiskers extend to the upper and lower adjacent values, the location of the furthest point within a distance of 1.5 interquartile ranges from the first and third quartiles. The parallel line plot contains 1 vertical line for each patient which extends from their baseline value to their 6-month value.[40] Descending lines indicate an improvement in symptoms. Baseline values are placed in ascending order for the surgery group and descending order for the ongoing medical management group. A, The AHI indicates the number of apnea and hypopnea events per hour of sleep (0-5 is classified as normal; 5-14, mild OSA; 15-30, moderate OSA; >30, severe OSA; a change of at least 15 is considered clinically meaningful and can move a patient 2 levels from severe to mild with established benefit for health[2]). B, The ESS[19] evaluates sleepiness (range, 0-24; higher scores indicate greater severity; score >10 signifies pathological sleepiness; a change of 2 is the minimally important clinical difference[20]).

CI, 2.5-4.4]; $P < .001$), and self-rated general health status (EQ visual analog scale score: 10.5 [95% CI, 5.6-15.4]; $P < .001$), but were not statistically significantly different for general health-related quality of life (EQ-5D-5L score: 0.06 [95% CI, 0.00-0.12]; $P = .054$) (Table 2). Participants in the surgery group identified specific health benefits from the surgical procedures (mean [SD] Glasgow Benefit Inventory, 32.1 [22.8]) (eTable 7 in Supplement 2). Mean blood pressure values were normal at baseline and showed no significant between-group difference at 6 months (in-office systolic blood pressure: −2.5 [95% CI, −7.1 to 2.2]; $P = .30$; in-office diastolic blood pressure: −3.9 [95% CI, −8.1 to 0.3]; $P = .07$) (Table 2 and eTable 7 in Supplement 2).

Very few participants attempted to use CPAP (n = 7) or a mandibular advancement device (n = 1) during the trial, and there was no significant weight loss in either study group (between-group difference, −1.2 kg [95% CI, −2.8 to 0.4]; $P = .14$) or change in health behaviors (eTable 7 in Supplement 2).

Post Hoc Sensitivity Analyses

There was no significant difference in treatment effect among surgeons for the AHI ($\chi^2_5 = 5.66$; $P = .34$) or the ESS ($\chi^2_5 = 4.51$; $P = .48$). There was also no significant difference in the treatment effects after incorporating a random intercept for surgeons into each model or when using data only for participants who met the original inclusion criteria (eTable 8 in Supplement 2).

Table 2. Estimated Treatment Effects in the Primary and Secondary Outcomes in a Study of the Effect of Upper Airway Surgery vs Ongoing Medical Management on Patients With Obstructive Sleep Apnea

Outcome	Surgery group (n = 50)			Medical management group (n = 49)			Baseline-adjusted difference of surgery vs medical management group, mean (95% CI)[a]	*P* value[b]
	Baseline, mean (SD)	6 mo, mean (SD)	Change from baseline, mean (95% CI)	Baseline, mean (SD)	6 mo, mean (SD)	Change from baseline, mean (95% CI)		
Primary outcomes								
Apnea-hypopnea index[c]	47.9 (23.1)	20.8 (18.4)	−27.4 (−33.8 to −21.0)	45.3 (23.9)	34.5 (23.0)	−9.8 (−16.5 to −3.1)	−17.6 (−26.8 to −8.4)	<.001
Apnea-hypopnea index, median (IQR)	43.2 (28.7 to 74.7)	15.6 (9.8 to 24.8)		34.1 (25.0 to 61.8)	26.7 (17.1 to 48.0)			
Epworth Sleepiness Scale[d]	12.4 (3.6)	5.3 (3.0)	−7.2 (−8.3 to −6.2)	11.1 (4.7)	10.5 (4.7)	−0.51 −1.6 to 0.6)	−6.7 (−8.2 to −5.2)	<.001
Secondary outcomes								
Secondary polysomnography OSA severity and sleep quality[e]								
Apnea-hypopnea index <10, n (%)	0 (0)	13 (26)	13 (26)	0	4 (8)	4 (8)	18% (4 to 32)	.01[f]
Apnea index	12.7 (14.5)	3.1 (5.6)	−13.4 (−20.5 to −6.3)	13.2 (21.1)	6.8 (12.1)	−6.0 (−9.6 to −2.4)	−7.4 (−14.3 to −0.5)	.04[g]
Apnea index, median (IQR)	8.3 (1.5 to 21.0)	0.4 (0.0 to 4.4)		3.6 (1.0 to 17.2)	2.75 (0.7 to 6.25)			
3% oxygen desaturation index (desaturation events/h of sleep)	29.0 (17.9)	11.4 (12.7)	−17.7 (−22.2 to −13.3)	27.0 (19.8)	21.6 (17.8)	−4.2 (−8.9 to 0.4)	−13.5 (−20.0 to −7.1)	.001
3% oxygen desaturation index (desaturation events/h of sleep), median (IQR)	25.1 (15.3 to 43.7)	6.7 (4.8 to 14.4)		22.1 (11.0 to 40.6)	15.7 (9.3 to 34.2)			
4% oxygen desaturation index (desaturation events/h of sleep)	23.2 (17.8)	8.1 (11.1)	−15.3 (−19.6 to −11.0)	21.5 (19.9)	15.9 (16.2)	−4.3 (−8.8 to 0.1)	−11.0 (−17.2 to −4.7)	.003
4% oxygen desaturation index (desaturation events/h of sleep), median (IQR)	17.8 (7.6 to 33.6)	3.6 (2.0 to 8.2)		12.6 (7.2 to 36.0)	9.5 (4.1 to 23.6)			
Lowest oxygen saturation, %	79.0 (8.7)	83.9 (6.1)	4.9 (2.8 to 7.0)	80.7 (9.1)	81.5 (8.9)	0.4 (−1.7 to 2.6)	4.5 (1.5 to 7.5)	<.001
Total sleep time spent with oxygen desaturation <90%, %	8.8 (10.9)	3.8 (8.1)	−8.5 (−14.1 to −2.8)	8.3 (13.5)	6.0 (12.2)	−0.8 (−2.3 to 0.7)	−7.7 (−13.3 to −2.0)	.008[g]
Total sleep time spent with oxygen desaturation <90% median (IQR), %	4.3 (0.6 to 11.8)	0.7 (0.1 to 2.1)		2.3 (0.2 to 8.7)	1.2 (0.1 to 4.1)			
Arousal index (arousals/h of sleep)	33.4 (18.6)	19.0 (10.8)	−15.1 (−18.0 to −12.2)	31.7 (18.1)	25.8 (15.0)	−3.9 (−6.2 to −1.5)	−11.2 (−14.9 to −7.5)	<.001[g]
Sleep stages (% total sleep time)								$\chi^2_3 = 13.46$; $P = .004$[h]
N1	11.4	11.8	−0.4 (−1.3 to −2.1)	12.1	9.4	−2.7 (−−7.2 to 1.8)	−3.9 (−5.5 to −2.3)	<.001[i]
N2	63.0	56.1	−6.9 (−13.9 to −1.3)	61.3	57.7	−3.6 (−8.8 to 1.6)	0.1 (−0.8 to 0.9)	.87[i]
N3	10.7	14.0	3.3 (−1.7 to 8.3)	11.7	14.4	2.7 (−1.8 to 7.2)	2.6 (−1.3 to 1.8)	.75[i]
REM	14.9	18.1	3.2 (−1.7 to 8.1)	14.9	18.5	3.6 (−1.6 to 8.8)	0.8 (−1.0 to 2.5)	.38[i]

