National Institute for Health and Care Excellence

Final

Obstructive sleep apnoea/hypopnoea syndrome and obesity hypoventilation syndrome in over 16s

Evidence review J: Surgery

NICE guideline NG202

Intervention evidence review

August 2021

Final

Developed by the National Guideline Centre, hosted by the Royal College of Physicians

Disclaimer

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and, where appropriate, their carer or guardian.

Local commissioners and providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

ISBN: 978-1-4731-4229-9

Contents

1 Surgery

1.1 Review question: What is the clinical and cost effectiveness of upper airway surgical interventions for people with obstructive sleep apnoea/hypopnoea syndrome (OSAHS)?

1.2 Introduction

Management of OSAHS may require input from ENT (Ear, Nose and Throat) and maxillofacial surgical specialities, particularly when conventional treatments such as CPAP or oral devices have failed or if a patient has large tonsils. ENT surgeons can evaluate the upper airway and identify any obvious upper airway obstruction.

Indications for surgical intervention may be two fold: to perform curative upper airway surgery to alleviate the need for CPAP or oral devices; or secondly, as adjunct therapy, whereby the surgery improves the upper airway dimensions to facilitate CPAP or oral device use and thus improve treatment adherence and efficacy.

Broadly speaking, surgery can be divided into three categories:

a. Soft tissue surgery (nasal surgery, tonsillectomy, palatal surgery, (tonsillectomy and palatal surgery known as oropharyngeal surgery), tongue-base surgery)
b. Skeletal framework surgery (e.g. bilateral maxilla-mandibular advancement, mandibular osteotomy).
c. Other (e.g. hypoglossal nerve stimulation)

Historically, the most commonly performed ENT procedure is palatoplasty and its numerous variants. Developments in surgical tools such as lasers, radiofrequency technology and robotic surgery have added to the surgical options. Other advances include the use of hypoglossal nerve stimulation for which long-term data in carefully selected patients has been encouraging, although at present this treatment option is not widely used in the UK.

Previously the benefit of surgery has been difficult to evaluate due to a lack of randomised controlled trials and the variation in surgical techniques used in different studies. It has therefore not been regarded as a successful treatment option for many patients. Research has now progressed to include randomised controlled trials, longer follow up and objective documentation of outcomes. The aim of this review was to re-evaluate the benefits of the available surgical interventions for people with OSAHS.

1.3 PICO table

For full details see the review protocol in appendix A.

Table 1: PICO characteristics of review question

Population	Inclusion: People (16 and older) with obstructive sleep apnoea/hypopnoea syndrome (OSAHS) Population will be stratified by: • severity- mild, moderate, severe (based on AHI Apnoea–Hypopnoea Index / ODI Oxygen desaturation index) • treatment stage – failed previous OSAHS treatment vs general population • site of obstruction – multilevel vs single level

Interventions	Any specific surgical intervention for OSAHS (nasal surgeries, uvulopalatopharyngoplasty, tonsillectomy, palatal implant, tongue reduction, genioglossus advancement, radiofrequency ablation, maxillomandibular advancement, hyoid suspension, upper airway stimulation).
Comparisons	• Any non-surgical intervention (positive airway pressure devices, positional modifiers, oral devices, • No intervention/usual care as defined in the studies (including lifestyle advice etc)
Outcomes	**Critical** • generic or disease specific validated quality of life measures (continuous) • mortality (dichotomous) **Important** • sleepiness scores (continuous, e.g. Epworth Sleepiness Scale (ESS)) • apnoea-Hypopnoea index (continuous) • oxygen desaturation index (continuous) • CO_2 control (continuous) • permanent adverse effects (e.g. neural dysfunction, open nasality, globus sensation, dichotomous) • reversible adverse effects (e.g. pain, infection, secondary bleeding, dichotomous) • driving outcomes (continuous) • neurocognitive outcomes (continuous) • impact on co-existing conditions: ○ HbA1c for diabetes (continuous) ○ cardiovascular events for cardiovascular disease (dichotomous) ○ systolic blood pressure for hypertension (continuous)
Study design	• RCTs • systematic review of RCTs Minimum duration of follow-up 1 month

1.4 Clinical evidence

1.4.1 Included studies

Eleven studies (15 papers) were included in the review[9, 23, 27, 38, 42, 45, 47, 70, 78, 82, 83, 85, 87-89]. Evidence from these studies is summarised in the clinical evidence table below (Table 3).

Overall, participants had moderate to severe sleep apnoea in the trials. Two studies assessed the effects of uvulopalatopharyngoplasty (UPPP) (Lojander 1996, Tegelberg 1999)[45, 78]. One study assessed the effects of multilevel surgery (modified uvulopalatopharyngoplasty and minimally invasive tongue volume reduction in adults with symptomatic moderate or severe OSA in whom conventional treatments had failed (MacKay, 2020)[47];One study assessed the effects of barbed repositioning pharyngoplasty (BRP) (Vicini 2020)[83]. One study (Ferguson 2003)[23] assessed the effects of laser assisted uvulopalatoplasty (LAUP). One study (Woodson 2003)[89] performed temperature-controlled radiofrequency tissue ablation (TCRFTA). One study (Vicini 2010)[82] performed maxillomandibular advancement [MMA]. Two studies (Joar 2018, Sommer 2016)[38, 70]

performed tonsillectomy with uvulopalatopharyngoplasty (TE-UPPP). One study (Koutsourelakis 2008)[42] performed sub mucous resection of the deviated nasal septum. One study (Friedman 2008)[27] performed surgically inserted palatal implants.

The following comparators were used in the studies: no treatment initially/delayed surgery in four studies (5 papers) (Browaldh 2013, Joar 2018, Sommer 2016, Ferguson 2003,Vicini 2020)[9, 23, 38, 45, 70, 83]; sham procedure in three studies, placebo implants in Friedman 2008[27] and sham surgery in Koutsourelakis 2008[42] and (Woodson 2003)[89]; conservative management in two studies (Lojander 1996, MacKay, 2020)[45, 47]; oral devices in one study (4 papers) (Wilhelmsson 1999, Tegelberg 1999, Walker-Engstrom 2002 and Walker-Engstrom 2000)[78, 85, 87, 88]; APAP in one study (Vicini, 2010)[82]; and CPAP in one study (Woodson 2003).[89] The study Woodson 2003 had 3 arms – surgery, sham surgery and CPAP.

One study (Sommer 2016) included patients with OSAHS and CPAP intolerance. One study (Joar 2018) included patients who were non-adherent to CPAP and mandibular repositioning devices (MRD) treatments. One study (Koutsourelakis 2008) excluded patients from treatment of OSAHS with continuous positive airway pressure (CPAP) during the course of the study. One study (Woodson 2003) included people with no prior surgical or CPAP treatment. One study (MacKay 2020) included participants in whom medically supervised attempts to use CPAP and, when deemed appropriate, a mandibular advancement device, failed or were refused.

Studies comparing surgery with sham surgery/no surgery/delayed surgery/conservative management have been combined in the analysis.

Studies were broadly categorised as follows: Nasal surgery (to include septal surgery and turbinate surgery); Oro-pharyngeal surgery – this includes tonsillectomy on its own or combined with any palatal surgery, UPPP, any form of palatoplasty, expansion sphincter palatoplasty, laser or radiofrequency palate surgery; trans-oral robotic surgery for tongue and/or epiglottis; hypoglossal nerve stimulation/upper airway stimulation; skeletal framework surgery - maxillomandibular advancement.

The participants in the studies had mixed levels of AHI. Studies were stratified based on the AHI/ODI severity of the population. When a mixed severity population was included the severity of the majority of the population was used by taking the mean AHI of the patients included and the study was downgraded for indirectness.

The included studies used the following airway obstruction inclusion criteria**:** Participants in Friedman 2008 were required to have soft palate implants; Woodson 2003 excluded participants if they had tonsillar hypertrophy or nasal/supraglottic obstruction on examination; Tegelberg 1999 excluded participants with significant nasal obstruction; Lojander 1996 considered appropriate for surgery participants with more than 50% obstruction at palatal level in the Mueller manoeuvre, with or without obstruction at the epiglottic level; Ferguson 2003 and Vicini 2010 did not stipulate airway obstruction as an entry criterion; Joar 2018 included participants with Friedman stage I or II; Koutsourelakis 2008 included participants with nasal septum deviation with or without inferior turbinate hypertrophy, as assessed by clinical examination and flexible fibre optic nasopharyngoscopy along with nasal resistance values exceeding normal limits at baseline (symptomatic fixed nasal obstruction)**;** Sommer 2016 included participants with tonsillar hypertrophy with velopharyngeal obstruction confirmed by clinical examination; Vicini 2020 included patients with certain degree of nasal obstruction planned for BRP and tonsillectomy, with nasal surgery (septoturbinoplasty).

Follow-up in the studies ranged from 2 months to 4 years.

There was no evidence available for the following outcomes: permanent adverse effects, CO2 control, driving outcomes, neurocognitive outcomes and impact on co-existing conditions such as HbA1c for diabetes, cardiovascular events for cardiovascular disease, and systolic blood pressure for hypertension.

See also the study selection flow chart in Appendix C:, study evidence tables in Appendix D:, forest plots in appendix E and GRADE tables in appendix F.

1.4.2 Excluded studies

See the excluded studies list in appendix I.

1.4.3 Summary of clinical studies included in the evidence review

Table 2: Clinical studies included in the evidence review

Study	Intervention and comparison	Population	Outcomes	Comments
Ferguson 2003[23] RCT Canada	Laser-assisted uvulopalatoplasty (LAUP) group: The LAUP procedure was repeated at 1- to 2-month intervals. The end points for the LAUP procedure were (*1*) when the snoring was significantly reduced or eliminated, (*2*) no more tissue could be safely removed, or (*3*) the patient refused further surgery. Control: The control subjects were not offered any therapy but were offered LAUP at the end of the study Control group were offered surgery after 6 months	N=45 subjects who had mild OSA (AHI 10-27 per hour) and complained of loud snoring. LAUP: 21; No treatment: 24 Mean age: 44.6 (SD 8.1) BMI 36 kg/m^2 (SD 4.5) AHI: LAUP: 18.6 (SD 4.3); Control 16.1 (SD 4) ESS: LAUP: 10.7 (SD 3.7); Control: 10 (SD 5.2) One participant in each group had previously had CPAP.	• AHI • ESS • Dysphagia • Infection • Bleeding • Pain • Nasal regurgitation • Withdarwals • Quality of life	Moderate OSAHS strata based on mean AHI Questionnaires, scales, and the polysomnogram were repeated 3 months after the last LAUP procedure or 6 months after baseline in the control group. Category of surgery: oro-pharyngeal surgery
Friedman 2008[27] RCT Taiwan	N=31 Surgically inserted palatal implants The palatal implant insertion tools provided by the manufacturer for the placebo control group did not include the palatal implants, but they were in all other aspects	N = 62. Palatal implants: 31; placebo: 31. BMI: 29 kg/m^2 AHI: 22 Inclusion criteria: History of OSAHS and/or symptoms of OSAHS (mainly snoring & excessive daytime	AHI ESS Quality of life	Moderate OSAHS strata based on mean AHI Outcomes assessed 12 weeks post-operatively. Category of surgery: oropharyngeal surgery

Study	Intervention and comparison	Population	Outcomes	Comments
	identical to the implant insertion tools used in the treatment group receiving the implant. Group assignment associated with insertion tools was not distinguishable by study participants and investigators because the implants are deployed from within the hollow needle of a delivery tool. The devices were all identified by a lot number before distribution. All patients underwent a preoperative mouth rinse with chlorhexidine and received a 5-day postoperative course of prophylactic antibiotics. versus N =31 Sham procedure (active and placebo implants were indistinguishable).	sleepiness); Friedman tongue position (FTP) I, II, or III; diagnosis of mild or moderate OSAHS (AHI 5-40); a soft palate 2 cm, but less than 3.5 cm; BMI 32 kg/m^2. Exclusion criteria: Clinical and physiological presentation of severe OSAHS (ESS 20, frequent choking and gasping during sleep; OSAHS (AHI 40)); unwilling to be randomly assigned to placebo; FTP IV; tonsil size 3 or 4; classified stage IV of Friedman staging system. Pre-operative AHI: Palatal implants: 23.8 (5.5) Placebo: 20.1 (5.4)		
Joar 2018[38] Browaldh 2013[9] RCT Sweden	N = 32 modified UPPP, including tonsillectomy, within 1 month A modification of the method initially described by Fujita, required only minor resections of the soft palate and uvula using the cold steel technique4 and suturing of the tonsillar	N=65 All OSA patients referred to the Ear, Nose, and Throat Department of the Karolinska University Hospital, Stockholm, Sweden from June 2007 to May 2011 for UPPP were eligible for this single-centre study. All patients underwent clinical investigations by ENT specialists, with fibre-	FOSQ AHI	Severe OSAHS strata based on mean AHI 6- and 24-months follow-up Category of surgery: oropharyngeal surgery

Study	Intervention and comparison	Population	Outcomes	Comments
	pillars including the palatopharyngeal muscle. All surgeons used the same technique. Versus N = 33 no treatment for 6 months (control group) before surgery.	endoscopy of the upper airways. Patients who were considered suitable (no other obvious anatomical abnormality), and willing to undergo pharyngeal surgery, were asked to participate. The following inclusion criteria were used: males and females > 18 years of age; AHI ≥ 15 events/hour of sleep (from PSG); ESS score ≥ 8; excessive daytime sleepiness three times a week or more; BMI of less than 36 kg/m^2; Friedman stage I or II (includes Friedman tongue position (FTP), tonsil size, and BMI); and non-adherence with CPAP and MRD treatments, with the exception of patients with Friedman stage I and BMI of less than 30 kg/m^2. The following exclusion criteria were used: serious psychiatric, cardiopulmonary, or neurological disease or an American Society of Anaesthesiologists (ASA) classification of >3; patients who decline surgery; insufficient knowledge of		

Study	Intervention and comparison	Population	Outcomes	Comments
		Swedish language to complete questionnaires; nightshift workers; patients who could be dangerous in traffic according to responses in a non-standardised questionnaire; severe nasal congestion (could be included after topical nasal treatment); previous tonsillectomy (as such patients were considered partially treated); Friedman stage III; and severe clinical worsening of OSA during the study. AHI: surgery: 53.3 (19.7); control: 52.6 (21.7)		
Koutsourelakis 2008[42] RCT Greece	N=27 Surgery group All patients underwent sub mucous resection of the deviated nasal septum. In 18 out of 27 patients, sub mucous resection of the bilateral inferior turbinates was also performed. Nasal packing was removed on the second post-operative day, and routine saline nasal irrigation and debridement were performed. Versus	A total of 51 consecutive subjects who referred to the Centre of Sleep Disorders of the ‘‘Evangelismos’’ General Hospital of Athens, Greece for suspected sleep-disordered breathing were recruited. Enrolment criteria were: 1) nasal septum deviation with or without inferior turbinate hypertrophy, as assessed by clinical examination and flexible fibreoptic nasopharyngoscopy along with nasal resistance values exceeding normal limits at	AHI ESS	Severe OSAHS strata based on mean AHI. Follow-up 3-4 months after surgery. Category of surgery- nasal surgery

Study	Intervention and comparison	Population	Outcomes	Comments
	N=22 Placebo group (sham surgery) To ensure blinding, a standard submucosal resection of the nasal septum was simulated. After the infiltration of the nasal septum with 10 mL lidocaine 1% containing epinephrine 1:200,000, the surgeon asked for all instruments and manipulated the nose as if submucosal resection was being performed. Patients remained in the operating room for the same amount of time required for the surgery group. Patients spent the night after the procedure in the hospital and were cared for by nurses who were unaware of the treatment group assignment. Nasal packing was removed on the second post-operative day and routine saline nasal irrigation and debridement were performed.	baseline (symptomatic fixed nasal obstruction); 2) AHI .5 events/h-1 at baseline; 3) no upper or lower respiratory tract disease, including a history of nasal allergy; 4) no recent surgery involving the upper airways; 5) no use of medications known to influence nasal resistance (antihistamine, decongestants, etc.); and 6) no history of neuromuscular or cardiovascular disease. Exclusion criterion was the treatment of OSA with continuous positive airway pressure (CPAP) during the course of the study. Mean AHI: surgery 31.5 (16.7); control 30.6 (13.8)		
Lojander 1996[45] RCT Finland	Surgery: N=18 Patients with more than 50% obstruction at the palatal level in the Mueller manoeuvre but less than 50% obstruction at the epiglottic level were considered to be suitable for UPPP alone. Mandibular osteotomy with hyoid myotomy suspension was performed	N = 32 adults with moderate to severe OSAS. Diagnosis confirmed by sleep study. Age range 27-65 BMI: 24-41 kg/m^2 ODI4 in control -median (range):34 (20-68)	ODI Dysphagia Tracheotomy Re-operations	Severe OSAHS strata based on mean ODI Participants were assessed by a team of specialists (including physicians and surgeons). Follow-up: 3 months and 1 year

Study	Intervention and comparison	Population	Outcomes	Comments
	together with UPPP, if the patient had a narrow posterior airspace, an inferiorly positioned hyoid, and a sharp sella nose mandibular angle. It was the policy of the study to operate only on patients with moderate to severe OSAHS i.e. patients with more than 20 desaturations of 4% or more per hour (ODI 4>20). UPPP was performed according to the method of Fujita. Mandibular osteotomy was performed according to the method of Powell et al. Versus N=14 conservative management conservative management consists of weight loss and positional therapy as well as avoidance of tranquilizers and alcohol at bedtime. Study duration: 1 year.	ODI4 in surgery group- median (range):45 (21-72) Participants with more than 50% obstruction at palatal level in the Mueller's manoeuvre, and those with or without obstruction at the epiglottic level were considered appropriate for surgery. Inclusion criteria: confirmed diagnosis of OSAS, periodic breathing pattern in both static charge sensitive bed and thermistor channels. Participants with BMI more than 40 kg/m^2 were excluded. Patients COPD/ asthma, other serious concomitant illnesses, and participants where somnolence would cause risk or incapacity to work were excluded.		Category of surgery: oropharyngeal surgery
MacKay 2020[47] RCT Australia	(n=51) Intervention 1: Surgery. The surgery intervention consisted of a modified uvulopalatopharyngoplasty to widen and stabilize the velopharynx and 7 to 9 submucosal insertions of a	n=102 Adults with symptomatic moderate or severe OSA in whom conventional treatments had failed were enrolled from August 2014 to	AHI ESS FOSQ Serious adverse events 24 h ambulatory systolic blood pressure	Severe OSAHS strata based on mean AHI Follow-up: 6 months