(continued)

Table 2. Estimated Treatment Effects in the Primary and Secondary Outcomes in a Study of the Effect of Upper Airway Surgery vs Ongoing Medical Management on Patients With Obstructive Sleep Apnea (continued)

Outcome	Surgery group (n = 50)			Medical management group (n = 49)			Baseline-adjusted difference of surgery vs medical management group, mean (95% CI)[a]	*P* value[b]
	Baseline, mean (SD)	6 mo, mean (SD)	Change from baseline, mean (95% CI)	Baseline, mean (SD)	6 mo, mean (SD)	Change from baseline, mean (95% CI)		
Objective daytime sleepiness								
Mean sleep latency (min)[j]	10.5 (5.2)	11.1 (4.3)	0.8 (−0.5 to 2.1)	11.6 (5.3)	11.9 (4.7)	0.2 (−1.1 to 1.6)	0.6 (−1.3 to 2.5)	.54[k]
Partner- and patient-reported outcomes								
Snoring severity scale[l]	7.7 (0.9) (n = 35)	2.6 (2.3) (n = 23)	−5.1 (−5.8 to −4.4)	7.1 (1.5) (n = 32)	6.6 (1.7) (n = 26)	−0.56 (−1.23 to 0.10)	−4.0 (−4.9 to −3.1)	<.001
Functional outcomes of Sleep questionnaire[n]	15.1 (2.8)	18.6 (1.8)	3.6 (2.9 to 4.2)	16.1 (2.7)	16.3 (2.4)	0.1 (- 0.6 to 0.8)	3.4 (2.5 to 4.4)	<.001[m]
EQ-5D-5L[o]	0.86 (0.20)	0.93 (0.12)	0.06 (0.02 to 0.11)	0.86 (0.10)	0.86 (0.10)	0.00 (−0.03 to 0.04)	0.06 (−0.00 to 0.12)	.054
EQ visual analog scale[p]	76.4 (16.4)	85.0 (12.1)	8.8 (5.4 to 12.2)	78.7 (13.2)	76.8 (14.0)	−1.7 (−5.3 to 1.8)	10.5 (5.6 to 15.4)	<.001
Blood pressure[q]								
24 h ambulatory systolic blood pressure	121.9 (8.4)	120.0 (12.0)	−1.5 (−5.0 to 1.9)	125.4 (12.2)	124.7 (13.6)	−1.3 (−4.6 to 1.9)	−0.2 (−4.9 to 4.6)	.94
24 h ambulatory diastolic blood pressure	75.3 (6.0)	74.1 (8.0)	−1.3 (−3.5 to 1.0)	77.4 (9.1)	77.7 (10.9)	−0.4 (−2.6 to 1.7)	−0.9 (−4.0 to 2.3)	.59
Anthropometry								
Body mass index	30.7 (3.9)	30.6 (4.2)	−0.14 (−0.51 to 0.23)	29.5 (3.7)	29.4 (3.7)	0.29 (−0.10 to 0.68)	0.43 (−0.10 to 0.97)	.11

Abbreviations: EQ, EuroQol; EQ-5D-5L, EuroQol 5-dimension 5-level questionnaire; N, nonrapid eye movement stage; OSA, obstructive sleep apnea; REM, rapid eye movement sleep.

[a] Using the group × visit interaction term from a linear mixed-effects model with group, visit (baseline or 6 months), and group × visit as fixed effects and participant as random intercept.

[b] For group × visit interaction term.

[c] For definition see footnotes for Table 1.

[d] For definition see footnotes for Table 1.

[e] Available age-appropriate normal values for secondary polysomnographic parameters: lowest oxygen saturation, 90.5% (95% CI, 89.3%-91.7%); arousal index, 12.5 (95% CI, 10.7-14.2); N1 sleep, 8.0% of sleep time (95% CI, 6.9%-9.2%); N2 sleep, 52.2% of sleep time (95% CI, 50.6%-53.8%); N3 sleep, 20.4% of sleep time (95% CI, 18.5%-22.2%); REM sleep, 19.3% of sleep time (95% CI, 18.2%-20.3%).[21] There are no agreed minimum important clinical difference values for secondary polysomnography parameters.

[f] Assessed using Fisher exact test.

[g] Assessed using mixed-effects model with a Poisson distribution to account for data nonnormality, the mean differences were estimated from these models.

[h] Using χ^2 test of association for overall difference between groups across the 4 stages of sleep at 6 months.

[i] Using a mixed-effects logistic regression model for each separate sleep stage with adjustment for baseline percentages.

[j] Mean sleep latency is the average time to fall asleep (measured by polysomnography in a sleep laboratory) across 5 equally spaced 20-minute daytime nap opportunities when the patient is instructed to try to fall asleep. Values below 10 minutes signify an increased propensity to fall asleep. The minimum important clinical difference is 1 minute[2].

[k] Measured by the Multiple Sleep Latency Test, mean sleep latency calculated as the mean time to sleep onset after lights out on 5 daytime 20-minute sleep opportunities spaced 2 hours apart.

[l] Snoring severity scale[22] evaluates snoring (range, 0-9; higher scores indicate more intense snoring).

[m] Assessed using a linear mixed effects model with bias corrected and accelerated bootstrapped 95% CIs.

[n] Functional Outcomes of Sleep Questionnaire[23] is a measure for sleep-specific quality of life (range, 5-20; higher scores indicate better functional status; normal is considered a score of ≥17.9; a score change of 1 is considered to be the minimal clinically important difference[24]).

[o] EuroQol group EQ-5D-5L[26] questionnaire is a measure of general health-related quality of life (utility index range, −0.594 to 1[25]; higher index score indicates better quality of life; score changes ranging from 0.03-0.52 are considered to be the minimally important differences in other diseases [there is no minimal clinically important difference established for OSA][27]).

[p] EQoL-5D-5L visual analogue scale[26] is a measure of self-rated general health status (range, 0 [worst possible health] to 100 [best possible health]; score changes ranging from 3.5-10.1 are considered to be the minimal clinically important differences in other diseases).[28,29]

[q] The minimal clinically important difference for 24-hour ambulatory blood pressure is 1 mm Hg.[2]

Adverse Events

There were 6 serious adverse events in 4 participants in the surgery group and no serious adverse events in the ongoing medical management group. Three serious adverse events occurred in the same patient: myocardial infarction on postoperative day 5, tonsillar fossa bleeding after initiation of anticoagulation therapy on postoperative day 14, and recurrent angina requiring a second coronary artery stent on postoperative day 21. Another serious adverse event in a different patient was hospital readmission lasting more than 24 hours for observation (a criterion for serious adverse event) after hematemesis of old blood on postoperative day 10.

Table 3. Adverse Events in a Study of the Effect of Upper Airway Surgery vs Ongoing Medical Management on Patients With Moderate or Severe Obstructive Sleep Apnea

Adverse event	No. of events	
	Surgery group	Medical management group
Serious[a] and related to surgical procedure		
Results in death	0	
Life-threatening	1	
Requires in-patient hospitalization (>24 h) or prolongation of existing hospitalization	2	
Results in persistent or significant disability or incapacity	0	
Serious[a] and unrelated to surgical procedure		
Results in death	0	0
Life-threatening	0	0
Requires in-patient hospitalization (>24 h) or prolongation of existing hospitalization	3	0
Results in persistent or significant disability or incapacity	0	0
Nonserious and related to surgical procedure		
Minor postoperative tonsil bleeds (Stammberger grade A-B; hospitalization for observation)	5	
Hospital admission <24 h (for observation only)	1	
Globus pharyngeus[b] (at 6 mo)	0	
Taste changes (at 6 mo)	1	
Tongue numbness (at 6 mo)	0	
Tongue weakness (at 6 mo)	0	
VPI speech (at 6 mo)	0	
VPI swallowing/regurgitation (at 6 mo)	1	
VPI hypernasality (at 6 mo)	0	
Other (excess saliva, feeling of scar tissue on roof of mouth)	2	
Nonserious and unrelated to surgical procedure		
Planned elective or hospital admission (<24 h)	2	3
Injuries or other medical conditions	6	5

Abbreviation: VPI, velopharyngeal insufficiency.

[a] There were 6 serious adverse events in 4 participants in the surgery group and 0 serious adverse events in the ongoing medical management group. Three serious adverse events occurred in the same patient. Two other patients had a serious adverse event prior to surgery. Further clinical details concerning all adverse events are located in eTable 10 in Supplement 2.

[b] Globus pharyngeus is a persistent or intermittent nonpainful sensation of a lump or foreign body in the throat.