Study	Intervention and comparison	Population	Outcomes	Comments
	radiofrequency-in-saline wand to reduce tongue volume. A training workshop was conducted to standardise the surgical techniques among the 7 participating surgeons. (n=51) Intervention 2: Medical management. 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.	November 2017, with follow-up until August 2018. Eligible adults were aged 18 to 70 years with moderate or severe OSA (defined as apnoea-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 Age - Mean (SD): surgery group: 42.7 (12.8); control group - 46.4 (12.6) Gender (Men): surgery group: 41 (80%); control group : 43 (84%) BMI: Men: surgery: 30.1 (4.0); control: 30.0 (3.6); Women: surgery: 33.3 (2.8); control: 26.6 (2.9) mean AHI- surgery: 47.9 (23.1); Control: 45.3 (23.9) Epworth Sleepiness Scale, mean (SD): surgery: 12.4 (3.6); control: 11.1 (4.7)	24 h ambulatory diastolic blood pressure	Category of surgery: oropharyngeal surgery

Study	Intervention and comparison	Population	Outcomes	Comments
		Previous OSA treatment: Tried CPAP: surgery: 38 (75); control: 37 (73) Refused CPAP: surgery: 13 (25); control:14 (27) Tried mandibular advancement device: surgery:16 (31); control: 12 (24)		
Sommer 2016[70] Germany RCT	Surgery Patients in the treatment arm underwent tonsillectomy with uvulopalatopharyngoplasty (TE-UPPP) within one month after inclusion. After cold steel tonsillectomy using general anaesthesia, uvulopalatopharyngoplasty according to the modifications by Pirsig was performed. Tonsil size was determined immediately following surgery using volume displacement. Complications occurring during inpatient stay, particularly haemorrhages were recorded by type and severity. versus Control	N= 42 (23 in the treatment group, 19 in the control group). Patients of both sexes aged between 18 and 65 years, were enrolled between 2010 and 2014. Inclusion criteria were obstructive sleep apnoea confirmed by polysomnography (PSG) with AHI above 15, according to the second edition of the International Classification of Sleep Disorders valid at that time and tonsillar hypertrophy with velopharyngeal obstruction confirmed by clinical examination. A further very important inclusion criterion was rejection of or poor compliance with ventilation therapy and an explicit wish	AHI ESS SpO2	Severe OSAHS strata based on mean AHI Follow-up- 3 months Category of surgery: oropharyngeal surgery

Study	Intervention and comparison	Population	Outcomes	Comments
	Patients in the control arm initially received no treatment and underwent repeat polysomnography again after three months, then underwent TE-UPPP.	on the part of the patient for a different approach (second-line therapy). All enrolled patients had tried CPAP treatment without success for at least one night. Because of the large number of patients who coped well with CPAP treatment, it is not possible to draw a CONSORT diagram for this trial. The most important exclusion criteria were body mass index above 34 kg/m2, increased anaesthetic risk according to the criteria of the American Society of Anaesthesiologists (ASA), specifically ASA class above III, and other relevant types of obstruction or significant malformations of the facial skeleton confirmed by clinical examination. AHI : control 35.7 ± 19.4; surgery 33.7 ± 14.5		
Vicini 2010[82] RCT Italy	N=25 Maxillomandibular advancement [MMA] All the surgical procedures were performed under general anaesthesia after routine fibreoptic orotracheal intubation.	Fifty patients with PSG classified as severe OSAHS (AHI N30) were prospectively enrolled. Inclusion criteria included the presence of severe OSAHS (AHI >30) regardless of BMI	AHI ESS Complications of surgery and APAP	Severe OSAHS strata based on mean AHI Mean follow-up was 13 ± 2.5 SD months

Study	Intervention and comparison	Population	Outcomes	Comments
	Temporary tracheotomy was routinely carried out to avoid possible episodes of dyspnoea in the first 24 to 48 hours after surgery and to facilitate the possibility of having to suction mucous secretion. As a first step, a sagittal split ramus mandibular osteotomy according to Obwegeser-Dal Pont was performed with a powered reciprocating saw and a Lindemann cutting burr (in the ramus inner cortex area). The fixed amount of advancement, 11 mm for all the cases, was checked by means of a customized intermediated splint. To stabilise the achieved advancement, we inserted 3 to 4 bicortical screws. In 3 cases, a titanium plate was added on each side to improve stabilization. As a second step, a low Le Fort I maxillary osteotomy was carried out step by step using a powered reciprocating saw and different kinds of special osteotomies. The final position of the maxilla was stabilised by 4 titanium screwed plates. The surgical team leader for all the procedures was always the same assisted by 2 surgeons. All the patients were postoperatively managed in the	(which was usually abnormally high) and no formal contraindication for surgery according to the Stanford protocol (pre-existing local and general medical conditions that could increase the risk of surgery or might compromise the final outcome, fear of surgery, concern over pain and discomfort, loss of work or income during convalescence, advancing age) and no formal contraindication for APAP Chronic Obstructive Pulmonary Disease (COPD), heart dysrhythmia, heart failure, restrictive lung disease, neuromuscular disease, previous surgery for SDB). Mean AHI was 56. 8 ± 16.5 SD,		Category of surgery: Skeletal Framework Surgery

Study	Intervention and comparison	Population	Outcomes	Comments
	ENT ward with continuous monitoring of pulse rate, blood pressure, and pO2 during the first 24 hours. Elastomeric release of morphine hydrochloride was the routine choice for pain relief. The rigid intermaxillary fixation was removed after 24 hours, and oral intake of food was immediately encouraged. The tracheotomy was removed usually on the fourth/fifth day. Discharge was possible within 1 week for all the patients. versus N=25 ventilatory treatment modality (auto titrating positive airway pressure [APAP]). The patients enrolled in the conservative section of the present study were submitted to automatic APAP application with a nasal mask, held in position with an elastic headgear and attached to a flow generator by elephant tubing. This APAP was able to detect 3 different parameters for pressure auto titration: (1)			

Study	Intervention and comparison	Population	Outcomes	Comments
	forced oscillation, (2) flow limitation, (3) snoring. The patients were requested to continue using the Auto-CPAP only after a successful test period, usually 1 week, checked by interview and smart card evaluation (the smart card records the true time of utilization and different operative parameters). All the patients were routinely recalled every 3 weeks to check the effectiveness and the use of the Auto-CPAP.			
Vicini 2020[83] RCT Italy	(n=25) Intervention 1: Surgery. Barbed repositioning pharyngoplasty (BRP) After bilateral tonsillectomy meticulous sparing of the palatoglossus and palatopharyngeus muscles was performed. Two weakening or releasing partial incisions were done by a pinpoint bowie (Colorado) at the inferior (caudal) part of the palatopharyngeal muscle. The centre of the palate was marked at palatal spine, and the pterygomandibular raphe in both sides were located by digital palpation and marked. Single barbed suture, bidirectional polydioxanone	Inclusion criteria: Patients suffering from moderate to severe OSA (AHI ≥15 events/h) with certain degree of nasal obstruction planned for BRP and tonsillectomy, with nasal surgery (septoturbinoplasty), Grades 1–2 tonsillar hypertrophy, aged between 18 and 65 years old, BMI ≤ 35, failure of CPAP or low adherence to this treatment during the last 3 months (< 4 h per night), mainly palatal/pharyngeal collapses at DISE (severe circular palatal collapses and severe transversal pharyngeal collapses with	AHI ESS	Severe OSAHS strata based on mean AHI Follow-up- 6 months Category of surgery: oropharyngeal surgery

Study	Intervention and comparison	Population	Outcomes	Comments
	absorbable monofilament, size 2.0, with transition zone in the middle was generally used. One needle was introduced at the centre point then passed laterally within the palate, turning around pterygomandibular raphe till it came out at the most superior part of the raphe at one side; the thread was pulled until it hung at the central transition. The needle again was re-introduced close to point of exit, passing around the pterygomandibular raphe, till it came out into the tonsillectomy bed, then through the upper part of the palatopharyngeus muscle and came out near to mucosa of posterior pillar not through it. The posterior pillar was entered at the junction between the upper third and the lower two-thirds. Then, again, the needle was passed back through the tonsillectomy, bed and then this suture would be suspended around the raphe again. The opposite side was done by the same way. Finally, each thread came out at the raphe of the same side, for locking of the stitches and looseness prevention; a superficial stitch in the opposite direction was taken, and then,	none or mild tongue collapses). Age – Mean yrs: surgery- 44.64; control- 50. Gender (M:F): surgery- 22/3; control- 20/1. mean AHI-: surgery 25.58 ± 14.60; control: 36.83 ± 23.82. ESS: surgery: 9.28 ± 3.10 ; control: 10.4 ± 23.68		

Study	Intervention and comparison	Population	Outcomes	Comments
	the thread was cut while pushing the tissue downward for more traction (n=25) Intervention 2: No intervention Observation. No further information.			
Wilhelmsson 1999[88] Tegelberg 1999[78] Walker-Engstrom 2000[87] Walker-Engstrom 2002[85] RCT Sweden	Participants were randomised to either oral appliance or surgical intervention (uvulopalatopharyngoplasty). Participants randomised to receive UPPP were followed up at regular intervals. Oral devices – before the intervention a clinical examination of the stomatognathic system was carried out. The same dentist treated all patients and one dental technician was responsible for the manufacture of the dental appliances. The appliances were carefully designed and fabricated on dental casts of acrylic polymer at a dental laboratory. The appliances were used at night times only and advanced the mandible by 50% of the patient's maximum protrusive capacity. Each patient was given an appointment for	95 male participants were recruited. Age: 20-65 Baseline AHI: oral device: 18.2 (15.7-20.8); surgery: 20.4 (17.44-23.3). Inclusion criteria: AHI between 5 and 25, age between 20 and 65. Exclusion criteria: Mental illness, drug misuse, significant nasal obstruction, insufficient teeth, pronounced dental malocclusion, severe cardiovascular disease, neurological disease, respiratory disease. At 4-year follow-up, OA group: N = 32, UPPP group: N = 40	AHI Quality of life Oxygen desaturation index Dysphagia Withdrawals Nasal regurgitation	Moderate OSAHS strata based on mean AHI Comparison of surgery versus oral appliance. Follow-up 1 year Category of surgery: oro-pharyngeal surgery

Study	Intervention and comparison	Population	Outcomes	Comments
	adjustment and adaptation of the device. N=49 Surgery – The Uvulopalatopharyngoplasty (UPPP) was performed by the same ear, nose and throat surgeon using a standardised procedure described by Frjita. The procedure involved tonsillectomy regardless of the size of the tonsils, and resection of excess fat and mucosa of the soft palate, including the uvula. The palpable musculature was saved, and several sutures approximated the anterior and posterior tonsillar pillars. The UPPP surgery was performed under general anaesthesia. N=46 Study duration: 1 year.			
Woodson 2003[89] RCT USA	N=30 TCRFTA (radiofrequency energy delivered to create non-overlapping lesions in two/three tongue sites, occurring at 4-week intervals. Data recorded after last treatment session. Palate sessions also included) Active temperature-controlled radiofrequency tissue ablation	N = 90 (CPAP: 30 and TCRFTA: 30; placebo: 30); Mean age: Placebo: 46 yrs; TCRFTA: 49.4 yrs; CPAP 51.7 yrs BMI: Placebo: 28.5 kg/m^2 (4.2); TCRFTA: 27.7 kg/m^2 (3.6);	CPAP machine usage FOSQ ESS AHI	Moderate OSAHS strata based on mean AHI CPAP outcome data reported at 8 weeks. Category of surgery: oro-pharyngeal surgery

Study	Intervention and comparison	Population	Outcomes	Comments
	(TCRFTA) was performed with the Somnoplasty radiofrequency generator (Gyrus- ENT, Memphis, TN). Five tongue and 2 palate sessions were planned for each active subject. Subjects were treated perioperatively with oral antibiotics, prednisone, antiseptic oral rinse, analgesic (as needed), and nonsteroidal anti-inflammatory medication (as needed). A local anaesthetic mixture (2.5 mL of 2% lidocaine with 1:100,000 epinephrine, 2.0 mL of normal saline, and 0.5 mL of 8.4% sodium bicarbonate) was injected into each tongue treatment site, and 1% lidocaine with 1:100,000 epinephrine (1 to 2 mL) was injected into each palate site. Radiofrequency energy was delivered to create non-overlapping lesions in 2 or 3 tongue sites (1000 or 750 J, respectively per site; target temperature 85° C; maximum power 10 W) per tongue treatment session, which occurred at 4-week intervals. Radiofrequency energy was delivered to create 1 midline and 2 lateral lesions (non-overlapping) to the soft palate	CPAP 29.1 kg/m^2 (3.7) AHI: TCRFTA: 21.3 (11.1); Placebo: 15.4 (7.8); CPAP: 19.8 (9.9) ESS: Placebo:11.6 (3.5); TCRFTA: 11.9 (4.6); CPAP: 12.6 (5) FOSQ: Placebo: 16.8 (2.1); TCRFTA: 16.5 (2); CPAP: 16 (2.6) Inclusion criteria: mild to moderate OSA (AHI 10 to 30; age 18-65; self-reported daytime sleepiness; BMI of 34 kg/m^{2}or less; no prior surgical or CPAP treatment Exclusion criteria: co-existing significant sleep disorder; tonsillar hypertrophy; nasal supraglottic obstruction on examination; ASA class IV/V; claustrophobia; latex allergy; pregnancy; major depression; drug/alcohol abuse; history of an accident		

Study	Intervention and comparison	Population	Outcomes	Comments
	(650 J and 325 J, respectively) in each palate treatment session. Investigators were instructed to adjust lesion numbers per treatment session based on clinical judgment and patient tolerance. When tongue and palate sessions were combined, the subject was offered overnight hospital admission. Investigators were instructed to perform sequential and not simultaneous tongue and palate treatments if there were concerns about airway oedema or patient tolerance. Attempts were made to apply similar levels of energy in all patients irrespective of the timing of sessions. N=30 Sham TRCFTA. Sham-placebo TCRFTA was performed as described above for tongue TCRFTA except that a blocking control box on the radiofrequency generator was set to “off” to prevent delivery of energy. Three tongue sessions were planned for each sham-placebo subject at 4-week intervals with 3 tongue lesions created per session. Subjects were anesthetised and	secondary to sleepiness; participation in another study		

Study	Intervention and comparison	Population	Outcomes	Comments
	medicated as described for active tongue TCRFTA. The sham treatment sessions were limited to 3 to balance the risk of hematoma, oedema, or abscess formation at the site of anaesthetic injection or TCRFTA probe insertion versus the goal of providing a realistic placebo N=30 CPAP Nasal CPAP therapy was titrated unattended over 3 or more nights with the AutoSet T device. Final constant CPAP pressure was set as the 95-percentile pressure and was continued for 8 weeks. Subjects were seen at 1, 2, and 4 weeks to troubleshoot and optimise compliance. Both active and sham surgical interventions were compared with nasal CPAP.			

1.4.4 Quality assessment of clinical studies included in the evidence review

Table 3: Clinical evidence summary: Surgery versus conservative management/no surgery/sham surgery- Moderate OSAHS [Category of surgery: oropharyngeal surgery][9]

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Conservative management/no surgery/sham surgery- MODERATE	Risk difference with Surgery versus conservative management/no surgery/sham surgery (95% CI)
AHI(events/hr) Lower is better	152 (3 studies)	⊕⊖⊖⊖ VERY LOW[1,2,8] due to risk of bias, imprecision, indirectness		Mean AHI in control group was 13.9	The mean AHI in the intervention groups was 5.19 lower (7.94 to 2.44 lower)
Epworth sleepiness score (ESS) Scale 0 to 24 Higher is worse	154 (3 studies)	⊕⊕⊖⊖ LOW [1,8] due to risk of bias		Mean ESS in control group was 6.9	The mean ESS in the intervention groups was 1.02 lower (2.16 lower to 0.12 higher)
Vitality (SF-36) Scale from 0-100 Higher is better	62 (1 study)	⊕⊕⊕⊖8 MODERATE due to indirectness		Mean in control group was -3.8	The mean vitality (SF-36) in the intervention groups was 27.4 higher (19.17 to 35.63 higher)
SF- 36- Physical health Scale from 0-100 Higher is better	113 (2 studies)	⊕⊖⊖⊖ VERY LOW[32,,8] due to inconsistency, imprecision, indirectness		Mean in control group was 1.9	The mean SF-36 physical health in the intervention groups was 5.96 higher (5.50 lower to 17.43 higher)
SF-36- mental health Scale from 0-100 Higher is better	113 (2 studies)	⊕⊖⊖⊖ VERY LOW[3,2,8] due to inconsistency,		Mean in control group was 0.28	The mean SF-36 mental health in the intervention groups was 10.50 higher (5.53 lower to 26.53 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Conservative management/no surgery/sham surgery- MODERATE	Risk difference with Surgery versus conservative management/no surgery/sham surgery (95% CI)
		imprecision, indirectness			
FOSQ(Functional Outcome of Sleep Questionnaire) Scale from 5-20 Higher is better	54 (1 study)	⊕⊕⊕⊖8 MODERATE due to indirectness		Mean in control group was 0.4	The mean FOSQ in the intervention groups was 0.8 higher (0.16 lower to 1.76 higher)
Quality of life (sleep apnoea quality of life index) Higher is better	45 (1 study)	⊕⊖⊖⊖ VERY LOW[1,8] due to risk of bias, indirectness		Mean in control group was 4.3	The mean quality of life (SAQLI) in the intervention groups was 0.3 higher (0.41 lower to 1.01 higher)
Dysphagia	45 (1 study)	⊕⊕⊖⊖ VERY LOW[1,8] due to risk of bias, indirectness	OR 9.97 (1.3 to 76.29)5	Moderate	
				0 per 1000	190 more per 1000 (from 13 more to 360 more)[4] (4/21 events in surgery, zero events in control)
Infection	99 (2 studies)	⊕⊕⊖⊖ LOW [1,8] due to risk of bias, indirectness	OR 9.97 (1.3 to 76.29)5	Moderate	
				0 per 1000	80 more per 1000 (from 1 fewer to 170 more)[4] (4/21 events in surgery from one study; second study had zero events in both groups)
Bleeding (mild-severe)	99 (2 studies)	⊕⊕⊖⊖ LOW1,8 due to risk of bias, indirectness	OR 13.72 (3.23 to 58.37)5	Moderate	
				0 per 1000	190 more per 1000 (from 75 more to 300 more)[4] (9/21events in surgery from one study; second study had zero events in both groups)
Pain[6]				Moderate	