The patient was hemodynamically stable with a normal hemoglobin concentration and was admitted to the hospital for observation. The remaining 2 serious adverse events (hospital admission for asthma/bronchitis and colitis) occurred in 2 participants after randomization but before the surgery. Thus, 2 of the 50 participants (4%) who underwent the surgery were considered to have serious adverse events possibly related to surgery and 0 of the 49 participants in the ongoing medical management group experienced a serious adverse event (between-group difference, 4.0% [95% CI, -3.8% to 13.5%]). At 6 months, 4 participants in the surgery group reported oral or pharyngeal symptoms, none of which were associated with significant functional impairment (**Table 3**; see eTable 9 in Supplement 2 for further details of all serious and nonserious adverse events).

Discussion

In this randomized clinical trial, multilevel upper airway surgery, compared with ongoing medical management, resulted in significant reductions in the frequency of sleep apneas and hypopneas and daytime sleepiness in patients with moderate or severe OSA in whom prior attempts at conventional medical device treatment had failed. Surgery was also associated with improvements in most other polysomnography measures (arterial oxygen saturation measures and cortical arousal frequency), partner-reported snoring, and patient-reported sleep-specific quality of life and general health status. Mean sleep latency and blood pressure values were in the normal range and were unchanged at 6 months.

The statistically significant between-group difference in the AHI was above the established minimal clinically important difference (15 events/h)[2] but less than the a priori hypothesized difference of 20 events per hour.[12] The accompanying significant decrease in ESS exceeded the a priori sufficiently important difference of 3 for the trial.[12] The significant improvements in self-reported sleepiness and sleep-specific quality of life were considerably greater than those found in patients with OSA treated with CPAP.[2] Although the lack of change in the multiple sleep latency test might appear discordant with the marked change in ESS, these 2 tests measure different aspects of daytime sleepiness as demonstrated by their poor correlation in OSA populations.[32] Compared with ESS, mean sleep latency is relatively insensitive to change with OSA treatment[2] and, unlike ESS, was within normal limits at baseline.

The significant surgery-related improvements in this study are similar to those reported in 2 previous surgery randomized clinical trials[8,9] conducted in patients with similar OSA severity. However, these trials used uvulopalatopharyngoplasty alone in patients because of predominant retropalatal obstruction and/or palatine tonsil enlargement (ie, Friedman stage I or II). Most patients with OSA have multilevel obstruction, including increased tongue size due to fat deposition.[33] Thus, this trial of multilevel surgery supports a broader role for upper airway surgery to manage OSA and expands on an earlier randomized clinical trial of multilevel minimally invasive surgery in patients with mild disease.[10] The more extensive approach used in the present study had a greater treatment effect, whereas the serious adverse event risk was similar to a large observational cohort of predominantly uvulopalatopharyngoplasy alone.[34] None of the participants reported significant long-term functional difficulties. Two participants experienced serious adverse events potentially related to surgery (another 2 participants randomized to the surgery group had serious adverse events before undergoing the surgical procedures).

Based on the preliminary study cohort,[14] and that nonanatomical as well as anatomical factors contribute to

OSA,[4] it was anticipated that this surgical intervention would reduce, not eliminate, obstructive breathing events. The reduction in AHI in the surgery group was substantial and similar to the net effect of CPAP treatment[35] and the decrease in AHI achieved with oral appliances[36] and hypoglossal nerve stimulation.[37] Most importantly, this study has shown substantial improvements in patient-centered outcomes in patients unable to use conventional OSA treatment. The surgery does not preclude reintroducing CPAP or other therapies later, if required. The perception that CPAP treatment is problematic after uvulopalatopharyngoplasty arises from a single early report of more mouth leaks with nasal CPAP after excisional uvulopalatopharyngoplasty.[38] However, the modified uvulopalatopharyngoplasty used in this study reduces retropalatal obstruction while preserving the palate and velopharyngeal sphincter function such that problems with reapplication of CPAP are rare.[39]

The strengths of this study include the randomized clinical trial design, sufficient sample size and duration of follow-up to ascertain treatment effects independent of short-term postoperative discomfort and physiological disturbance, systematic recording of adverse events, and low rates of participant withdrawal and loss to follow-up. Standardization of the surgical intervention was done to try to ensure consistency among the trial surgeons at multiple sites. The surgical technique is similar to other contemporary, widely used uvulopalatopharyngoplasties that open the lateral palate,[8,9] and radiofrequency is a common otolaryngologic intervention. The findings may have broader clinical relevance because the surgical exclusion criteria were relatively limited and most anatomical subtypes were included, as 84% of patients meeting general screening and medical criteria were surgically eligible and enrolled in the trial.

Limitations

This study has several limitations. First, although the study was adequately powered to establish efficacy, generalizability from any randomized clinical trial is inherently limited by numbers of patients and surgeons, and this study did not meet the sufficiently important difference for AHI (but did surpass the minimal clinically important difference). Thus, further studies will be needed to establish the long-term effectiveness, safety, and cost-effectiveness of this surgical treatment for OSA. Second, this study included a select population that excluded patients with severe obesity (ie, body mass index of 38 or greater), patients older than 70 years, and patients with retrognathia and significant comorbidities; women were underrepresented in the trial. Therefore, the results may not generalize to the larger OSA population. Third, because the surgical intervention caused pain and anatomical changes, it was not possible to blind patients and some of the clinical assessors to treatment group randomization. Hence, some of the observed effects, such as self-reported sleepiness, may have been influenced by participant expectancy, although this bias is less likely to have affected the blinded objective measures of OSA (polysomnography).

Conclusions

In this preliminary study of adults with moderate or severe OSA who had failed conventional treatment, combined palatal and tongue surgery, compared with ongoing medical management, reduced the number of apnea and hypopnea events and patient-reported sleepiness at 6 months. Further research is needed to confirm these findings in additional populations and to understand clinical utility, long-term efficacy, and safety of multilevel upper airway surgery for treatment of patients with OSA.

ARTICLE INFORMATION

Accepted for Publication: July 15, 2020.

Published Online: September 4, 2020. doi:10.1001/jama.2020.14265

Author Affiliations: Illawarra Shoalhaven Local Health District, Wollongong, NSW, Australia (MacKay, Jones); University of Wollongong, NSW, Australia (MacKay, Jones); Illawarra ENT Head and Neck Clinic, Wollongong, NSW, Australia (MacKay); Southern ENT & Adelaide Sinus Centre, Flinders Private Hospital, Adelaide, SA, Australia (Carney); College of Medicine and Public Health, Flinders University, Adelaide, SA, Australia (Carney, Ooi, Woods); Adelaide Institute for Sleep Health, College of Medicine and Public Health, Flinders University, Adelaide, SA, Australia (Catcheside, Chai-Coetzer, Pinczel, McEvoy); Respiratory and Sleep Service, Southern Adelaide Local Health Network, Adelaide, SA, Australia (Chai-Coetzer, McEvoy); Department of Thoracic Medicine, Royal Adelaide Hospital, Adelaide, SA, Australia (Chia, Yeo); Charles Perkins Centre, Faculty for Medicine and Health, University of Sydney, NSW, Australia (Cistulli); Department of Respiratory and Sleep Medicine, Royal North Shore Hospital, Sydney, NSW, Australia (Cistulli); Ear Nose and Throat Department, Royal Adelaide Hospital, Adelaide, SA, Australia (Hodge); Illawarra Sleep Medicine Centre, Wollongong, NSW, Australia (Jones); Health Economics Unit, College of Medicine and Public Health, Flinders University, Adelaide, SA, Australia (Kaambwa); Hollywood Medical Centre, Perth, WA, Australia (Lewis); Department of Otolaryngology, Head & Neck Surgery, Royal Perth Hospital, Perth, WA, Australia (Lewis); Department of Otolaryngology, Head and Neck Surgery, Flinders Medical Centre, Adelaide, SA, Australia (Ooi, Woods); West Australian Sleep Disorders Research Institute, Queen Elizabeth II Medical Centre, Perth, WA, Australia (McArdle, Singh); Department of Pulmonary Physiology and Sleep Medicine, Sir Charles Gairdner Hospital, Perth, WA, Australia (McArdle, Singh); The Memorial Hospital, Adelaide, SA, Australia (Rees); Faculty of Human Sciences, University of Western Australia, Perth, WA, Australia (Singh); The Woolcock Clinic, University of Sydney, NSW, Australia (Stow); Department of Otolaryngology/Head and Neck Surgery, University of Washington, Seattle (Weaver); Staff Surgeon, Seattle Veterans Affairs Medical Center, Seattle, Washington (Weaver); Flinders Centre for Epidemiology and Biostatistics, College of Medicine and Public Health, Flinders University, Adelaide, SA, Australia (Woodman).

Author Contributions: Drs McEvoy and Woodman had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: MacKay, Carney, Catcheside, Chai-Coetzer, Chia, Cistulli, Hodge, Weaver, Woods, Yeo, McEvoy.
Acquisition, analysis, or interpretation of data: All authors.
Drafting of the manuscript: MacKay, Carney, Catcheside, Chia, Cistulli, Hodge, Jones, Woodman, Woods, Yeo, McEvoy.
Critical revision of the manuscript for important intellectual content: MacKay, Catcheside, Chai-Coetzer, Chia, Cistulli, Hodge, Jones, Kaambwa, Lewis, Ooi, Pinczel, McArdle, Rees, Singh, Stow, Weaver, Woodman, Woods, Yeo, McEvoy.
Statistical analysis: Catcheside, Kaambwa, Weaver, Woodman, Woods, Yeo.
Obtained funding: MacKay, Carney, Catcheside, Chai-Coetzer, Cistulli, Ooi, Weaver, Woods, McEvoy.
Administrative, technical, or material support: MacKay, Carney, Catcheside, Chai-Coetzer, Chia, Jones, Ooi, Pinczel, McArdle, Rees, Stow, Woods.
Supervision: MacKay, Chai-Coetzer, Chia, Cistulli, Hodge, Jones, Ooi, Rees, McEvoy.
Other - Trial PI: McEvoy.