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Conservative management/no surgery/sham surgery- MODERATE	Risk difference with Surgery versus conservative management/no surgery/sham surgery (95% CI)
	45 (1 study)	⨁⨀⨀⨀ VERY LOW[1,8] due to risk of bias, indirectness	OR 29 (8.78 to 95.76)5	0 per 1000	800 more per 1000 (from 630 more to 980 more)[4] (17/21 events in surgery)
Nasal regurgitation	45 (1 study)	⨁⨀⨀⨀ VERY LOW[1,8] due to risk of bias, indirectness	OR 10.56 (1.67 to 66.68)5	Moderate	
				0 per 1000	230 more per 1000 (from 40 more to 420 more)[4] (5 /21 events in surgery)
Ulcerations	54 (1 study)	⨁⨁⨁⨀ MODERATE 8 due to indirectness	Not estimable	See comment	See comment (zero events in both groups)
Pain at 1 week (10 cm visual analog scale range 0-10) 7	54 (1 study)	⨁⨁⨀⨀ VERY LOW[2,8] due to imprecision, indirectness		Mean in control group was 1.84	The mean pain at 1 week (10 cm visual analog scale range 0-10) in the intervention groups was 0.2 lower (1.41 lower to 1.01 higher)
Pain at 3 weeks (10 cm visual analog scale range 0-10) 7	54 (1 study)	⨁⨁⨀⨀ LOW [2,8] due to imprecision, indirectness		Mean in control group was 0.33	The mean pain at 3 weeks (10 cm visual analog scale range 0-10) in the intervention groups was 0.38 higher (0.12 lower to 0.88 higher)
Swallowing at 1 week (10 cm visual analog scale range 0-10) 7	54 (1 study)	⨁⨀⨀⨀ VERY LOW[2,8] due to imprecision, indirectness		Mean in control group was 1.73	The mean swallowing at 1 week (10 cm visual analog scale range 0-10) in the intervention groups was 0.41 higher (0.91 lower to 1.73 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Conservative management/no surgery/sham surgery- MODERATE	Risk difference with Surgery versus conservative management/no surgery/sham surgery (95% CI)
Swallowing at 3 weeks (10 cm visual analog scale range 0-10) 7	54 (1 study)	⨁⊖⊖⊖ VERY LOW[2,8] due to imprecision, indirectness		Mean in control group was 0.57	The mean swallowing at 3 weeks (10 cm visual analog scale range 0-10) in the intervention groups was 0.28 higher (0.36 lower to 0.92 higher)
Mortality					Not reported

1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
2 Downgraded by one increment if the confidence interval crossed one MID and downgraded by two increments if the confidence interval crossed both MIDs. Established MIDs for SF-36 physical/mental- 2/3; FOSQ- 2; SAQLI-2; ESS -2.5. Default MID (0.5XSD)used for AHI, swallowing and pain outcomes.
3 I2=95%. Downgraded by 1 or 2 increments for heterogeneity Sub-group analysis could not be conducted due to insufficient number of studies in the comparison. Random effects analysis used.
4 Risk difference calculated in Revman
5 Peto odds ratio used as zero events in one/both groups.
6 moderate to severe pain immediately after the procedure (despite analgesia
7 Pain and swallow side effects 10 cm visual analog scale (pain:0= no pain; and 10=severe pain; swallowing 0= normal swallow; and 10= unable to swallow without pain even with medication.
8 Downgraded by 1 or 2 increments because the majority of the evidence included an indirect population (downgrade by one increment) or a very indirect population (downgrade by two increments). The population was deemed to be indirect when the outcome included evidence from studies with mixed severity OSAHS populations. 9 surgeries in this comparison: surgically inserted palatal implants; temperature-controlled radiofrequency tissue ablation (TCRFTA)- tongue base and palate; laser assisted uvulopalatoplasty (LAUP)

Table 4: Clinical evidence summary: Surgery versus conservative management/no surgery/sham surgery- severe OSAHS [Category of surgery: oro-pharyngeal surgery and nasal surgery][9]

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Conservative management/no surgery/sham surgery- SEVERE	Risk difference with Surgery versus conservative management/no surgery/sham surgery (95% CI)
AHI (events/hr) (all studies) Lower is better	297 (5 studies)	⊕⊖⊖⊖ VERY LOW[1,2,,6] due to risk of bias, inconsistency, indirectness		Mean in control group was 34.78	The mean AHI in the intervention groups was 15.01 lower (23.67 to 6.34 lower)
AHI (events/hr) sub-group analysis oropharyngeal surgery Lower is better	248 (4 studies)	⊕⊕⊖⊖ LOW[1], [6] due to risk of bias, indirectness,		Mean in control group was 35.45	The mean AHI - oropharyngeal surgery in the intervention groups was 18.72 lower (24.79 to 12.64 lower)
AHI (events/hr) -subgroup analysis nasal surgery Lower is better	49 (1 study)	⊕⊕⊖⊖ LOW[1],[6] due to risk of bias, indirectness		Mean in control group was 32.1	The mean AHI - nasal surgery in the intervention groups was 0.6 lower (9.7 lower to 8.5 higher)
AHI (events/hr) sub-group analysis- all people with OSAHS (including people who are tolerant and not tolerant to CPAP)Lower is better	164 (3 studies)	⊕⊖⊖⊖ VERY LOW[1,2,,6] due to risk of bias, inconsistency, indirectness		Mean in control group was 36.94	The mean ahi in sub-group all people with OSAHS (including people who are tolerant and not tolerant to CPAP) in the intervention groups was 15.59 lower (21.02 to 10.17 lower)
AHI (events/hr) sub-group analysis- not tolerant/adherent to CPAP	133 (2 studies)	⊕⊖⊖⊖ VERY LOW[1,3,6]		Mean in control group was 31.55	The mean ahi in sub-group not tolerant/adherent to CPAP

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Conservative management/no surgery/sham surgery-SEVERE	Risk difference with Surgery versus conservative management/no surgery/sham surgery (95% CI)
Lower is better		due to risk of bias, imprecision, indirectness			in the intervention groups was 13.53 lower (20.21 to 6.85 lower)
FOSQ (Functional Outcome of Sleep Questionnaire) Scale from 5-20 Higher is better	164 (2 studies)	⊕⊖⊖⊖ VERY LOW[1,3,6] due to risk of bias, imprecision, indirectness		Mean in control group was 8.05	The mean FOSQ in the intervention groups was 2.07 higher (1.42 to 2.71 higher)
ESS (all studies) Scale 0 to 24 Higher is worse	233 (4 studies)	⊕⊖⊖⊖ VERY LOW[1,,6,2] due to risk of bias, indirectness, inconsistency		Mean in control group was 10.86	The mean ESS in the intervention groups was 4.13 lower (6.80 to 1.46 lower)
ESS – sub-group analysis oropharyngeal surgery Scale 0 to 24 Higher is worse	184 (3 studies)	⊕⊖⊖⊖ VERY LOW[1,,2,6] due to risk of bias, indirectness, inconsistency		Mean in control group was 10.3	The mean ESS - oropharyngeal surgery in the intervention groups was 5.37 lower (7.14 to 3.59 lower)
ESS – sub-group analysis nasal surgery Scale 0 to 24 Higher is worse	49 (1 study)	⊕⊕⊕⊖ VERY LOW[1,,3,6] due to risk of bias, imprecision, indirectness		Mean in control group was 12.5	The mean ESS - nasal surgery in the intervention groups was 0.8 lower (2.81 lower to 1.21 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Conservative management/no surgery/sham surgery- SEVERE	Risk difference with Surgery versus conservative management/no surgery/sham surgery (95% CI)
ESS sub-group analysis- all people with OSAHS (including people who are tolerant and not tolerant to CPAP)Scale 0 to 24 Higher is worse	99 (2 studies)	⨁⊖⊖⊖ VERY LOW[1],[2,6] due to risk of bias, indirectness, inconsistency		Mean in control group was 11.6	The mean ess (sub-group analysis in all comers in the intervention groups was 3.92 lower (10.02 lower to 2.25 higher)
ESS sub-group analysis- not tolerant/adherent to CPAP Scale 0 to 24 Higher is worse	134 (2 studies)	⨁⨁⊖⊖ LOW[1,6,] due to risk of bias, indirectness		Mean in control group was 10.05	The mean ess sub-group analysis in not tolerant/adherent to CPAP in the intervention groups was 4.74 lower (6.28 to 3.21 lower)
Velopharyngeal insufficiency (speech abnormalities)	32 (1 study)	⨁⊖⊖⊖ VERY LOW[1,3],[6] due to risk of bias, imprecision, indirectness	OR 6.28 (0.37 to 107.44)[5]	Moderate 0 per 1000	111 more per 1000 (from 60 fewer to 280 more)[4]
SpO2 (peripheral capillary oxygen saturation)	33 (1 study)	⨁⊖⊖⊖ VERY LOW[1,3,6] due to risk of bias, imprecision, indirectness		Mean in control group was 91.7	The mean SpO2 in the intervention groups was 1.2 higher (2.25 lower to 4.65 higher)
Serious adverse events (Lojander 1996 -CV events (non Q myocardial infarction and transient ischemic cerebral attack; MacKay 2020- myocardial infarction and hematemesis of old blood)	125 (2 studies)	⨁⊖⊖⊖ VERY LOW[1,3,6] due to risk of bias, imprecision, indirectness	OR 6.38 (0.85 to 47.69)[5]	Moderate 0 per 1000	60 more per 1000 (from3 fewer to 120 more)[4] (4/66 in surgery group and 0/59 in control group)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Conservative management/no surgery/sham surgery-SEVERE	Risk difference with Surgery versus conservative management/no surgery/sham surgery (95% CI)
Tracheostomy	32 (1 study)	⊕⊖⊖⊖ VERY LOW[1,3,6] due to risk of bias, imprecision, indirectness	OR 5.92 (0.11 to 307.57) 5	Moderate	
				0 per 1000	50 more per 1000 (from 90 fewer to 200 more)4 (1/18 in surgery group and 0/14 in control group)
Re-operations[10]	32 (1 study)	⊕⊖⊖⊖ VERY LOW[1,3,6] due to risk of bias, imprecision, indirectness	OR 6.28 (0.37 to 107.44) 5	Moderate	
				0 per 1000	111 more per 1000 (from 60 fewer to 280 more)[4] (2/18 in surgery group and 0/14 in control group)
Mortality	32 (1 study)	⊕⊕⊖⊖ LOW[1,6,] due to risk of bias, indirectness	Not estimable	Not estimable	Zero events in both groups
24 h ambulatory systolic blood pressure	99 (1 study)	⊕⊖⊖⊖ VERY LOW[1,3,4] due to risk of bias, indirectness, imprecision		Mean in control group was 124.7	The mean 24 h ambulatory systolic blood pressure (copy) in the intervention groups was 4.7 lower (9.76 lower to 0.36 higher)
24 h ambulatory diastolic blood pressure	99 (1 study)	⊕⊖⊖⊖ VERY LOW[1,3,4] due to risk of bias, indirectness, imprecision		Mean in control group was 77.7	The mean 24 h ambulatory diastolic blood pressure (copy) in the intervention groups was 3.6 lower (7.37 lower to 0.17 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Conservative management/no surgery/sham surgery- SEVERE	Risk difference with Surgery versus conservative management/no surgery/sham surgery (95% CI)

1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias

2 Downgraded by 1 or 2 increments for heterogeneity, unexplained by subgroup analysis. Random effects analysis used

3 Downgraded by one increment if the confidence interval crossed one MID and downgraded by two increments if the confidence interval crossed both MIDs. Established MIDs for SF-36 physical/mental- 2/3; FOSQ- 2; ESS -2.5 Default MIDs (0.5XSD) used for all other outcomes.

4 Risk difference calculated in Revman

5 Peto odds ratio used when zero events in one/both groups.

6 Downgraded by 1 or 2 increments because the majority of the evidence included an indirect population (downgrade by one increment) or a very indirect population (downgrade by two increments). The population was deemed to be indirect when the outcome included evidence from studies with mixed severity OSAHS populations.

9 surgeries in this comparison: tonsillectomy with uvulopalatopharyngoplasty (TE-UPPP); sub mucous resection of the deviated nasal septum; tonsillectomy with uvulopalatopharyngoplasty (TE-UPPP); UPPP; barbed repositioning pharyngoplasty (BRP); modified uvulopalatopharyngoplasty and minimally invasive tongue volume reduction.

10 In the conservative management group one patient needed CPAP, and three underwent surgery due to worsening of

symptoms.

Table 5: Clinical evidence summary: surgery versus APAP (auto titrating positive airway pressure)- severe OSAHS [category of surgery- Skeletal Framework Surgery][7]

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Surgery versus APAP- SEVERE (95% CI)
AHI (events/hour) Lower is better	50 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2,] due to risk of bias, imprecision,		Mean in control group was 6.3	The mean AHI in the intervention groups was 1.8 higher (1.04 lower to 4.64 higher)
ESS(Epworth Sleepiness Scale) Scale 0 to 24 Higher is worse	50 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2.] due to risk of bias, imprecision,		Mean in control group was 5.9	The mean ESS in the intervention groups was 1.8 higher (0.99 to 2.61 higher)
Bleeding	50 (1 study)	⊕⊖⊖⊖ VERY LOW[1,] due to risk of bias	Not estimable	See comment	See comment (zero events in both the groups)
Dyspnoea	50 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2,] due to risk of bias, imprecision,	OR 7.39 (0.15 to 372.38)4	Moderate	
				0 per 1000	40 more per 1000 (from 60 fewer to 140 more)[3]
Persistent paresthesia[5]	50 (1 study)	⊕⊖⊖⊖ VERY LOW[1] due to risk of bias,	OR 9.77 (2.01 to 47.5)4	Moderate	
				0 per 1000	280 more per 1000 (from 90 more to 460 more)[3]
Lesion in facial skin (with APAP)[6]	50 (1 study)	⊕⊖⊖⊖ VERY LOW[1,2,] due to risk of bias, imprecision	OR 0.14 (0 to 6.82)4	Moderate	
				40 per 1000	40 fewer per 1000 (from 14 fewer to 60 more)[3]
Mortality					Not reported

1 Downgraded by 1 increment if the majority of the evidence was at high risk of bias and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
2 Downgraded by one increment if the confidence interval crossed one MID and downgraded by two increments if the confidence interval crossed both MIDs. Established MIDs for SF-36 physical/mental- 2/3 ; FOSQ- 2 ; ESS -2.5 . AHI- different severity groups, likely true MCID will vary, qualitatively

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Surgery versus APAP- SEVERE (95% CI)
considered in decision making throughout. GRADE default MIDs (0.5XSD)used for all other continuous outcomes. 3 Risk difference calculated in Revman 4 Peto odds ratio used when zero events in one/both groups. 5 Persistent but not disturbing paresthaesia around the chin. 6 Lesion to the facial skin was overcome with a new type of mask. 7 surgery in this comparison- Maxillomandibular advancement [MMA].					

Table 6: Clinical evidence summary: Surgery versus CPAP- moderate OSAHS [category of surgery- oro-pharyngeal Surgery][4]

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Surgery versus CPAP- MODERATE (95% CI)
SF-36 physical Scale from 0-100 Higher is better	48 (1 study)	⨁⨀⨀⨀ VERY LOW[1,2] due to imprecision, indirectness		Mean in control group was 0.1	The mean SF-36 physical in the intervention groups was 0.4 higher (3.71 lower to 4.51 higher)
SF-36 mental Scale from 0-100 Higher is better	48 (1 study)	⨁⨁⨀⨀ LOW [1,2] due to imprecision, indirectness		Mean in control group was 2	The mean SF-36 mental in the intervention groups was 0.9 higher (2.91 lower to 4.71 higher)
FOSQ (Functional Outcomes of Sleep Questionnaire) Scale from 5-20 Higher is better	51 (1 study)	⨁⨁⨁⨀ MODERATE[2] due to indirectness		Mean in control group was 1.5	The mean FOSQ (functional outcomes of sleep questionnaire) in the intervention groups was 0.3 lower (1.33 lower to 0.73 higher)
ESS (Epworth Sleepiness Score)[3] Scale 0 to 24 Higher is worse	51 (1 study)	⨁⨁⨀⨀ LOW [1,2] due to		Mean in control group was -2.3	The mean ESS in the intervention groups was 0.2 higher (2.33 lower to 2.73 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Surgery versus CPAP- MODERATE (95% CI)
		imprecision, indirectness			
Mortality					Not reported

1 Downgraded by one increment if the confidence interval crossed one MID and downgraded by two increments if the confidence interval crossed both MIDs. Established MIDs for SF-36 physical/mental- 2/3; FOSQ- 2 ; ESS -2.5 . GRADE default MID (0.5XSD) used for all other continous outcomes.