Conflict of Interest Disclosures: Dr MacKay reported receiving grants from National Health and Medical Research Council of Australia, Flinders

University, and The Repat Foundation 2013 Prabha Seshadri Research Grant during the conduct of the study and nonfinancial support from Genio-Nyxoah and grants from Garnett-Passe Rodney Williams Foundation and Illawarra Health and Medical Research Institute outside the submitted work. Dr Carney reported receiving grants from National Health and Medical Research Council of Australia, Flinders University, and The Repat Foundation Prabha Seshadri Research Grant during the conduct of the study. Dr Catcheside reported receiving grants from National Health and Medical Research Council of Australia and from Flinders University during the conduct of the study; receiving grants from the Cooperative Research Centre for Alertness, Safety and Productivity, the National Health and Medical Research Council of Australia, and the Flinders Foundation Health Seed Grant and nonfinancial support from Philips Respironics, Air Liquide Healthcare, and Gorman Promed outside the submitted work; having a patent to Sleep Disorder Diagnosis, App No. AU 2018904007, US 62/812066 pending and a patent to decision support software for sleep disorder identification, February 28, 2019, US 61/815562 Sleep Disorder Data Fidelity Management System, March 8, 2019, pending. Dr Chai-Coetzer reported receiving grants from the National Health and Medical Research Council of Australia, Flinders University, and The Repat Foundation during the conduct of the study. Dr Cistulli reported receiving grants from the National Health and Medical Research Council of Australia, Flinders University, and The Repat Foundation Prabha Seshadri Research Grant during the conduct of the study and nonfinancial support from ResMed Inc; nonfinancial support from Zephyr Sleep Technologies; personal fees from SomnoMed, Signifier Medical, Wolter Kluwers, and Quintessence Publishing; and a provision of clinical trial services from Bayer outside the submitted work. Dr Jones reported receiving grants from Illawarra Sleep Medicine Centre during the conduct of the study. Dr Kaambwa reported receiving grants from the Australian National Health and Medical Research Council during the conduct of the study. Dr Lewis reported receiving consulting fees for Nyxoah SA (Belgium) as an investigator in 3 trials of a new hypoglossal nerve stimulator for obstructive sleep apnea. Dr Ooi reported receiving grants from the National Health and Medical Research Council and Flinders University during the conduct of the study and personal fees from Medtronic outside the submitted work. Dr Pinczel reported receiving grants from the National Health and Medical Research Council of Australia, Flinders University, and The Repat Foundation during the conduct of the study. Dr McArdle reported receiving grants from the National Health and Medical Research Council of Australia, Flinders University, and The Repat Foundation Prabha Seshadri Research Grant during the conduct of the study and grants from Oventus (Brisbane, Australia), Nyxoah (Mont-Saint Guibert, Belgium), and Zelda Therapeutics (Australia) outside the submitted work. Dr Weaver reported being a research consultant for Nyxoah SA outside the submitted work and practicing sleep surgery. Dr Yeo reported receiving grants from the National Health and Medical Research Council of Australia, Flinders University, Flinders University, and The Repat Foundation Prabha Seshadri Research Grant during the conduct of the study. Dr McEvoy reported receiving grants from the National Health and Medical Research Council of Australia, Flinders University, and The Repat Foundation during the conduct of the study and grants from Philips Respironics and ResMed outside the submitted work. No other disclosures were reported.

Funding/Support: This study is supported by a project grant from the National Health and Medical Research Council of Australia (1059510). Supplementary trial funding was provided by Flinders University and The Repat Foundation Prabha Seshadri Research Grant. Dr McEvoy was supported by a practitioner fellowship of the National Health and Medical Research Council (GNT 1079785). Dr Catcheside was supported by an Australian Research Council Future Fellowship (FT120100510). Dr. Weaver's effort was supported by the Unites States National Institutes of Health (R01 HL084139, PI Weaver) and the Veterans Affairs Puget Sound Health Care System.

Role of the Funder/Sponsor: The sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. The contents of this article do not represent the views of the US government.

Group Information: The Sleep Apnea Multilevel Surgery (SAMS) Trial Group *Investigators*: Stuart MacKay, MD; A. Simon Carney, DM; Peter G. Catcheside, PhD; Ching Li Chai-Coetzer, MD, PhD; Michael Chia, MD; Peter A. Cistulli, MD, PhD; John-Charles Hodge, MD; Andrew Jones, MD; Billingsley Kaambwa, PhD; Richard Lewis, MD; Eng H. Ooi, MD, PhD; Alison Pinczel, PhD; Nigel McArdle, MD; Guy Rees, MD; Bhajan Singh, MD, PhD; Nicholas Stow, MD; Edward M. Weaver, MD, MPH; Richard J. Woodman, PhD; Charmaine M. Woods, PhD; Aeneas Yeo, MD; Nick A. Antic, PhD; R. Doug McEvoy, MD. *Data and safety monitoring board:* Matthew Naughton, MD (sleep and respiratory medicine specialist; Alfred Hospital and Monash University, Melbourne, Australia); Nathaniel Marshall, PhD (clinical trialist, epidemiologist, sleep disorders specialist; University of Sydney and The Woolcock Medical Institute, Sydney, Australia); and William Coman, AM, MD (consultant ENT surgeon; University of Queensland). *Sleep laboratory managers:* Jeremy Mercer, PhD (Flinders Medical Centre); Terry Sands, MD (Wollongong Hospital); Pam Singh, BSc(Hons) (Royal Adelaide Hospital); Neill Madeira, BHealthSc (Sir Charles Gairdner Hospital); Gary Cohen, PhD (for Sydney site). *Corelab manager:* Jeremy Mercer, PhD. *Corelab scoring:* Laura Bandick, BPsych(Hons); Michaela O'Keefe, BPsych(Hons). *Additional site coordinators*: Alison Teare, BSc, BEd; Kelsey Johnson, BPsych(Hons); Natasha Umrigar, BmedHlthSciAdv(Hons); Sue-Ellen Holmes, BSc (Hons); Kate Sutherland, PhD; Aimee Lowth, BSc (Hons); Clare Rea, BSc, BA.