2 Downgraded by 1 or 2 increments because the majority of the evidence included an indirect population (downgrade by one increment) or a very indirect population (downgrade by two increments). The population was deemed to be indirect when the outcome included evidence from studies with mixed severity OSAHS populations.3 Change values: post treatment mean -baseline mean. Baseline ESS: surgery: 11.9 (4.6); CPAP: 12.6 (5.0)

4 surgery in this comparison: temperature-controlled radiofrequency tissue ablation (TCRFTA).- tongue base and palate-

Table 7: Clinical evidence summary: Surgery versus oral devices- moderate OSAHS [category of surgery- oro-pharyngeal Surgery][5]

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Surgery versus oral devices- MODERATE (95% CI)
AHI (events/hour) - 6 months Lower is better	80 (1 study)	⨁⊖⊖⊖ VERY LOW[1,4] due to imprecision, indirectness		Mean in control group was 6.6	The mean AHI - 6 months in the intervention groups was 2 higher (1.54 lower to 5.54 higher)
AHI (events/hour) - 1 year Lower is better	80 (1 study)	⨁⨁⊖⊖ LOW[1,4] due to imprecision, indirectness		Mean in control group was 5.9	The mean AHI - 1 year in the intervention groups was 4.5 higher (0.52 to 8.48 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Surgery versus oral devices- MODERATE (95% CI)
AHI (events/hour) - 4 years Lower is better	72 (1 study)	⨁⨁⊖ MODERATE[,4] due to indirectness		Mean in control group was 7.2	The mean AHI - 4 years in the intervention groups was 7 higher (5.61 to 8.39 higher)
Oxygen desaturation index (ODI) - 6 months Lower is better	80 (1 study)	⨁⊖⊖⊖ VERY LOW[1,4] due to imprecision, indirectness		Mean in control group was 6.4	The mean ODI - 6 months in the intervention groups was 1.6 higher (2.04 lower to 5.24 higher)
Oxygen desaturation index (ODI) - 1 year Lower is better	80 (1 study)	⨁⨁⊖⊖ LOW[1,4] due to imprecision, indirectness		Mean in control group was 6.1	The mean ODI - 1 year in the intervention groups was 3.2 higher (1.13 lower to 7.53 higher)
Oxygen desaturation index (ODI) - 4 years Lower is better	72 (1 study)	⨁⨁⊖⊖ LOW[1,4] due to imprecision, indirectness		Mean in control group was 6.7	The mean ODI - 4 years in the intervention groups was 6.4 higher (2.33 to 10.47 higher)
Quality of life: Vitality (minor symptoms evaluation profile MSE) at 1 year Scale 0-100 Lower is better	80 (1 study)	⨁⨁⊖⊖ LOW[1,4] due to imprecision, indirectness		Mean in control group was 31.6	The mean quality of life: vitality (minor symptoms evaluation profile MSE) at 1 year in the intervention groups was 5.2 lower (10.81 lower to 0.41 higher)
Quality of life: Sleep (minor symptoms evaluation profile MSE) at 1 year Scale 0-100	80 (1 study)	⨁⊖⊖⊖ VERY LOW[1,4] due to		Mean in control group was 29.2	The mean quality of life: sleep (minor symptoms evaluation profile MSE) at 1 year in the intervention groups was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Surgery versus oral devices- MODERATE (95% CI)
Lower is better		imprecision, indirectness			4 lower (11.1 lower to 3.1 higher)
Quality of life: Contentment (minor symptoms evaluation profile MSE) at 1 year Scale 0-100 Lower is better	80 (1 study)	⨁⨁⨀⨀ LOW[1,4] due to imprecision, indirectness		Mean in control group was 33.7	The mean quality of life: contentment (minor symptoms evaluation profile MSE) at 1 year in the intervention groups was 6.3 lower (11.91 to 0.69 lower)
Dysphagia at 4 years	72 (1 study)	⨁⨁⨁⨀ MODERATE [4] due to indirectness	OR 6.55 (0.87 to 49.14)3	Moderate	
				0 per 1000	100 more per 1000 (from 4 fewer to 200 more)[2]
Nasopharyngeal regurgitation at 4 years	72 (1 study)	⨁⨁⨁⨀ LOW [1,4] due to imprecision ,indirectness	OR 6.37 (0.63 to 64.21)3	Moderate	
				0 per 1000	75 more per 1000 (from 200 fewer to 170 more)[2]
Mortality					Not reported

1 Downgraded by one increment if the confidence interval crossed one MID and downgraded by two increments if the confidence interval crossed both MIDs. Established MIDs for SF-36 physical/mental- 2/3 ; FOSQ- 2 ; ESS -2.5 .. GRADE default MIDs (0.5XSD)used for all other continous outcomes.
2 Risk difference calculated in Revman
3 Peto odds ratio used when zero events in one/both groups.

4 Downgraded by 1 or 2 increments because the majority of the evidence included an indirect population (downgrade by one increment) or a very indirect population (downgrade by two increments). The population was deemed to be indirect when the outcome included evidence from studies with mixed severity OSAHS populations. 5 surgeries in this comparison: UPPP

See appendix F for full GRADE tables.

Narrative results:

Data has been presented narratively for studies where the data could not be analysed in GRADE. Narrative data was considered alongside the GRADE evidence by the committee when making recommendations. The overall study quality was taken into account as GRADE analysis for each outcome could not be performed.

Woodson 2003: TCRFTA vs sham TCRFTA vs CPAP (n=90) (very low quality)

Data could not be analysed as AHI not measured in the same way in the 2 groups

AHI available for surgery group from polysomnography and AHI for CPAP group AHI from built-in CPAP monitor (not polysomnography).

AHI in surgery (change values from baseline): -1.8 (11.5); Base line AHI: 21.3 (11.1)

CPAP: AHI at 8 weeks after CPAP started: 4.6 (2.7); Baseline AHI: 19.8 (9.9)

1.5 Economic evidence

1.5.1 Included studies

No health economic studies were included.

1.5.2 Excluded studies

No relevant health economic studies were excluded due to assessment of limited applicability or methodological limitations.

See also the health economic study selection flow chart in Appendix G: G.

1.5.2.1 Unit costs

Relevant unit costs were provided to the committee for consideration of cost effectiveness.

Table 8: Unit costs of surgery

Category	Healthcare Resource Group (HRG)	Description	Activity	Mean cost
Day case	CA64Z	Uvulopalatoplasty or Uvulopalatopharyngoplasty	192	£1,525
Elective inpatient	CA64Z	Uvulopalatoplasty or Uvulopalatopharyngoplasty	163	£2,568
Day case	CA60C	Tonsillectomy, 4 years and over	17,283	£1,492
Elective inpatient	CA60C	Tonsillectomy, 4 years and over	6,858	£1,913

Source: National schedule of NHS costs 2018/19[21]

1.5.3 Economic model

A threshold analysis was conducted to see what quality of life improvement would be required for surgery to be considered cost effective (at £20,000 per QALY gained) - Table 9.

Table 9: Improvement required for surgery to be considered cost effective

Healthcare Resource Group description	Cost	QALYs required (=cost /20000)	EQ-5D improvement required by duration of improvement (=QALYs required / years)		
			2 years	5 years	10 years
Day case Uvulopalatoplasty or Uvulopalatopharyngoplasty	£1,525	0.076	0.038	0.015	0.008
Inpatient Uvulopalatoplasty or Uvulopalatopharyngoplasty	£2,568	0.128	0.064	0.026	0.013
Day case Tonsillectomy, 4 years and over	£1,492	0.075	0.037	0.015	0.007
Inpatient Tonsillectomy, 4 years and over	£1,913	0.096	0.048	0.019	0.010

As a benchmark an increase of 1 point in ESS is associated with a fall in EQ-5D of 0.01.[49] So, on that basis, if the 5.4 point reduction in ESS observed in Table 4 above for oropharyngeal surgery in a severe population was sustained for 2.4 years (0.128 QALYs/0.054) or more it would be considered cost effective (3.4 years if a day of intensive care is required).

1.5.4 Health economic evidence statements

No relevant economic evaluations were identified.

1.6 The committee's discussion of the evidence

1.6.1 Interpreting the evidence

1.6.1.1 The outcomes that matter most

The committee considered the outcomes of health-related quality of life and mortality as critical outcomes for decision making. Other important outcomes included: sleepiness scores (e.g. Epworth Sleepiness Scale (ESS)), Apnoea-Hypopnoea index (AHI), oxygen desaturation index (ODI), CO2 control, permanent adverse effects (e.g. neural dysfunction, open nasality, globus sensation, reversible adverse effects (e.g. pain, infection, secondary bleeding), driving outcomes, neurocognitive outcomes. The committee were also interested in the impact on co-existing conditions such as HbA1c for diabetes, cardiovascular events for cardiovascular disease and systolic blood pressure for hypertension.

No evidence was identified for the outcomes of permanent adverse effects, CO2 control, driving outcomes, neurocognitive outcomes and impact on co-existing conditions such as HbA1c for diabetes, cardiovascular events for cardiovascular disease, and systolic blood pressure for hypertension.

1.6.1.2 The quality of the evidence

The quality of the evidence varied from moderate to very low quality. The majority of the evidence was downgraded due to risk of bias, imprecision and indirectness. Risk of bias was most commonly due to selection bias and performance bias as there was a lack of blinding in the studies due to the nature of the interventions. When a mixed severity population was included the severity of the whole population was assumed to correspond to the mean AHI of the patients included' and the study was downgraded for indirectness. The committee also acknowledged that some uncertainty existed across the effect sizes seen within the evidence, with some confidence intervals crossing the MID thresholds or line of no effect. Studies included a variety of different procedures including uvulopalatopharyngoplasty (UPPP) (two studies), multilevel surgery modified uvulopalatopharyngoplasty and minimally invasive tongue volume reduction (one study); barbed repositioning pharyngoplasty (BRP) (one study), laser assisted uvulopalatoplasty (LAUP) (one study), temperature-controlled radiofrequency tissue ablation (TCRFTA) (one study), maxillomandibular advancement [MMA] (one study), tonsillectomy with uvulopalatopharyngoplasty (TE-UPPP) (two studies). sub mucous resection of the deviated nasal septum (one study) and surgically inserted palatal implants (one study). There were a range of surgical procedures within each strata , which could have influenced the observed effectiveness of each type of surgery within the evidence, leading to greater uncertainty around the point estimates.

There was evidence from ten studies (14 papers); eight studies compared surgery with no treatment initially/delayed surgery/conservative management; one study compared oral devices with surgery; one study compared surgery with APAP and one study compared surgery with CPAP.
Two studies included patients who were unable to tolerate or adhere to CPAP. The committee felt this was important information as CPAP intolerance is an important clinical problem.

Overall in the studies, participants suffered from moderate to severe sleep apnoea, with mixed levels of AHI. Studies were stratified based on the AHI/ODI severity of the population. Follow-up in the studies ranged from 2 months to 4 years. The committee considered the clinical importance for AHI on a case by case basis, taking into consideration the baseline AHI and the improvement in severity of sleep apnoea.

1.6.1.3 Benefits and harms

Surgery versus conservative management/no surgery/sham surgery - Moderate OSAHS [category of surgery: oro-pharyngeal surgery]

The evidence suggested that there was a small reduction in AHI with surgery in people with moderate OSAHS which is possibly clinically significant and was associated with some measures of quality of life improvement (SF-36 vitality, physical health and mental health). The committee agreed that even though the reduction in AHI may be small, it is likely that the proportion of apnoea to hypopneas may have changed favourably. The committee observed that in this population lowering AHI may be of clinical benefit as it could enable some patients to utilise CPAP when they have been not tolerant previously.

The evidence suggested that there was no clinically important difference between surgery and conservative management/no surgery/sham surgery for the outcomes of ESS, quality of life (FOSQ, SAQLI). The evidence suggested adverse effects such dysphagia, infection, pain, nasal regurgitation, ulcerations, and bleeding with surgery but these were not found to be clinically important. Dysphagia in most of these procedures was expected and due to oedema of the pharynx and pain post-operatively, as both would make the swallowing process difficult but as healing took place swallowing returned to normal. Infection and bleeding can be appropriately treated and hence the adverse effects were considered not to be clinically harmful in the medium or long term. There was no evidence for the outcome mortality. The committee therefore felt that the clinical benefits of this surgery outweighed the side effects and would consider this in people with moderate OSAHS not tolerant of optimised CPAP.

Surgery versus conservative management/no surgery/sham surgery - severe OSAHS [category of surgery: oro-pharyngeal surgery and nasal surgery]

The evidence suggested that there was a reduction in AHI with surgery in people with severe OSAHS, which is possibly clinically significant. Sub-group analysis based on type of surgery suggested that this benefit was limited to oropharyngeal surgery and not to nasal surgery. Sub-group analysis based on selected population in the studies suggested that there was small reduction in AHI for both all people with OSAHS (including people who are tolerant and not tolerant to CPAP)and people who are not tolerant/adherent to CPAP.

The evidence suggested clinical benefit of surgery for ESS and quality of life (FOSQ). Sub-group analysis for ESS based on site surgery suggested that there was benefit for both oropharyngeal surgery and nasal surgery. Sub-group analysis based on selected population in the studies suggested that there was benefit of ESS for both all people with OSAHS (including people who are tolerant and not tolerant to CPAP)and people who are not tolerant/adherent to CPAP.

The evidence suggested that there was no clinically important difference between surgery and conservative management/no surgery/sham surgery for the outcomes of SPO2 (peripheral capillary oxygen saturation), re-operations, 24 h ambulatory systolic blood pressure and 24 h ambulatory diastolic blood pressure. There were adverse effects such as speech abnormalities, cardiovascular events and need for tracheotomy with surgery and although they were not found to be clinically significant the committee considered that this could be due to the small number of participants/studies. The committee noted that one study of maxillomandibular advancement (Vicini 2010) routinely carried out temporary tracheostomy, indicating the extensive nature of the surgery. Evidence from one study suggested that there was no mortality in either group.

Surgery versus APAP - severe OSAHS [category of surgery - skeletal framework surgery]

The evidence suggested that there was no clinically important difference between surgery and APAP for values of AHI and ESS. The evidence suggested reversible adverse effects with surgery such as dyspnoea and paraesthesia but these were not found to be clinically important. Evidence from one study suggested that there were no bleleding events in either group. There was no evidence for the outcome mortality.

Surgery versus CPAP - moderate OSAHS [category of surgery - oropharyngeal surgery]

The evidence suggested that there was no clinically important difference between surgery and CPAP for quality of life outcomes (SF-36 physical, mental; FOSQ) and ESS.

Narrative evidence from one study showed that there was a greater reduction in AHI in CPAP group compared to surgery group. The committee however were not confident of this outcome as it was based on very low quality evidence. There was no evidence for the outcome mortality.

Surgery versus oral devices - moderate OSAHS [category of surgery - oropharyngeal surgery]

The evidence suggested that there was clinically important benefit of oral devices for quality of life (vitality, sleep, contentment) compared to surgery, although there was some uncertainty around the effect estimates. The evidence suggested that there was no clinically important difference between surgery and oral devices for AHI and ODI at 6 months, 1 year and 4 years. The evidence suggested reversible adverse effects with surgery, such as dysphagia and nasopharyngeal regurgitation but were not found to be clinically significant. There was no evidence for the outcome mortality.

Overall conclusions - people to consider for surgery

As there was limited evidence on each of the surgical interventions for OSAHS the committee also used their clinical knowledge and experience to make the recommendations. The evidence suggested that oropharyngeal surgery (including tonsillectomy) was effective in some people with moderate and severe OSAHS, but not more effective than CPAP or mandibular advancement splints/oral devices. Of note however, oropharyngeal surgery was effective in people with moderate or severe OSAHS who are unable to tolerate or adhere to CPAP and mandibular advancement splints. There were some adverse effects associated with surgery, but the committee agreed that they were not clinically significant.

Based on their knowledge and experience the committee agreed that tonsillectomy could be considered in people with large obstructive tonsils with OSAHS. People with a BMI of 35 kg/m^2 or above are unlikely to benefit from this surgery because they are more likely to have multi-level upper airway obstruction, and therefore surgery can be considered in patients with BMI of less than 35 kg/m^2.. The committee decided that there was enough evidence of benefit to make a recommendation to consider tonsillectomy, but the evidence was not of high enough quality to justify a stronger recommendation to offer tonsillectomy in all cases.

There was no direct evidence for people with mild OSAHS but the committee agreed that tonsillectomy should be applicable to all severities when tonsils are clearly causing obstruction.

Based on the evidence and their knowledge and experience the committee also agreed that oropharyngeal surgery other than, or in addition to, tonsillectomy could be considered for people with severe OSAHS who have been unable to tolerate CPAP and a bespoke mandibular advancement splint despite medically supervised attempts. The best trial evidence included people with either moderate or severe OSAHS, but the majority of the subjects were in the severe category and the committee agreed that benefit was more likely in this group. There are no other treatment options for people with severe OSAHS who cannot tolerate CPAP and mandibular advancement splints, and the committee agreed that

surgery for the right people would improve their quality of life. The committee also noted the potential risks of surgical intervention in people with severe OSAHS, and stressed that prior to considering surgery, patients who are not tolerant/non-adherent to CPAP should have fully explored all CPAP options under medical supervision for a sufficient period of time. The committee agreed that selection of patients who could be considered for oropharyngeal surgery, and selection of the correct procedure is critical.

This includes an assessment of anaesthetic risk and of the type and extent of surgery, which is critical because the outcome will depend on the anatomical and physiological phenotype of OSAHS. They therefore made a recommendation for referral for surgical consideration rather than surgery itself, acknowledging that precise individual assessment by the surgical team would be needed.

The committee's recommendation for tonsillectomy is broadly in line with current practice. However, people who are unable to tolerate/non-adherent to CPAP and mandibular advancement splint are currently not usually referred for oro-pharyngeal surgery so there is likely to be a change in practice for some providers. This recommendation is likely to only affect a small minority of people with severe OSAHS who are not helped by other treatments, have few comorbidities and for whom surgery is a suitable option.

The committee also discussed various other surgical scenarios but decided against making recommendations for these. They noted that people with retrognathia or maxillary retrusion could be considered for assessment for skeletal framework surgery but agreed that data is limited and that this procedure is associated with considerable morbidity. They agreed that nasal surgery alone does not ameliorate OSAHS, although it may facilitate adherence to CPAP therapy in people who are not tolerant of this.

The committee agreed to make a research recommendation on upper airway surgical interventions for people with obstructive sleep apnoea/hypopnoea syndrome who are unable to tolerate or adhere to CPAP, as there was limited evidence for the applicability of this approach.

1.6.2 Cost effectiveness and resource use

There were no published economic evaluations found and therefore the cost effectiveness of surgery is uncertain. In the absence of cost effectiveness evidence, unit costs of surgery were presented to the committee.

The committee decided that, on balance, tonsillectomy was likely to be cost effective for patients with OSAHS and large obstructive tonsils. This was thought to be current practice.

The committee also decided that referral for oropharyngeal surgery could be cost effective for carefully selected people with severe OSAHS who have been unable to tolerate CPAP and a bespoke mandibular advancement splint despite medically supervised attempts, if the treatment effects are maintained for long enough. The committee did not think a daycase operation would be suitable given the complexity of the surgery and the severity of OSAHS. A simple threshold analysis showed that inpatient oropharyngeal surgery in this population would be cost effective if the measured effect on sleepiness at 6 months (and therefore quality of life) was sustained for 2.4 years on average (3.4 years if a day of intensive care is required). This assumes that there are no other significant impacts on resource use. We know there will be some additional costs associated with preoperative and postoperative assessment. However, we would also expect some cost savings and QALY gains from reduced cardiovascular events and reduced road traffic accidents.

There are no other treatment options for these people and the committee agreed that, for the right people, oropharyngeal surgery improves quality of life. On that basis, the committee

concluded that if the benefits are there at 6 months then it is reasonable to assume they will be maintained in the long-term and surgery should be cost effective. This recommendation is likely to be a change in practice, although it is likely to only affect a small minority of people, who are not helped by CPAP, who have little comorbidity but who are robust enough to have surgery and choose to have it.

1.6.3 Other factors the committee took into account

The committee stressed the importance of experienced anaesthetic input to all surgical procedures, including post-operative care, as airway swelling may be more marked acutely with a risk of bleeding, in a population who may be at significant cardiovascular risk.

References

1. Ayers CM, Lohia S, Nguyen SA, Gillespie MB. The effect of upper airway surgery on continuous positive airway pressure levels and adherence: A systematic review and meta-analysis. ORL: Journal of Oto-Rhino-Laryngology and Its Related Specialties. 2016; 78(3):119-125

2. Back LJJ, Liukko T, Rantanen I, Peltola JS, Partinen M, Ylikoski J et al. Radiofrequency surgery of the soft palate in the treatment of mild obstructive sleep apnea is not effective as a single-stage procedure: A randomized single-blinded placebo-controlled trial. Laryngoscope. 2009; 119(8):1621-1627

3. Balcerzak J, Arcimowicz M, Byśkiniewicz K, Bielicki P, Korczyński P, Niemczyk K et al. Effect of nasal obstruction on subjective assessment of sleep quality by the patients with obstructive sleep apnea. Polski Merkuriusz Lekarski. 2005; 19(111):286-287

4. Baradaranfar MH, Edalatkhah M, Dadgarnia MH, Atighechi S, Behniafard N, Mirvakili A et al. The effect of uvulopalatopharyngoplasty with tonsillectomy in patients with obstructive sleep apnea. Indian Journal of Otolaryngology & Head & Neck Surgery. 2015; 67(Suppl 1):29-33

5. Bayir O, Acar M, Yuksel E, Yuceege M, Saylam G, Tatar EC et al. The effects of anterior palatoplasty on floppy eyelid syndrome patients with obstructive sleep apnea. Laryngoscope. 2016; 126(9):2171-2175

6. Binar M, Karakoc O. Anterior palatoplasty for obstructive sleep apnea: A systematic review and meta-analysis. Otolaryngology - Head & Neck Surgery. 2018; 158(3):443-449

7. Bostanci A, Turhan M. A systematic review of tongue base suspension techniques as an isolated procedure or combined with uvulopalatopharyngoplasty in obstructive sleep apnea. European Archives of Oto-Rhino-Laryngology. 2016; 273(10):2895-2901

8. Bridgman S, Dunn K. Systematic review of the surgical treatment of obstructive sleep apnoea. Journal of Clinical Excellence. 2000; 1:223-226

9. 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

10. Camacho M, Certal V, Brietzke SE, Holty JE, Guilleminault C, Capasso R. Tracheostomy as treatment for adult obstructive sleep apnea: a systematic review and meta-analysis. Laryngoscope. 2014; 124(3):803-811

11. Camacho M, Li D, Kawai M, Zaghi S, Teixeira J, Senchak AJ et al. Tonsillectomy for adult obstructive sleep apnea: A systematic review and meta-analysis. Laryngoscope. 2016; 126(9):2176-2186

12. Camacho M, Nesbitt NB, Lambert E, Song SA, Chang ET, Liu SY et al. Laser-assisted uvulopalatoplasty for obstructive sleep apnea: A systematic review and meta-analysis. Sleep. 2017; 40(3):01

13. Camacho M, Noller MW, Del Do M, Wei JM, Gouveia CJ, Zaghi S et al. Long-term results for maxillomandibular advancement to treat obstructive sleep apnea: A meta-analysis. Otolaryngology - Head & Neck Surgery. 2019; 160(4):580-593

14. Camacho M, Riaz M, Capasso R, Ruoff CM, Guilleminault C, Kushida CA et al. The effect of nasal surgery on continuous positive airway pressure device use and therapeutic treatment pressures: a systematic review and meta-analysis. Sleep. 2015; 38(2):279-286

15. Camacho M, Teixeira J, Abdullatif J, Acevedo JL, Certal V, Capasso R et al. Maxillomandibular advancement and tracheostomy for morbidly obese obstructive sleep apnea: a systematic review and meta-analysis. Otolaryngology - Head & Neck Surgery. 2015; 152(4):619-630

16. Caples SM, Rowley JA, Prinsell JR, Pallanch JF, Elamin MB, Katz SG 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

17. Certal VF, Zaghi S, Riaz M, Vieira AS, Pinheiro CT, Kushida C et al. Hypoglossal nerve stimulation in the treatment of obstructive sleep apnea: A systematic review and meta-analysis. Laryngoscope. 2015; 125(5):1254-1264

18. Chisholm E, Kotecha B. Oropharyngeal surgery for obstructive sleep apnoea in CPAP failures. European Archives of Oto-Rhino-Laryngology. 2007; 264(1):51-55

19. Choi JH, Cho SH, Kim SN, Suh JD, Cho JH. Predicting outcomes after uvulopalatopharyngoplasty for adult obstructive sleep apnea: A meta-analysis. Otolaryngology - Head & Neck Surgery. 2016; 155(6):904-913

20. Couch ME, Senior B. Nonsurgical and surgical treatments for sleep apnea. Anesthesiology Clinics of North America. 2005; 23(3):525-534

21. Department of Health. National Schedule of NHS Costs 2018/19 2020. Available from: https://www.england.nhs.uk/national-cost-collection/#ncc1819 Last accessed: 12/06/2020.

22. DRKS. The surgical treatment of obstructive sleep apnea (OSA) using tonsillectomy with uvulapalatopharyngoplasty - A clinical randomized study [DRKS00000549]. 2010. Available from: https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00000549 Last accessed: 20/02/2020.

23. Ferguson KA, Heighway K, Ruby RR. A randomized trial of laser-assisted uvulopalatoplasty in the treatment of mild obstructive sleep apnea. American Journal of Respiratory and Critical Care Medicine. 2003; 167(1):15-19

24. Fernandez-Ferrer L, Montiel-Company JM, Pinho T, Almerich-Silla JM, Bellot-Arcis C. Effects of mandibular setback surgery on upper airway dimensions and their influence on obstructive sleep apnoea - a systematic review. Journal of Cranio-Maxillo-Facial Surgery. 2015; 43(2):248-253

25. Franklin KA, Anttila H, Axelsson S, Gislason T, Maasilta P, Myhre KI et al. Effects and side-effects of surgery for snoring and obstructive sleep apnea--a systematic review. Sleep. 2009; 32(1):27-36

26. Friedman M, Schalch P, Joseph NJ. Palatal implant for the treatment of OSAHS and snoring. Otolaryngology - Head & Neck Surgery. 2006; 135(2 (Suppl 1)):P211-212

27. Friedman M, Schalch P, Lin HC, Kakodkar KA, Joseph NJ, Mazloom N. Palatal implants for the treatment of snoring and obstructive sleep apnea/hypopnea syndrome. Otolaryngology - Head & Neck Surgery. 2008; 138(2):209-216

28. Gaddam S, Gunukula SK, Mador MJ. Post-operative outcomes in adult obstructive sleep apnea patients undergoing non-upper airway surgery: a systematic review and meta-analysis. Sleep & Breathing. 2014; 18(3):615-633

29. Gakwaya S, Borel JC, Masse JF, Cesar S, Series F. Impact of mandibular advancement device associated with CPAP interface on upper airway mechanical properties assessed by phrenic nerve stimulation in obstructive sleep apnea patients: preliminary results. Sleep Medicine. 2011; 12(Suppl 1):S107-S108 [W-I-052]

30. Gao YN, Wu YC, Lin SY, Chang JZ, Tu YK. Short-term efficacy of minimally invasive treatments for adult obstructive sleep apnea: A systematic review and network meta-analysis of randomized controlled trials. Journal of the Formosan Medical Association. 2019; 118(4):750-765

31. Gerek M, Akçam T, Durmaz A. Radiofrequency thermal ablation of the soft palate in patients with snoring and mild sleep apnea. Journal of Ear, Nose, and Throat: KBB. 2005; 14(3-4):67-78

32. Halle TR, Oh MS, Collop NA, Quyyumi AA, Bliwise DL, Dedhia RC. Surgical treatment of OSA on cardiovascular outcomes: A systematic review. Chest. 2017; 152(6):1214-1229

33. Ho LI, Leung YY. Skeletal advancement surgery in patients with moderate-to-severe obstructive sleep apnea. Current Sleep Medicine Reports. 2020; 6(2):61-66

34. Holmlund T, Franklin KA, Levring Jaghagen E, Lindkvist M, Larsson T, Sahlin C et al. Tonsillectomy in adults with obstructive sleep apnea. Laryngoscope. 2016; 126(12):2859-2862

35. Holty JE, Guilleminault C. Maxillomandibular advancement for the treatment of obstructive sleep apnea: a systematic review and meta-analysis. Sleep Medicine Reviews. 2010; 14(5):287-297

36. Iftikhar IH, Bittencourt L, Youngstedt SD, Ayas N, Cistulli P, Schwab R et al. Comparative efficacy of CPAP, MADs, exercise-training, and dietary weight loss for sleep apnea: a network meta-analysis. Sleep Medicine. 2017; 30:7-14

37. Ishii L, Roxbury C, Godoy A, Ishman S, Ishii M. Does nasal surgery improve osa in patients with nasal obstruction and osa? A meta-analysis. Otolaryngology - Head & Neck Surgery. 2015; 153(3):326-333

38. Joar S, Danielle F, Johan B, Arne L, Roberta N, Nanna B. Sleep quality after modified uvulopalatopharyngoplasty: Results from the SKUP3 randomized controlled trial. Sleep. 2018; 41(1):01

39. Justin GA, Chang ET, Camacho M, Brietzke SE. Transoral robotic surgery for obstructive sleep apnea: A systematic review and meta-analysis. Otolaryngology - Head & Neck Surgery. 2016; 154(5):835-846

40. Kompelli AR, Ni JS, Nguyen SA, Lentsch EJ, Neskey DM, Meyer TA. The outcomes of hypoglossal nerve stimulation in the management of OSA: A systematic review and meta-analysis. World Journal Of Otorhinolaryngology - Head And Neck Surgery. 2019; 5(1):41-48

41. Kotecha BT, Hall AC. Role of surgery in adult obstructive sleep apnoea. Sleep Medicine Reviews. 2014; 18(5):405-413

42. Koutsourelakis I, Georgoulopoulos G, Perraki E, Vagiakis E, Roussos C, Zakynthinos SG. Randomised trial of nasal surgery for fixed nasal obstruction in obstructive sleep apnoea. European Respiratory Journal. 2008; 31(1):110-117

43. Koutsourelakis I, Perraki E, Vagiakis E, Dimitrokalli P, Zakynthinos S, Roussos C. Nasal surgery versus sham nasal surgery for obstructive sleep apnoea. European Respiratory Society Annual Congress, Berlin, Germany,October 4-8. 2008:1740

44. Li H. Nasal surgery for obstructive sleep apnea syndrome. Sleep. 2008; 31(Suppl):A146

45. Lojander J, Maasilta P, Partinen M, Brander PE, Salmi T, Lehtonen H. Nasal-CPAP, surgery, and conservative management for treatment of obstructive sleep apnea syndrome. A randomized study. Chest. 1996; 110(1):114-119

46. Lu CN, Friedman M, Lin HC, Bi KW, Hsueh TP, Chang HW et al. Alternative therapy for patients with obstructive sleep apnea/hypopnea syndrome: A 1-year, single-blind, randomized trial of Tui Na. Alternative Therapies in Health and Medicine. 2017; 23(4):16-24

47. MacKay S, Carney AS, Catcheside PG, Chai-Coetzer CL, Chia M, Cistulli PA et al. 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. JAMA. 2020; 134(12):1168-1179

48. Marzetti A, Tedaldi M, Passali FM. Preliminary findings from our experience in anterior palatoplasty for the treatment of obstructive sleep apnea. Clinical and Experimental Otorhinolaryngology. 2013; 6(1):18-22

49. McDaid C, Griffin S, Weatherly H, Duree K, van der Burgt M, van Hout S et al. Continuous positive airway pressure devices for the treatment of obstructive sleep apnoea-hypopnoea syndrome: a systematic review and economic analysis. Health Technology Assessment. 2009; 13(4)

50. Meccariello G, Cammaroto G, Montevecchi F, Hoff PT, Spector ME, Negm H et al. Transoral robotic surgery for the management of obstructive sleep apnea: a systematic review and meta-analysis. European Archives of Oto-Rhino-Laryngology. 2017; 274(2):647-653

51. Miller SC, Nguyen SA, Ong AA, Gillespie MB. Transoral robotic base of tongue reduction for obstructive sleep apnea: A systematic review and meta-analysis. Laryngoscope. 2017; 127(1):258-265

52. Mulholland GB, Jeffery CC, Ziai H, Hans V, Seikaly H, Pang KP et al. Multilevel palate and tongue base surgical treatment of obstructive sleep apnea: A systematic review and meta-analysis. Laryngoscope. 2019; 129(7):1712-1721

53. Murphey AW, Baker AB, Soose RJ, Padyha TA, Nguyen SA, Xiao CC et al. Upper airway stimulation for obstructive sleep apnea: The surgical learning curve. Laryngoscope. 2016; 126(2):501-506

54. Murphey AW, Kandl JA, Nguyen SA, Weber AC, Gillespie MB. The effect of glossectomy for obstructive sleep apnea: A systematic review and meta-analysis. Otolaryngology - Head & Neck Surgery. 2015; 153(3):334-342

55. Mwenge GB, Rombaux P, Lengele B, Rodenstein D. Hypoglossal nerve stimulation for obstructive sleep apnea. Progress in Neurological Surgery. 2015; 29:94-105

56. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual [Updated 2018]. London. National Institute for Health and Care Excellence, 2014. Available from: http://www.nice.org.uk/article/PMG20/chapter/1%20Introduction%20and%20overview

57. NCT. RCT of the effect of uvulopalatopharyngoplasty compared to expectancy in patients with obstructive sleep apnea. 2012. Available from: https://clinicaltrials.gov/show/NCT01659671 Last accessed: 27/07/2020.

58. Neruntarat C, Chantapant S. Palatal suspension and uvulopalatopharyngoplasty for obstructive sleep apnea. International Medical Journal. 2010; 17(3):213-217

59. Noller MW, Guilleminault C, Gouveia CJ, Mack D, Vivian C, Abdullatif J et al. Mandibular advancement for adult obstructive sleep apnea: A systematic review and meta-analysis. Journal of Cranio-Maxillo-Facial Surgery. 2017; 45(12):2035-2040

60. Pang KP, Pang EB, Pang KA, Rotenberg B. Anterior palatoplasty in the treatment of obstructive sleep apnoea - a systemic review. Acta Otorhinolaryngologica Italica. 2018; 38(1):1-6

61. Pang KP, Pang EB, Win MT, Pang KA, Woodson BT. Expansion sphincter pharyngoplasty for the treatment of OSA: a systemic review and meta-analysis. European Archives of Oto-Rhino-Laryngology. 2016; 273(9):2329-2333

62. Pang KP, Plaza G, Baptista JP, O'Connor Reina C, Chan YH, Pang KA et al. Palate surgery for obstructive sleep apnea: a 17-year meta-analysis. European Archives of Oto-Rhino-Laryngology. 2018; 275(7):1697-1707

63. Pengo MF, Xiao S, Ratneswaran C, Reed K, Shah N, Chen T et al. Randomised sham-controlled trial of transcutaneous electrical stimulation in obstructive sleep apnoea. Thorax. 2016; 71(10):923-931

64. Pietzsch JB, Liu S, Garner AM, Kezirian EJ, Strollo PJ. Long-term cost-effectiveness of upper airway stimulation for the treatment of obstructive sleep apnea: A model-based projection based on the STAR trial. Sleep. 2015; 38(5):735-744

65. Pietzsch JB, Richter AK, Randerath W, Steffen A, Liu S, Geisler BP et al. Clinical and economic benefits of upper airway stimulation for obstructive sleep apnea in a european setting. Respiration. 2019; 98(1):38-47

66. Reckley LK, Fernandez-Salvador C, Camacho M. The effect of tonsillectomy on obstructive sleep apnea: an overview of systematic reviews. Nature & Science of Sleep. 2018; 10:105-110

67. Rojo-Sanchis C, Almerich-Silla JM, Paredes-Gallardo V, Montiel-Company JM, Bellot-Arcis C. Impact of bimaxillary advancement surgery on the upper airway and on obstructive sleep apnea syndrome: A meta-analysis. Scientific Reports. 2018; 8(1):5756

68. Sharma S, Wormald JCR, Fishman JM, Andrews P, Kotecha BT. Rhinological interventions for obstructive sleep apnoea - a systematic review and descriptive meta-analysis. Journal of Laryngology and Otology. 2019; 133(3):168-176

69. Sommer JU, Maurer JT, Hormann K, Stuck BA. Randomized controlled trials in the surgical treatment of obstructive sleep apnea. HNO. 2012; 60(4):294-299

70. Sommer UJ, Heiser C, Gahleitner C, Herr RM, Hormann K, Maurer JT et al. Tonsillectomy with uvulopalatopharyngoplasty in obstructive sleep apnea. Deutsches Arzteblatt International. 2016; 113(1-02):1-8

71. Song SA, Wei JM, Buttram J, Tolisano AM, Chang ET, Liu SY et al. Hyoid surgery alone for obstructive sleep apnea: A systematic review and meta-analysis. Laryngoscope. 2016; 126(7):1702-1708

72. Steward DL, de Vries N, Froymovich O, Soose RJ, Gillespie MB, Maurer JT. Safety and effectiveness of upper airway stimulation via the hypoglossal nerve for obstructive sleep apnea. Otolaryngology - Head & Neck Surgery. 2013; 149(2 suppl):P140

73. Strollo P, Soose R, Strohl K. A randomized controlled study of upper airway stimulation therapy withdrawal effect in treatment of obstructive sleep apnea. European Respiratory Journal. 2013; 42(Suppl 57):p4975

74. Strollo PJ, Jr., Gillespie MB, Soose RJ, Maurer JT, de Vries N, Cornelius J et al. Upper airway stimulation for obstructive sleep apnea: Durability of the treatment effect at 18 months. Sleep. 2015; 38(10):1593-1598

75. Strollo PJ, Jr., Soose RJ, Maurer JT, de Vries N, Cornelius J, Froymovich O et al. Upper-airway stimulation for obstructive sleep apnea. New England Journal of Medicine. 2014; 370(2):139-149

76. Sundaram S, Lim J, Lasserson T. Surgery for obstructive sleep apnoea in adults. Cochrane Database of Systematic Reviews 2005, Issue 4. Art. No.: CD001004. DOI: 10.1002/14651858.CD001004.pub2.