Meeting Presentation: This study was presented at the Australasian Sleep Association Conference; October 19, 2018; Brisbane, Australia, and at the virtual European Respiratory Society International Congress; September 7, 2020.

Data Sharing Statement: See Supplement 3.

Additional Contributions: This work would not have been possible without both the late Samuel Robinson, MD (Flinders Medical Centre), who pioneered the surgical technique and the late Nick Antic, PhD (Flinders University). Dr Antic brought together the collaboration of sleep physicians and surgeons, led the funding proposal, and was the study's principal investigator before his untimely death (in 2016). We thank the patients who participated in this trial; the various clinical and research teams of the Sleep and Respiratory, and Otolaryngology Head and Neck Surgery departments; and the technicians in the core sleep laboratory at the Adelaide Institute of Sleep Health for their work on the study.

REFERENCES

1. Xie C, Zhu R, Tian Y, Wang K. Association of obstructive sleep apnoea with the risk of vascular outcomes and all-cause mortality: a meta-analysis. *BMJ Open*. 2017;7(12):e013983. doi:10.1136/bmjopen-2016-013983

2. Patil SP, Ayappa IA, Caples SM, Kimoff RJ, Patel SR, Harrod CG. Treatment of adult obstructive sleep apnea with positive airway pressure: an american academy of sleep medicine systematic review, meta-analysis, and GRADE assessment. *J Clin Sleep Med*. 2019;15(2):301-334. doi:10.5664/jcsm.7638

3. Tregear S, Reston J, Schoelles K, Phillips B. Obstructive sleep apnea and risk of motor vehicle crash: systematic review and meta-analysis. *J Clin Sleep Med*. 2009;5(6):573-581. doi:10.5664/jcsm.27662

4. Eckert DJ, White DP, Jordan AS, Malhotra A, Wellman A. Defining phenotypic causes of obstructive sleep apnea. Identification of novel therapeutic targets. *Am J Respir Crit Care Med*. 2013;188(8):996-1004. doi:10.1164/rccm.201303-0448OC

5. Weaver TE, Grunstein RR. Adherence to continuous positive airway pressure therapy: the challenge to effective treatment. *Proc Am Thorac Soc*. 2008;5(2):173-178. doi:10.1513/pats.200708-119MG

6. Sutherland K, Vanderveken OM, Tsuda H, et al. Oral appliance treatment for obstructive sleep apnea: an update. *J Clin Sleep Med*. 2014;10(2):215-227. doi:10.5664/jcsm.3460

7. 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(10):1396-1407. doi:10.1093/sleep/33.10.1396

8. Browaldh N, Nerfeldt P, Lysdahl M, Bring J, Friberg D. SKUP3 randomised controlled trial: polysomnographic results after uvulopalatopharyngoplasty in selected patients with obstructive sleep apnoea. *Thorax*. 2013;68(9):846-853. doi:10.1136/thoraxjnl-2012-202610

9. Sommer UJ, Heiser C, Gahleitner C, et al. Tonsillectomy with uvulopalatopharyngoplasty in obstructive sleep apnea. *Dtsch Arztebl Int*. 2016;113(1-02):1-8. doi:10.3238/arztebl.2016.0001

10. Woodson BT, Steward DL, Weaver EM, Javaheri S. A randomized trial of temperature-controlled radiofrequency, continuous positive airway pressure, and placebo for obstructive sleep apnea syndrome. *Otolaryngol Head Neck Surg*. 2003;128(6):848-861. doi:10.1016/S0194-5998(03)00461-3

11. Vicini C, Meccariello G, Montevecchi F, et al. Effectiveness of barbed repositioning pharyngoplasty for the treatment of obstructive sleep apnea (OSA): a prospective randomized trial.

Sleep Breath. 2020;24(2):687-694. doi:10.1007/s11325-019-01956-7

12. Carney AS, Antic NA, Catcheside PG, et al. Sleep Apnea Multilevel Surgery (SAMS) trial protocol: a multicenter randomized clinical trial of upper airway surgery for patients with obstructive sleep apnea who have failed continuous positive airway pressure. *Sleep*. 2019;42(6):1-8. doi:10.1093/sleep/zsz056

13. Saghaei M, Saghaei S. Implementation of an open-source customizable minimization program for allocation of patients to parallel groups in clinical trials. *J Biomedical Science and Engineering* 2011;4(11):6. doi:10.4236/jbise.2011.411090

14. MacKay SG, Carney AS, Woods C, et al. Modified uvulopalatopharyngoplasty and coblation channeling of the tongue for obstructive sleep apnea: a multi-centre Australian trial. *J Clin Sleep Med*. 2013;9(2):117-124. doi:10.5664/jcsm.2402

15. Friedman M, Ibrahim H, Joseph NJ. Staging of obstructive sleep apnea/hypopnea syndrome: a guide to appropriate treatment. *Laryngoscope*. 2004;114(3):454-459. doi:10.1097/00005537-200403000-00013

16. *Note for Guidance on Clinical Safety Data Management: Definitions and Standards for Expedited Reporting (CPMP/ICH/377/95)*. Therapeutic Goods Administration; 2000:13.

17. Iber C, Ancoli-Israel S, Chesson A, Quan SF. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. American Academy of Sleep Medicine; 2007.