77. Tang S, Wang Y, Qing J, Jiang Y, Lu X, Yao S et al. Exercise pharynx and genioglossus to treat obstructive sleep apnea and hypopnea syndrome. Journal of Clinical Otorhinolaryngology, Head, and Neck Surgery. 2013; 27(15):822-826

78. Tegelberg A, Wilhelmsson B, Walker-Engstrom ML, Ringqvist M, Andersson L, Krekmanov L et al. Effects and adverse events of a dental appliance for treatment of obstructive sleep apnoea. Swedish Dental Journal. 1999; 23(4):117-126

79. Terris DJ, Kunda LD, Gonella MC. Minimally invasive tongue base surgery for obstructive sleep apnoea. Journal of Laryngology and Otology. 2002; 116(9):716-721

80. Thomas A, Chavoya M, Terris D. Prospective randomised trial of tongue-based surgery for OSA. Annual Meeting of the American Academy of Otolaryngology - Head And Neck Surgery Foundation, Inc Otolaryngology- Head and Neck Surgery. 2002:117

81. Veer V, Yang WY, Green R, Kotecha B. Long-term safety and efficacy of radiofrequency ablation in the treatment of sleep disordered breathing: a meta-analysis. European Archives of Oto-Rhino-Laryngology. 2014; 271(11):2863-2870

82. Vicini C, Dallan I, Campanini A, De Vito A, Barbanti F, Giorgiomarrano G et al. Surgery vs ventilation in adult severe obstructive sleep apnea syndrome. American Journal of Otolaryngology. 2010; 31(1):14-20

83. Vicini C, Meccariello G, Montevecchi F, De Vito A, Frassineti S, Gobbi R et al. Effectiveness of barbed repositioning pharyngoplasty for the treatment of obstructive sleep apnea (OSA): a prospective randomized trial. Sleep & Breathing. 2020; 24(2):687-694

84. Volner K, Dunn B, Chang ET, Song SA, Liu SY, Brietzke SE et al. Transpalatal advancement pharyngoplasty for obstructive sleep apnea: a systematic review and meta-analysis. European Archives of Oto-Rhino-Laryngology. 2017; 274(3):1197-1203

85. Walker-Engstrom ML, Tegelberg A, Wilhelmsson B, Ringqvist I. 4-year follow-up of treatment with dental appliance or uvulopalatopharyngoplasty in patients with obstructive sleep apnea: a randomized study. Chest. 2002; 121(3):739-746

86. Walker-Engstrom ML, Tegelberg A, Wilhelmsson B, Ringqvist I. Four-year follow-up of treatment with dental appliance or UPPP in obstructive sleep apnea - a randomized study. Swedish Dental Journal. 2001; 25(4):181

87. Walker-Engstrom ML, Wilhelmsson B, Tegelberg A, Dimenas E, Ringqvist I. Quality of life assessment of treatment with dental appliance or UPPP in patients with mild to moderate obstructive sleep apnoea. A prospective randomized 1-year follow-up study. Journal of Sleep Research. 2000; 9(3):303-308

88. Wilhelmsson B, Tegelberg A, Walker-Engstrom ML, Ringqvist M, Andersson L, Krekmanov L et al. A prospective randomized study of a dental appliance compared with uvulopalatopharyngoplasty in the treatment of obstructive sleep apnoea. Acta Oto-Laryngologica. 1999; 119(4):503-509

89. 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. Otolaryngology - Head & Neck Surgery. 2003; 128(6):848-861

90. Wu J, Zhao G, Li Y, Zang H, Wang T, Wang D et al. Apnea-hypopnea index decreased significantly after nasal surgery for obstructive sleep apnea: A meta-analysis. Medicine. 2017; 96(5):e6008

91. Yang D, Zhou HF, Xie Y. Efficacy of uvulopalatopharyngoplasty combined with oral appliance in treatment of obstructive sleep apnea-hypopnea syndrome. Irish Journal of Medical Science. 2015; 184(2):329-334

92. Zhang M, Liu Y, Liu Y, Yu F, Yan S, Chen L et al. Effectiveness of oral appliances versus continuous positive airway pressure in treatment of OSA patients: An updated meta-analysis. Cranio. 2019; 37(6):347-364

93. Zhao MM, Tan XX, Ding N, Zhang XL. Comparison of efficacy between continuous positive airway pressure and renal artery sympathetic denervation by radiofrequency ablation in obstructive sleep apnea syndrome patients with hypertension. Zhonghua yi xue za zhi. 2013; 93(16):1234-1237

94. Zhu Z, Hofauer B, Wirth M, Hasselbacher K, Frohnhofen H, Heiser C et al. Selective upper airway stimulation in older patients. Respiratory Medicine. 2018; 140:77-81

Appendices

Appendix A: Review protocols

Table 10: Review protocol: surgery

ID	Field	Content
0.	PROSPERO registration number	Not registered
1.	Review title	Surgery
2.	Review question	What is the clinical and cost effectiveness of upper airway surgical interventions for people with obstructive sleep apnoea/hypopnoea syndrome?
3.	Objective	To determine the clinical and cost effectiveness of upper airway surgical interventions for people with obstructive sleep apnoea/hypopnoea syndrome.
4.	Searches	The following databases will be searched: • Cochrane Central Register of Controlled Trials (CENTRAL) • Cochrane Database of Systematic Reviews (CDSR) • Embase • MEDLINE • Epistemonikos Searches will be restricted by: • English language studies The searches may be re-run 6 weeks before the final committee meeting and further studies retrieved for inclusion if relevant. The full search strategies will be published in the final review.
5.	Condition or domain being studied	Obstructive sleep apnoea/hypopnoea syndrome is the most common form of sleep disordered breathing. The guideline will also cover obesity hypoventilation syndrome and COPD-OSAHS overlap syndrome (the coexistence of obstructive

		sleep apnoea/hypopnoea syndrome and chronic obstructive pulmonary disease).
6.	Population	Inclusion: People (16 and older) with OSAHS Population will be stratified by: Severity- Mild, moderate, severe (based on AHI/ODI) Treatment stage – failed previous OSAHS treatment vs general population Site of obstruction – multilevel vs single level Severity: Mild OSAHS: AHI >5 but <15 Moderate OSAHS: AHI >/= 15 but <30 Severe OSAHS: AHI >/= 30 When a mixed severity population is included the severity of the majority of the population will be used by taking the mean AHI of the patients included and the study will be downgraded for indirectness. Exclusion: Children and young people (under 16 years old) People with OHS and COPD-OSAHS overlap syndrome
7.	Intervention/Exposure/Test	• Any specific surgical intervention for OSAHS (nasal surgeries, uvulopalatopharyngoplasty, tonsillectomy, palatal implant, tongue reduction, genioglossus advancement, radiofrequency ablation, maxillomandibular advancement, hyoid suspension, upper airway stimulation)

		List of surgeries (from ERS) : • Nasal surgeries • functional rhinoplasty, septoplasty, turbinate reduction, polypectomy; • laryngeal surgeries • tracheotomy, epiglottoplasty • Tonsils • Tonsillectomy & tonsillotomy • Radiofrequency surgery of tonsils • Palate • Uvulopalatopharyngoplasty • Radiofrequency surgery of soft palate • Pillar method • Tongue base and hypopharnyx • Radiofrequency surgery of tongue base • Hyoid suspension • Partial tongue base resection et al • Maxfacs • Genioglossus advancement • Maxillo-mandibular advancement • Distraction osteogenesis

		• Multilevel surgery Expansion sphincter palatoplasty Barbed suture palatoplasty Hypoglossal nerve stimulation surgery for OSA Transoral robotic surgery for OSA
8.	Comparator/Reference standard/Confounding factors	• Any non-surgical intervention (positive airway pressure devices, positional modifiers, oral devices, • No intervention/usual care as defined in the studies (including lifestyle advice etc)
9.	Types of study to be included	RCTs Systematic review of RCTs Minimum duration of follow-up 1 months
10.	Other exclusion criteria	Exclusion: Non-English language studies. Abstracts will be excluded as it is expected there will be sufficient full text published studies available.
11.	Context	N/A
12.	Primary outcomes (critical outcomes)	• Generic or disease specific validated quality of life measures (continuous) • Mortality (dichotomous)
13.	Secondary outcomes (important outcomes)	• Sleepiness scores (continuous, e.g. Epworth) • Apnoea-Hypopnoea index (continuous) • Oxygen desaturation index (continuous) • CO2 control (continuous)

		• Permanent adverse effects (e.g. nerval dysfunction, open nasality, globus sensation, dichotomous) • Reversible adverse effects (e.g. pain, infection, secondary bleeding, dichotomous) • Driving outcomes (continuous) • Neurocognitive outcomes (continuous) • Impact on co-existing conditions: o HbA1c for diabetes (continuous) o Cardiovascular events for cardiovascular disease (dichotomous) o Systolic blood pressure for hypertension (continuous) Outcomes will be separated into short term (latest follow-up to 6 months) and long term (latest follow-up beyond 6 months)
14.	Data extraction (selection and coding)	EndNote will be used for reference management, sifting, citations and bibliographies. All references identified by the searches and from other sources will be screened for inclusion. 10% of the abstracts will be reviewed by two reviewers, with any disagreements resolved by discussion or, if necessary, a third independent reviewer. The full text of potentially eligible studies will be retrieved and will be assessed in line with the criteria outlined above. EviBASE will be used for data extraction.
15.	Risk of bias (quality) assessment	Risk of bias will be assessed using the appropriate checklist as described in Developing NICE guidelines: the manual. • Systematic reviews: Risk of Bias in Systematic Reviews (ROBIS) • Randomised Controlled Trial: Cochrane RoB (2.0) 10% of all evidence reviews are quality assured by a senior research fellow. This includes checking: • papers were included /excluded appropriately • a sample of the data extractions • correct methods are used to synthesise data

		• a sample of the risk of bias assessments Disagreements between the review authors over the risk of bias in particular studies will be resolved by discussion, with involvement of a third review author where necessary.
16.	Strategy for data synthesis	• Pairwise meta-analyses will be performed using Cochrane Review Manager (RevMan5). • GRADEpro will be used to assess the quality of evidence for each outcome, taking into account individual study quality and the meta-analysis results. The 4 main quality elements (risk of bias, indirectness, inconsistency and imprecision) will be appraised for each outcome. Publication bias is tested for when there are more than 5 studies for an outcome. The risk of bias across all available evidence was evaluated for each outcome using an adaptation of the 'Grading of Recommendations Assessment, Development and Evaluation (GRADE) toolbox' developed by the international GRADE working group http://www.gradeworkinggroup.org/ • Where meta-analysis is not possible, data will be presented, and quality assessed individually per outcome. • WinBUGS will be used for network meta-analysis, if possible, given the data identified. Heterogeneity between the studies in effect measures will be assessed using the I^2 statistic and visually inspected. An I^2 value greater than 50% will be considered indicative of substantial heterogeneity. Sensitivity analyses will be conducted based on pre-specified subgroups using stratified meta-analysis to explore the heterogeneity in effect estimates. If this does not explain the heterogeneity, the results will be presented pooled using random-effects.
17.	Analysis of sub-groups	• High risk occupational groups (for example heavy goods vehicle drivers) vs general population • Sleepiness – Epworth >9 vs Epworth 9 or less • Coexisting conditions – type 2 diabetes vs atrial fibrillation vs hypertension vs none • BMI – obese vs non-obese • Population – pre-selected vs failed CPAP vs general OSAHS • Type of surgery (example nasal, palate, tonsils)
18.	Type and method of review	☒ Intervention ☐ Diagnostic ☐ Prognostic ☐ Qualitative ☐ Epidemiologic

		☐ Service Delivery ☐ Other (please specify)
19.	Language	English
20.	Country	England
21.	Anticipated or actual start date	NA – not registered on PROSPERO
22.	Anticipated completion date	NA – not registered on PROSPERO
24.	Named contact	5a. Named contact National Guideline Centre 5b Named contact e-mail SleepApnoHypo@nice.org.uk 5e Organisational affiliation of the review National Institute for Health and Care Excellence (NICE) and the National Guideline Centre
25.	Review team members	From the National Guideline Centre: Carlos Sharpin, Guideline lead Sharangini Rajesh, Senior systematic reviewer Audrius Stonkus, Systematic reviewer Emtiyaz Chowdhury (until January 2020), Health economist David Wonderling, Head of health economics Agnes Cuyas, Information specialist (till December 2019) Jill Cobb, Information specialist
26.	Funding sources/sponsor	This systematic review is being completed by the National Guideline Centre which receives funding from NICE.
27.	Conflicts of interest	All guideline committee members and anyone who has direct input into NICE guidelines (including the evidence review team and expert witnesses) must declare any potential conflicts of interest in line with NICE's code of practice for declaring and dealing with conflicts of interest. Any relevant interests, or changes to interests, will also be declared publicly at the start of each guideline committee meeting. Before each meeting, any potential conflicts of interest will be considered by the guideline committee Chair and a senior member of the development team. Any decisions to exclude a person from all or part of a meeting will be documented. Any changes to a member's declaration of interests will be recorded in the minutes of the meeting. Declarations of interests will be published with the final guideline.
28.	Collaborators	Development of this systematic review will be overseen by an advisory committee who will use the review to inform the

		development of evidence-based recommendations in line with section 3 of Developing NICE guidelines: the manual. Members of the guideline committee are available on the NICE website: https://www.nice.org.uk/guidance/indevelopment/gid-ng10098
29.	Other registration details	NA – not registered
30.	Reference/URL for published protocol	NA – not registered
31.	Dissemination plans	NICE may use a range of different methods to raise awareness of the guideline. These include standard approaches such as: • notifying registered stakeholders of publication • publicising the guideline through NICE's newsletter and alerts • issuing a press release or briefing as appropriate, posting news articles on the NICE website, using social media channels, and publicising the guideline within NICE.
32.	Keywords	-
33.	Details of existing review of same topic by same authors	N/A
35..	Additional information	N/A
36.	Details of final publication	www.nice.org.uk

Table 11: Health economic review protocol

Review question	All questions – health economic evidence
Objectives	To identify health economic studies relevant to any of the review questions.
Search criteria	• Populations, interventions and comparators must be as specified in the clinical review protocol above. • Studies must be of a relevant health economic study design (cost–utility analysis, cost-effectiveness analysis, cost–benefit analysis, cost–consequences analysis, comparative cost analysis). • Studies must not be a letter, editorial or commentary, or a review of health economic evaluations. (Recent reviews will be ordered although not reviewed. The bibliographies will be checked for relevant studies, which will then be ordered.)

	• Unpublished reports will not be considered unless submitted as part of a call for evidence. • Studies must be in English.
Search strategy	A health economic study search will be undertaken using population-specific terms and a health economic study filter – see appendix B below.
Review strategy	Studies not meeting any of the search criteria above will be excluded. Studies published before 2003, abstract-only studies and studies from non-OECD countries or the USA will also be excluded. Each remaining study will be assessed for applicability and methodological limitations using the NICE economic evaluation checklist which can be found in appendix H of Developing NICE guidelines: the manual (2014).[56] **Inclusion and exclusion criteria** • If a study is rated as both 'Directly applicable' and with 'Minor limitations' then it will be included in the guideline. A health economic evidence table will be completed, and it will be included in the health economic evidence profile. • If a study is rated as either 'Not applicable' or with 'Very serious limitations' then it will usually be excluded from the guideline. If it is excluded, then a health economic evidence table will not be completed, and it will not be included in the health economic evidence profile. • If a study is rated as 'Partially applicable', with 'Potentially serious limitations' or both then there is discretion over whether it should be included. **Where there is discretion** The health economist will make a decision based on the relative applicability and quality of the available evidence for that question, in discussion with the guideline committee if required. The ultimate aim is to include health economic studies that are helpful for decision-making in the context of the guideline and the current NHS setting. If several studies are considered of sufficiently high applicability and methodological quality that they could all be included, then the health economist, in discussion with the committee if required, may decide to include only the most applicable studies and to selectively exclude the remaining studies. All studies excluded on the basis of applicability or methodological limitations will be listed with explanation in the excluded health economic studies appendix below. The health economist will be guided by the following hierarchies. *Setting:* • UK NHS (most applicable). • OECD countries with predominantly public health insurance systems (for example, France, Germany, Sweden). • OECD countries with predominantly private health insurance systems (for example, Switzerland). • Studies set in non-OECD countries or in the USA will be excluded before being assessed for applicability and methodological limitations. *Health economic study type:* • Cost–utility analysis (most applicable). • Other type of full economic evaluation (cost–benefit analysis, cost-effectiveness analysis, cost–consequences analysis). • Comparative cost analysis. • Non-comparative cost analyses including cost-of-illness studies will be excluded before being assessed for applicability and methodological limitations. *Year of analysis:* • The more recent the study, the more applicable it will be.

	• Studies published in 2003 or later but that depend on unit costs and resource data entirely or predominantly from before 2003 will be rated as 'Not applicable'. • Studies published before 2003 will be excluded before being assessed for applicability and methodological limitations. *Quality and relevance of effectiveness data used in the health economic analysis:* • The more closely the clinical effectiveness data used in the health economic analysis match with the outcomes of the studies included in the clinical review the more useful the analysis will be for decision-making in the guideline.

Appendix B: Literature search strategies

Sleep Apnoea search strategy 12 surgery

This literature search strategy was used for the following reviews;

- What is the clinical and cost effectiveness of upper airway surgical interventions for people with obstructive sleep apnoea/hypopnoea syndrome?

The literature searches for this review are detailed below and complied with the methodology outlined in Developing NICE guidelines: the manual.[56]

For more information, please see the Methods Report published as part of the accompanying documents for this guideline.

B.1 Clinical search literature search strategy

Searches were constructed using a PICO framework where population (P) terms were combined with Intervention (I) and in some cases Comparison (C) terms. Outcomes (O) are rarely used in search strategies for interventions as these concepts may not be well described in title, abstract or indexes and therefore difficult to retrieve. Search filters were applied to the search where appropriate.

Table 12: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline (OVID)	1946 – 6 July 2020	Exclusions Randomised controlled trials Systematic review studies
Embase (OVID)	1974 – 6 July 2020	Exclusions Randomised controlled trials Systematic review studies
The Cochrane Library (Wiley)	Cochrane Reviews to 2020 Issue 7 of 12 CENTRAL to 2020 Issue 7 of 12	None
Epistemonikos (Epistemonikos Foundation)	Inception – 29 November 2018	None

Medline (Ovid) search terms

1.	exp Sleep Apnea Syndromes/
2.	(sleep* adj4 (apn?ea* or hypopn?ea*)).ti,ab.
3.	(sleep* adj4 disorder* adj4 breath*).ti,ab.
4.	(OSAHS or OSA or OSAS).ti,ab.
5.	(obes* adj3 hypoventil*).ti,ab.
6.	pickwick*.ti,ab.
7.	or/1-6
8.	limit 7 to English language
9.	letter/
10.	editorial/
11.	news/
12.	exp historical article/
13.	Anecdotes as Topic/
14.	comment/
15.	case report/

16.	(letter or comment*).ti.
17.	or/9-16
18.	randomized controlled trial/ or random*.ti,ab.
19.	17 not 18
20.	animals/ not humans/
21.	exp Animals, Laboratory/
22.	exp Animal Experimentation/
23.	exp Models, Animal/
24.	exp Rodentia/
25.	(rat or rats or mouse or mice).ti.
26.	or/19-25
27.	8 not 26
28.	otorhinolaryngologic surgical procedures/ or adenoidectomy/ or laryngoplasty/ or laryngoscopy/ or nasal surgical procedures/ or rhinoplasty/ or tonsillectomy/ or tracheostomy/ or tracheotomy/
29.	exp Surgery, oral/ or Oral Surgical Procedures/
30.	Tongue/su or Nose/su or Mouth/su or Palate/su or Paranasal Sinuses/su
31.	((surg* or operat*) adj3 (pharyn* or nasal* or intranasal* or sinonasal* or paranasal* or turbinate* or palate* or palatal or uvula* or upper-airway* or upperairway* or tongue* or jaw* or adenoid* or tonsil* or endoscopic*)).ti,ab.
32.	(septoplast* or rhinoplast* or polypectom* or turbinectom*).ti,ab.
33.	((surg* or operat*) adj3 (oral* or transoral or trans-oral or oropharyng* or nasopharyng* or otorhinolaryng* or maxillofacial or maxillo-facial or hypopharyn* or facial)).ti,ab.
34.	(uvulopalatopharyngoplast* or uppp or uvpp or upp or uvulopalatal or upf or palatoplast* or pharyngoplast* or palatopharyngoplast* or ppp or uvulopalatoplast* or laup).ti,ab.
35.	(tonsillectom* or tonsillotom* or adenotonsillectom* or adenoidectom* or orthognathic).ti,ab.
36.	(palat* adj3 (advanc* or implant*)).ti,ab.
37.	(intrapalat* adj3 resection*).ti,ab.
38.	((tori or torus) adj3 (excis* or remov* or reduc* or surg*)).ti,ab.
39.	(sagittal adj3 (ramus or osteotom*)).ti,ab.
40.	Glossectomy/ or Osteotomy, Le Fort/ or Osteotomy, Sagittal Split Ramus/
41.	(glossectom* or lingualplast*).ti,ab.
42.	((hyoid or tongue*) adj3 (myotom* or suspens* or advanc* or reduc* or ablat* or stabili*)).ti,ab.
43.	(TCRFTA or tissue ablat*).ti,ab.
44.	(((genioglossus or genioglossal) adj3 advanc*) or genial tubercle advanc*).ti,ab.
45.	Mandibular Advancement/
46.	MMA.ti,ab.
47.	((maxillomandibular or maxillo-mandibular or maxillary or mandibular or bimaxillary or bi-maxillary) adj3 (advanc* or osteotom* or surg* or operat*)).ti,ab.
48.	(tracheotom* or tracheostom* or minitracheostom* or mini-tracheostom* or epiglottoplast*).ti,ab.
49.	Radiofrequency ablation/
50.	((radio frequency or radiofrequency) adj3 ablat*).ti,ab.
51.	((upper airway or upperairway or hypoglossal) adj3 stimulat*).ti,ab.
52.	or/28-51

53.	27 and 52
54.	randomized controlled trial.pt.
55.	controlled clinical trial.pt.
56.	randomi#ed.ti,ab.
57.	placebo.ab.
58.	randomly.ti,ab.
59.	Clinical Trials as topic.sh.
60.	trial.ti.
61.	or/54-60
62.	Meta-Analysis/
63.	exp Meta-Analysis as Topic/
64.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
65.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
66.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
67.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
68.	(search* adj4 literature).ab.
69.	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or science citation index or bids or cancerlit).ab.
70.	cochrane.jw.
71.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
72.	or/62-71
73.	53 and (61 or 72)

Embase (Ovid) search terms

1.	exp Sleep Disordered Breathing/
2.	(sleep* adj4 (apn?ea* or hypopn?ea*)).ti,ab.
3.	(sleep* adj4 disorder* adj4 breath*).ti,ab.
4.	(OSAHS or OSA or OSAS).ti,ab.
5.	(obes* adj3 hypoventil*).ti,ab.
6.	pickwick*.ti,ab.
7.	or/1-6
8.	limit 7 to English language
9.	letter.pt. or letter/
10.	note.pt.
11.	editorial.pt.
12.	case report/ or case study/
13.	(letter or comment*).ti.
14.	or/9-13
15.	randomized controlled trial/ or random*.ti,ab.
16.	14 not 15
17.	animal/ not human/
18.	nonhuman/
19.	exp Animal Experiment/
20.	exp Experimental Animal/

21.	animal model/
22.	exp Rodent/
23.	(rat or rats or mouse or mice).ti.
24.	or/16-23
25.	8 not 24
26.	ear nose throat surgery/ or endoscopic sinus surgery/ or nose surgery/ or nose reconstruction/ or throat surgery/ or adenoidectomy/ or laryngoplasty/ or laryngoscopy/ or tonsillectomy/ or tracheostomy/ or tracheotomy/
27.	oral surgery/ or orthognathic surgery/
28.	Tongue/su or Nose/su or Mouth/su or Palate/su or Paranasal Sinuses/su
29.	((surg* or operat*) adj3 (pharyn* or nasal* or intranasal* or sinonasal* or paranasal* or turbinate* or palate* or palatal or uvula* or upper-airway* or upperairway* or tongue* or jaw* or adenoid* or tonsil* or endoscopic*)).ti,ab.
30.	(septoplast* or rhinoplast* or polypectom* or turbinectom*).ti,ab.
31.	((surg* or operat*) adj3 (oral* or transoral or trans-oral or oropharyng* or nasopharyng* or otorhinolaryng* or maxillofacial or maxillo-facial or hypopharyn* or facial)).ti,ab.
32.	(uvulopalatopharyngoplast* or uppp or uvpp or upp or uvulopalatal or upf or palatoplast* or pharyngoplast* or palatopharyngoplast* or ppp or uvulopalatoplast* or laup).ti,ab.
33.	(tonsillectom* or tonsillotom* or adenotonsillectom* or adenoidectom* or orthognathic).ti,ab.
34.	(palat* adj3 (advanc* or implant*)).ti,ab.
35.	(intrapalat* adj3 resection*).ti,ab.
36.	((tori or torus) adj3 (excis* or remov* or reduc* or surg*)).ti,ab.
37.	(sagittal adj3 (ramus or osteotom*)).ti,ab.
38.	glossectomy/ or Le Fort osteotomy/ or sagittal split ramal osteotomy/
39.	(glossectom* or lingualplast*).ti,ab.
40.	((hyoid or tongue*) adj3 (myotom* or suspens* or advanc* or reduc* or ablat* or stabili*)).ti,ab.
41.	(TCRFTA or tissue ablat*).ti,ab.
42.	(((genioglossus or genioglossal) adj3 advanc*) or genial tubercle advanc*).ti,ab.
43.	mandibular advancement/
44.	MMA.ti,ab.
45.	((maxillomandibular or maxillo-mandibular or maxillary or mandibular or bimaxillary or bi-maxillary) adj3 (advanc* or osteotom* or surg* or operat*)).ti,ab.
46.	(tracheotom* or tracheostom* or minitracheostom* or mini-tracheostom* or epiglottoplast*).ti,ab.
47.	radiofrequency ablation/
48.	((radio frequency or radiofrequency) adj3 ablat*).ti,ab.
49.	((upper airway or upperairway or hypoglossal) adj3 stimulat*).ti,ab.
50.	or/26-49
51.	25 and 50
52.	random*.ti,ab.
53.	factorial*.ti,ab.
54.	(crossover* or cross over*).ti,ab.
55.	((doubl* or singl*) adj blind*).ti,ab.
56.	(assign* or allocat* or volunteer* or placebo*).ti,ab.
57.	crossover procedure/

58.	single blind procedure/
59.	randomized controlled trial/
60.	double blind procedure/
61.	or/52-60
62.	systematic review/
63.	meta-analysis/
64.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
65.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
66.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
67.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
68.	(search* adj4 literature).ab.
69.	(medline or pubmed or cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinahl or science citation index or bids or cancerlit).ab.
70.	cochrane.jw.
71.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
72.	or/62-71
73.	51 and (61 or 72)

Cochrane Library (Wiley) search terms

#1.	MeSH descriptor: [Sleep Apnea Syndromes] explode all trees
#2.	(sleep* near/4 (apnea* or apnoea* or hypopnea* or hypopnoea*)):ti,ab
#3.	(sleep* near/4 disorder* near/4 breath*):ti,ab
#4.	(OSAHS or OSA or OSAS):ti,ab
#5.	(obes* near/3 hypoventil*):ti,ab
#6.	pickwick*:ti,ab
#7.	(OR #1-#6)
#8.	MeSH descriptor: [Otorhinolaryngologic Surgical Procedures] this term only
#9.	MeSH descriptor: [Adenoidectomy] this term only
#10.	MeSH descriptor: [Laryngoplasty] this term only
#11.	MeSH descriptor: [Laryngoscopy] this term only
#12.	MeSH descriptor: [Nasal Surgical Procedures] this term only
#13.	MeSH descriptor: [Rhinoplasty] this term only
#14.	MeSH descriptor: [Tonsillectomy] this term only
#15.	MeSH descriptor: [Tracheostomy] this term only
#16.	MeSH descriptor: [Tracheotomy] this term only
#17.	MeSH descriptor: [Surgery, Oral] explode all trees
#18.	MeSH descriptor: [Oral Surgical Procedures] this term only
#19.	MeSH descriptor: [Tongue] this term only and with qualifier(s): [surgery - SU]
#20.	MeSH descriptor: [Nose] this term only and with qualifier(s): [surgery - SU]
#21.	MeSH descriptor: [Mouth] this term only and with qualifier(s): [surgery - SU]
#22.	MeSH descriptor: [Palate] this term only and with qualifier(s): [surgery - SU]
#23.	MeSH descriptor: [Paranasal Sinuses] this term only and with qualifier(s): [surgery - SU]

#24.	((surg* or operat*) NEAR/3 (pharyn* or nasal* or intranasal* or sinonasal* or paranasal* or turbinate* or palate* or palatal or uvula* or upper-airway* or upperairway* or tongue* or jaw* or adenoid* or tonsil* or endoscopic*)):ti,ab
#25.	(septoplast* or rhinoplast* or polypectom* or turbinectom*):ti,ab
#26.	((surg* or operat*) NEAR/3 (oral* or transoral or trans-oral or oropharyng* or nasopharyng* or otorhinolaryng* or maxillofacial or maxillo-facial or hypopharyn* or facial)):ti,ab
#27.	(uvulopalatopharyngoplast* or uppp or uvpp or upp or uvulopalatal or upf or palatoplast* or pharyngoplast* or palatopharyngoplast* or ppp or uvulopalatoplast* or laup):ti,ab
#28.	(tonsillectom* or tonsillotom* or adenotonsillectom* or adenoidectom* or orthognathic):ti,ab
#29.	(palat* NEAR/3 (advanc* or implant*)):ti,ab
#30.	(intrapalat* NEAR/3 resection*):ti,ab
#31.	((tori or torus) NEAR/3 (excis* or remov* or reduc* or surg*)):ti,ab
#32.	(sagittal NEAR/3 (ramus or osteotom*)):ti,ab
#33.	MeSH descriptor: [Glossectomy] this term only
#34.	MeSH descriptor: [Osteotomy, Le Fort] this term only
#35.	MeSH descriptor: [Osteotomy, Sagittal Split Ramus] this term only
#36.	(glossectom* or lingualplast*):ti,ab
#37.	((hyoid or tongue*) NEAR/3 (myotom* or suspens* or advanc* or reduc* or ablat* or stabili*)):ti,ab
#38.	(TCRFTA or tissue ablat*):ti,ab
#39.	(((genioglossus or genioglossal) NEAR/3 advanc*) or genial tubercle advanc*):ti,ab
#40.	MeSH descriptor: [Mandibular Advancement] this term only
#41.	MMA:ti,ab
#42.	((maxillomandibular or maxillo-mandibular or maxillary or mandibular or bimaxillary or bi-maxillary) NEAR/3 (advanc* or osteotom* or surg* or operat*)):ti,ab
#43.	(tracheotom* or tracheostom* or minitracheostom* or mini-tracheostom* or epiglottoplast*):ti,ab
#44.	MeSH descriptor: [Radiofrequency Ablation] this term only
#45.	((radio frequency or radiofrequency) near/3 ablat*):ti,ab
#46.	((upper airway or upperairway or hypoglossal) NEAR/3 stimulat*):ti,ab
#47.	(OR #8-#46)
#48.	#7 AND #47

Epistemonikos search terms

1.	((title:((sleep apnea syndromes) OR (sleep* AND (apn?ea* OR hypopn?ea*)) OR (sleep* AND (apn?ea* OR hypopn?ea*)) OR (sleep* AND (disorder* OR breath*)) OR (OSAHS OR OSA OR OSAS) OR (obes* AND hypoventil*) OR pickwick*) OR abstract:((sleep apnea syndromes) OR (sleep* AND (apn?ea* OR hypopn?ea*)) OR (sleep* AND (apn?ea* OR hypopn?ea*)) OR (sleep* AND (disorder* OR breath*)) OR (OSAHS OR OSA OR OSAS) OR (obes* AND hypoventil*) OR pickwick*)))

B.2 Health Economics literature search strategy

Health economic evidence was identified by conducting a broad search relating to sleep apnoea population in NHS Economic Evaluation Database (NHS EED – this ceased to be updated after March 2015) and the Health Technology Assessment database (HTA – this ceased to be updated after March 2018) with no date restrictions. NHS EED and HTA

databases are hosted by the Centre for Research and Dissemination (CRD). Additional searches were run on Medline and Embase for health economics and quality of life studies.

B.2.1 Health economic studies strategy

Table 13: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline	2014 – 6 July 2020	Exclusions Health economics studies
Embase	2014 – 6 July 2020	Exclusions Health economics studies
Centre for Research and Dissemination (CRD)	HTA - Inception – 31 March 2018 NHSEED - Inception to March 2015	None

Medline (Ovid) search terms

	exp Sleep Apnea Syndromes/
1.	(sleep* adj4 (apn?ea* or hypopn?ea*)).ti,ab.
2.	(sleep* adj4 disorder* adj4 breath*).ti,ab.
3.	(OSAHS or OSA or OSAS).ti,ab.
4.	(obes* adj3 hypoventil*).ti,ab.
5.	pickwick*.ti,ab.
6.	or/1-6
7.	limit 7 to English language
8.	letter/
9.	editorial/
10.	news/
11.	exp historical article/
12.	Anecdotes as Topic/
13.	comment/
14.	case report/
15.	(letter or comment*).ti.
16.	or/9-16
17.	randomized controlled trial/ or random*.ti,ab.
18.	17 not 18
19.	animals/ not humans/
20.	exp Animals, Laboratory/
21.	exp Animal Experimentation/
22.	exp Models, Animal/
23.	exp Rodentia/
24.	(rat or rats or mouse or mice).ti.
25.	or/19-25
26.	8 not 26
27.	Economics/

28.	Value of life/
29.	exp "Costs and Cost Analysis"/
30.	exp Economics, Hospital/
31.	exp Economics, Medical/
32.	Economics, Nursing/
33.	Economics, Pharmaceutical/
34.	exp "Fees and Charges"/
35.	exp Budgets/
36.	budget*.ti,ab.
37.	cost*.ti.
38.	(economic* or pharmaco?economic*).ti.
39.	(price* or pricing*).ti,ab.
40.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
41.	(financ* or fee or fees).ti,ab.
42.	(value adj2 (money or monetary)).ti,ab.
43.	or/28-43
44.	27 and 44

Embase (Ovid) search terms

1.	exp Sleep Disordered Breathing/
2.	(sleep* adj4 (apn?ea* or hypopn?ea*)).ti,ab.
3.	(sleep* adj4 disorder* adj4 breath*).ti,ab.
4.	(OSAHS or OSA or OSAS).ti,ab.
5.	(obes* adj3 hypoventil*).ti,ab.
6.	pickwick*.ti,ab.
7.	or/1-6
8.	limit 7 to English language
9.	letter.pt. or letter/
10.	note.pt.
11.	editorial.pt.
12.	case report/ or case study/
13.	(letter or comment*).ti.
14.	or/9-13
15.	randomized controlled trial/ or random*.ti,ab.
16.	14 not 15
17.	animal/ not human/
18.	nonhuman/
19.	exp Animal Experiment/
20.	exp Experimental Animal/
21.	animal model/
22.	exp Rodent/
23.	(rat or rats or mouse or mice).ti.
24.	or/16-23

25.	8 not 24
26.	health economics/
27.	exp economic evaluation/
28.	exp health care cost/
29.	exp fee/
30.	budget/
31.	funding/
32.	budget*.ti,ab.
33.	cost*.ti.
34.	(economic* or pharmaco?economic*).ti.
35.	(price* or pricing*).ti,ab.
36.	(cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab.
37.	(financ* or fee or fees).ti,ab.
38.	(value adj2 (money or monetary)).ti,ab.
39.	or/26-38
40.	25 and 39

NHS EED and HTA (CRD) search terms

#1.	MeSH DESCRIPTOR Sleep Apnea Syndromes EXPLODE ALL TREES
#2.	(sleep* adj4 (apn?ea* or hypopn?ea*))
#3.	(sleep* adj4 disorder* adj4 breath*)
#4.	(OSAHS or OSA or OSAS)
#5.	(obes* adj3 hypoventil*)
#6.	(pickwick*)
#7.	#1 OR #2 OR #3 OR #4 OR #5 OR #6

B.2.2 Quality of life studies strategy

Table 14: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline	1946 – 26 November 2019	Exclusions Quality of life studies
Embase	1974 – 26 November 2019	Exclusions Quality of life studies

Medline (Ovid) search terms

1.	exp Sleep Apnea Syndromes/
2.	(sleep* adj4 (apn?ea* or hypopn?ea*)).ti,ab.
3.	(sleep* adj4 disorder* adj4 breath*).ti,ab.
4.	(OSAHS or OSA or OSAS).ti,ab.
5.	(obes* adj3 hypoventil*).ti,ab.
6.	pickwick*.ti,ab.
7.	or/1-6
8.	limit 7 to English language
9.	letter/