18. Littner MR, Kushida C, Wise M, et al; Standards of Practice Committee of the American Academy of Sleep Medicine. Practice parameters for clinical use of the multiple sleep latency test and the maintenance of wakefulness test. *Sleep*. 2005;28(1):113-121. doi:10.1093/sleep/28.1.113

19. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*. 1991;14(6):540-545. doi:10.1093/sleep/14.6.540

20. Crook S, Sievi NA, Bloch KE, et al. Minimum important difference of the Epworth Sleepiness Scale in obstructive sleep apnoea: estimation from three randomised controlled trials. *Thorax*. 2019;74(4):390-396.

21. Boulos MI, Jairam T, Kendzerska T, Im J, Mekhael A, Murray BJ. Normal polysomnography parameters in healthy adults: a systematic review and meta-analysis. *Lancet Respir Med*. 2019;7(6):533-543. doi:10.1016/S2213-2600(19)30057-8

22. Lim PV, Curry AR. A new method for evaluating and reporting the severity of snoring. *J Laryngol Otol*. 1999;113(4):336-340. doi:10.1017/S0022215100143919

23. Weaver TE, Laizner AM, Evans LK, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep*. 1997;20(10):835-843.

24. Billings ME, Rosen CL, Auckley D, et al. Psychometric performance and responsiveness of the functional outcomes of sleep questionnaire and sleep apnea quality of life instrument in a randomized trial: the HomePAP study. *Sleep*. 2014;37(12):2017-2024. doi:10.5665/sleep.4262

25. van Hout B, Janssen MF, Feng YS, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health*. 2012;15(5):708-715. doi:10.1016/j.jval.2012.02.008

26. *EQ-5D-5L User Guide*. EuroQol Research Foundation;2019. Accessed April 1, 2020. https://euroqol.org/publications/user-guides

27. Coretti S, Ruggeri M, McNamee P. The minimum clinically important difference for EQ-5D index: a critical review. *Expert Rev Pharmacoecon Outcomes Res*. 2014;14(2):221-233. doi:10.1586/14737167.2014.894462

28. Mulhern B, O'Gorman H, Rotherham N, Brazier J. Comparing the measurement equivalence of EQ-5D-5L across different modes of administration. *Health Qual Life Outcomes*. 2015;13:191. doi:10.1186/s12955-015-0382-6

29. Nolan CM, Longworth L, Lord J, et al. The EQ-5D-5L health status questionnaire in COPD: validity, responsiveness and minimum important difference. *Thorax*. 2016;71(6):493-500. doi:10.1136/thoraxjnl-2015-207782

30. Robinson K, Gatehouse S, Browning GG. Measuring patient benefit from otorhinolaryngological surgery and therapy. *Ann Otol Rhinol Laryngol*. 1996;105(6):415-422. doi:10.1177/000348949610500601

31. Barrett B, Brown D, Mundt M, Brown R. Sufficiently important difference: expanding the framework of clinical significance. *Med Decis Making*. 2005;25(3):250-261. doi:10.1177/0272989X05276863

32. Li Y, Zhang J, Lei F, Liu H, Li Z, Tang X. Self-evaluated and close relative-evaluated Epworth Sleepiness Scale vs. multiple sleep latency test in patients with obstructive sleep apnea. *J Clin Sleep Med*. 2014;10(2):171-176. doi:10.5664/jcsm.3446

33. Kim AM, Keenan BT, Jackson N, et al. Tongue fat and its relationship to obstructive sleep apnea. *Sleep*. 2014;37(10):1639-1648. doi:10.5665/sleep.4072

34. Kezirian EJ, Weaver EM, Yueh B, et al. Incidence of serious complications after uvulopalatopharyngoplasty. *Laryngoscope*. 2004;114(3):450-453. doi:10.1097/00005537-200403000-00012

35. Ravesloot MJL, de Vries N. Reliable calculation of the efficacy of non-surgical and surgical treatment of obstructive sleep apnea revisited. *Sleep*. 2011;34(1):105-110. doi:10.1093/sleep/34.1.105

36. Zhu Y, Long H, Jian F, et al. The effectiveness of oral appliances for obstructive sleep apnea syndrome: a meta-analysis. *J Dent*. 2015;43(12):1394-1402. doi:10.1016/j.jdent.2015.10.008

37. Strollo PJ Jr, Soose RJ, Maurer JT, et al; STAR Trial Group. Upper-airway stimulation for obstructive sleep apnea. *N Engl J Med*. 2014;370(2):139-149. doi:10.1056/NEJMoa1308659

38. Mortimore IL, Bradley PA, Murray JA, Douglas NJ. Uvulopalatopharyngoplasty may compromise nasal CPAP therapy in sleep apnea syndrome. *Am J Respir Crit Care Med*. 1996;154(6 Pt 1):1759-1762. doi:10.1164/ajrccm.154.6.8970367

39. Khan A, Ramar K, Maddirala S, Friedman O, Pallanch JF, Olson EJ. Uvulopalatopharyngoplasty in the management of obstructive sleep apnea: the mayo clinic experience. *Mayo Clin Proc*. 2009;84(9):795-800. doi:10.4065/84.9.795

40. Schriger DL. Graphic portrayal of studies with paired data: a tutorial. *Ann Emerg Med*. 2018;71(2):239-246. doi:10.1016/j.annemergmed.2017.05.033