10.	editorial/
11.	news/
12.	exp historical article/
13.	Anecdotes as Topic/
14.	comment/
15.	case report/
16.	(letter or comment*).ti.
17.	or/9-16
18.	randomized controlled trial/ or random*.ti,ab.
19.	17 not 18
20.	animals/ not humans/
21.	exp Animals, Laboratory/
22.	exp Animal Experimentation/
23.	exp Models, Animal/
24.	exp Rodentia/
25.	(rat or rats or mouse or mice).ti.
26.	or/19-25
27.	8 not 26
28.	quality-adjusted life years/
29.	sickness impact profile/
30.	(quality adj2 (wellbeing or well being)).ti,ab.
31.	sickness impact profile.ti,ab.
32.	disability adjusted life.ti,ab.
33.	(qal* or qtime* or qwb* or daly*).ti,ab.
34.	(euroqol* or eq5d* or eq 5*).ti,ab.
35.	(qol* or hql* or hqol* or h qol* or hrqol* or hr qol*).ti,ab.
36.	(health utility* or utility score* or disutilit* or utility value*).ti,ab.
37.	(hui or hui1 or hui2 or hui3).ti,ab.
38.	(health* year* equivalent* or hye or hyes).ti,ab.
39.	discrete choice*.ti,ab.
40.	rosser.ti,ab.
41.	(willingness to pay or time tradeoff or time trade off or tto or standard gamble*).ti,ab.
42.	(sf36* or sf 36* or short form 36* or shortform 36* or shortform36*).ti,ab.
43.	(sf20 or sf 20 or short form 20 or shortform 20 or shortform20).ti,ab.
44.	(sf12* or sf 12* or short form 12* or shortform 12* or shortform12*).ti,ab.
45.	(sf8* or sf 8* or short form 8* or shortform 8* or shortform8*).ti,ab.
46.	(sf6* or sf 6* or short form 6* or shortform 6* or shortform6*).ti,ab.
47.	or/28-46
48.	27 and 47

Embase (Ovid) search terms

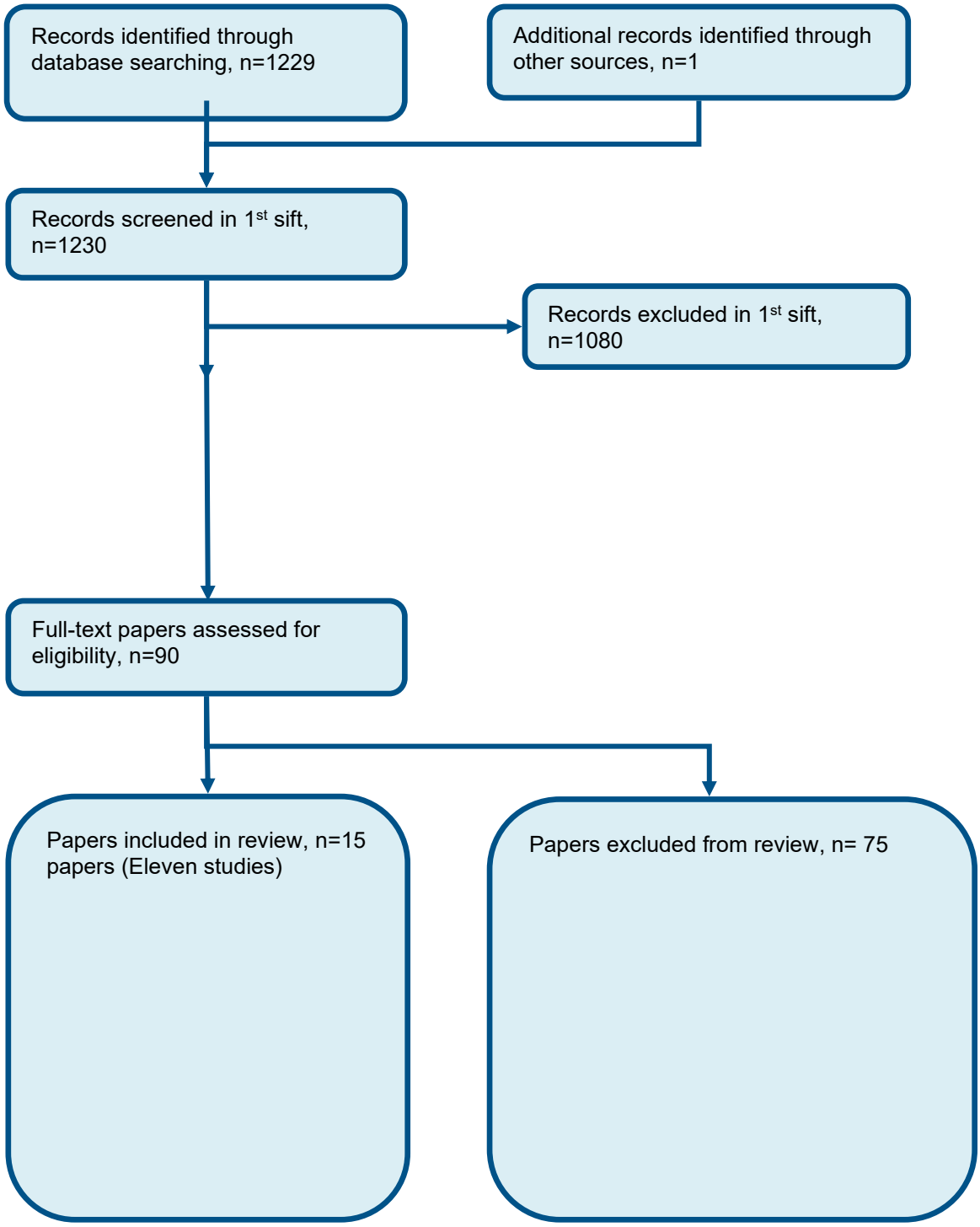
1.	exp Sleep Disordered Breathing/

2.	(sleep* adj4 (apn?ea* or hypopn?ea*)).ti,ab.
3.	(sleep* adj4 disorder* adj4 breath*).ti,ab.
4.	(OSAHS or OSA or OSAS).ti,ab.
5.	(obes* adj3 hypoventil*).ti,ab.
6.	pickwick*.ti,ab.
7.	or/1-6
8.	limit 7 to English language
9.	letter.pt. or letter/
10.	note.pt.
11.	editorial.pt.
12.	case report/ or case study/
13.	(letter or comment*).ti.
14.	or/9-13
15.	randomized controlled trial/ or random*.ti,ab.
16.	14 not 15
17.	animal/ not human/
18.	nonhuman/
19.	exp Animal Experiment/
20.	exp Experimental Animal/
21.	animal model/
22.	exp Rodent/
23.	(rat or rats or mouse or mice).ti.
24.	or/16-23
25.	8 not 24
26.	quality adjusted life year/
27.	"quality of life index"/
28.	short form 12/ or short form 20/ or short form 36/ or short form 8/
29.	sickness impact profile/
30.	(quality adj2 (wellbeing or well being)).ti,ab.
31.	sickness impact profile.ti,ab.
32.	disability adjusted life.ti,ab.
33.	(qal* or qtime* or qwb* or daly*).ti,ab.
34.	(euroqol* or eq5d* or eq 5*).ti,ab.
35.	(qol* or hql* or hqol* or h qol* or hrqol* or hr qol*).ti,ab.
36.	(health utility* or utility score* or disutilit* or utility value*).ti,ab.
37.	(hui or hui1 or hui2 or hui3).ti,ab.
38.	(health* year* equivalent* or hye or hyes).ti,ab.
39.	discrete choice*.ti,ab.
40.	rosser.ti,ab.
41.	(willingness to pay or time tradeoff or time trade off or tto or standard gamble*).ti,ab.
42.	(sf36* or sf 36* or short form 36* or shortform 36* or shortform36*).ti,ab.
43.	(sf20 or sf 20 or short form 20 or shortform 20 or shortform20).ti,ab.
44.	(sf12* or sf 12* or short form 12* or shortform 12* or shortform12*).ti,ab.

45.	(sf8* or sf 8* or short form 8* or shortform 8* or shortform8*).ti,ab.
46.	(sf6* or sf 6* or short form 6* or shortform 6* or shortform6*).ti,ab.
47.	or/26-46
48.	25 and 47

Appendix C: Clinical evidence selection

Figure 1: Flow chart of clinical study selection for the review of surgery



Appendix D: Clinical evidence tables

Study	Ferguson 2003[23]
Study type	RCT (Patient randomised; Crossover: N/A)
Number of studies (number of participants)	1 (n=46)
Countries and setting	Conducted in Canada; Setting: Department of Medicine, University of Western Ontario; and the Department of Otolaryngology, University of Western Ontario, London, Canada.
Line of therapy	Unclear
Duration of study	Follow up (post intervention): 6 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis.
Stratum	Moderate-severe
Subgroup analysis within study	Not applicable
Inclusion criteria	Patients who had mild OSA (apnoea/hypopnea index[AHI] 10.1–25) and complained of loud snoring were recruited.
Exclusion criteria	Exclusion criteria was the following; Patients less than 18 years of age, Patients with a diagnosis of another sleep disorder in addition to OSA (e.g., periodic limb movement disorder), Patients with previous palatal surgery for OSA, Patients in whom oral steroids would be contraindicated (e.g., diabetes), Patients on anticoagulants (e.g., coumadin or aspirin) that could not be safely discontinued

Study	Ferguson 2003[23]
	in the perioperative period.
Age, gender and ethnicity	Age - Mean (SD): surgery group - 43.7 (6.3) control group - 45.3 (9.5). Gender (M: F): 35/11. Ethnicity: unclear
Further population details	1. BMI: BMI of 30 2 kg/m^2 or more. mean AHI- AHI: LAUP: 18.6 (SD 4.3); Control 16.1 (SD 4)
Extra comments	-
Indirectness of population	Serious indirectness: patients included with AHI ranging from 10-25
Interventions	(n=21) Intervention 1: Surgery. The LAUP procedure was repeated at 1- to 2-month intervals. The end points for the LAUP procedure were (1) when the snoring was reported to be significantly reduced or eliminated, 2) no more tissue could be safely removed or 3) the patient refused further surgery. Subjects received preoperative pain medication. Topical anaesthesia was applied and lidocaine was injected into the uvula and soft palate. A series of full thickness vertical trenches were created with the CO2 laser on the free edge of the soft palate on either side of the uvula. The uvula was shortened and thinned and the soft palate was also reduced. The tonsils were not treated. Subjects received post-operative anti biotics, analgesics, anti-inflammatories and dilute hydrogen peroxide gargles for 7 days. . Duration 1 to 2-month intervals. Concurrent medication/care: Blood pressure was measured at baseline and at each follow-up visit. The blood pressure was measured in the patient after sitting for at least 10 minutes. All of the visits were in the daytime, although the visits varied between the morning and afternoon. Neck circumference was recorded at baseline and at each follow-up visit. It was measured in centimetres at the level of the cricothyroid membrane. Questionnaires, scales, and the polysomnogram were repeated 3 months after the last LAUP procedure or 6 months after baseline in the control group. Further details: 1. Intervention type: Electronic (laser). 2. Type of surgery: palatal (n=25) Intervention 2: No intervention - Usual care (lifestyle advice etc.). The control subjects were not offered any therapy but were offered LAUP at the end of the study.

Study	Ferguson 2003[23]
	. Duration 6 months. Concurrent medication/care: Blood pressure was measured at baseline and at each follow-up visit. The blood pressure was measured in the patient after sitting for at least 10 minutes. All of the visits were in the daytime, although the visits varied between the morning and afternoon. Neck circumference was recorded at baseline and at each follow-up visit. It was measured in centimetres at the level of the cricothyroid membrane. Questionnaires, scales, and the polysomnogram were repeated 3 months after the last LAUP procedure or 6 months after baseline in the control group.
Funding	Academic or government funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SURGERY versus USUAL CARE (LIFESTYLE ADVICE ETC)

Protocol outcome 1: Quality of life at >1 month
- Actual outcome: SAQLI at 8 months in controls and 15 months in surgery group; Group 1: mean 4.6 (SD 0.9); n=21,
Risk of bias: All domain - Very high, Selection - Low, Blinding - Very high, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of mild and moderate OSA included ; Blinding details: Follow up periods differed between groups; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: unrelated illness

Protocol outcome 2: Sleepiness score at >1 month
- Actual outcome: Epworth sleepiness scale at 8 months in controls and 15 months in surgery group; Group 1: mean 9.3 (SD 3.8); n=21, Group 2: mean 10.8 (SD 9.3); n=24
Risk of bias: All domain - Very high, Selection - Low, Blinding - Very high, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of mild and moderate OSA included ; Blinding details: Follow up periods differed between groups; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: unrelated illness

Protocol outcome 3: AHI/RDI at >1 month
- Actual outcome: AHI at 8 months in controls and 15 months in surgery group; Group 1: mean 14.7 (SD 7.5); n=21, Group 2: mean 22.7 (SD 15.2); n=24
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of mild and moderate OSA included ; Blinding details: Follow up periods differed between groups; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: unrelated illness

Study	Ferguson 2003[23]

Protocol outcome 4: Adverse effects of treatment at >1 month
- Actual outcome: infection
at 8 months in controls and 15 months in surgery group; Group 1: 4/21, Group 2: 0/24
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of mild and moderate OSA included ; Blinding details: Follow up periods differed between groups; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: unrelated illness
- Actual outcome: unusual sensation in the throat and discomfort with swallowing
at 8 months in controls and 15 months in surgery group; Group 1: 4/21, Group 2: 0/24
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of mild and moderate OSA included ; Blinding details: Follow up periods differed between groups; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: unrelated illness
- Actual outcome: mild, moderate and severe bleeding at 8 months in controls and 15 months in surgery group; Group 1: 9/21, Group 2: 0/24
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of mild and moderate OSA included ; Blinding details: Follow up periods differed between groups; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: unrelated illness
- Actual outcome: nasal regurgitation at 8 months in controls and 15 months in surgery group; Group 1: 5/21, Group 2: 0/24
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of mild and moderate OSA included ; Blinding details: Follow up periods differed between groups; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: unrelated illness
- Actual outcome: moderate to severe pain at 8 months in controls and 15 months in surgery group; Group 1: 17/21, Group 2: 0/24
Risk of bias: All domain - Very high, Selection - Low, Blinding - Very high, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of mild and moderate OSA included ; Blinding details: Follow up periods differed between groups; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: unrelated illness

Protocol outcome 5: Disruption of partners sleep at >1 month
- Actual outcome: Snoring intensity score (subjective) at 8 months in controls and 15 months in surgery group; Group 1: mean 4.8 (SD 2.8); n=21, Group 2: mean 8.5 (SD 1.2); n=24
Risk of bias: All domain - Very high, Selection - Low, Blinding - Very high, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of mild and moderate OSA included ; Blinding details: Follow up periods differed between groups; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: unrelated illness
- Actual outcome: Snoring frequency score (subjective) at 8 months in controls and 15 months in surgery group; Group 1: mean 5.5 (SD 2.8);

Study	Ferguson 2003[23]
n=21, Group 2: mean 8.5 (SD 1.3); n=24 Risk of bias: All domain - Very high, Selection - Low, Blinding - Very high, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of mild and moderate OSA included ; Blinding details: Follow up periods differed between groups; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: unrelated illness Protocol outcome 6: Systolic blood pressure for hypertension at >1 month - Actual outcome: Systolic BP at 8 months in controls and 15 months in surgery group; Group 1: mean 140.4 mmhg (SD 16.9); n=21, Group 2: mean 144 mmhg (SD 16.9); n=24 Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of mild and moderate OSA included ; Blinding details: Follow up periods differed between groups; Group 1 Number missing: 0; Group 2 Number missing: 1, Reason: unrelated illness	
Protocol outcomes not reported by the study	Mortality at >1 month; ODI at >1 month; CO2 control at >1 month; Driving outcomes at >1 month; Neurocognitive outcomes at >1 month; Adherence in hours of use at >1 month; Patient preference at >1 month; HbA1c for diabetes at >1 month; Cardiovascular events at >1 month

Study	Friedman 2008[27]
Study type	RCT (randomised; Parallel)
Number of studies (number of participants)	1 (n=62)
Countries and setting	Conducted in Taiwan; Setting: hospital
Line of therapy	1st line

Study	Friedman 2008[27]
Duration of study	Intervention + follow up:
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Moderate
Subgroup analysis within study	Not applicable
Inclusion criteria	Patients were selected to participate in the study if they had a history of OSAHS and/or symptoms of OSAHS, mainly significant snoring and excessive daytime sleepiness; Friedman tongue position (FTP) I, II, or III based on the previously described staging system1; diagnosis of mild or moderate OSAHS (apnoea/hypopnea index (AHI) _ 5 and _40) on the baseline PSG; a soft palate _ 2 cm, but less than 3.5 cm; BMI _32 kg/m^2
Exclusion criteria	Patients were selected to participate in the study if they had a history of OSAHS and/or symptoms of OSAHS, mainly significant snoring and excessive daytime sleepiness; Friedman tongue position (FTP) I, II, or III based on the previously described staging system1; diagnosis of mild or moderate OSAHS (apnoea/hypopnea index (AHI) _ 5 and _40) on the baseline PSG; a soft palate _ 2 cm, but less than 3.5 cm; BMI _32 kg/m^2.
Recruitment/selection of patients	The study was conducted at one site between January 2005 and April 2006. Patients willing to participate underwent a detailed history and physical examination, including full otolaryngologic examination with fiberoptic nasopharyngoscopy. Demographics were recorded at the initial visit for each patient. Patients filled out a baseline quality of life questionnaire (QOL, SF-36 v2),4 bed partners completed a visual analog scale (VAS) to determine preoperative snoring intensity, and patients completed an Epworth Sleepiness Scale (ESS)5 to determine the extent of daytime somnolence. Candidates were selected based on

Study	Friedman 2008[27]
	inclusion and exclusion criteria (see below) and were scheduled to undergo a baseline polysomnogram (PSG) that determined their eligibility.
Age, gender and ethnicity	Age - Mean (SD): Age (yrs.) L surgery- 48.1 (11.2); control-39 (9.9). Gender (M:F): surgery-18/13; control-15/16. Ethnicity: not stated
Further population details	BMI (kg/m^2)surgery 29.3 (1.9); control- 28.7 (2.3); mean AHI-Palatal implants: 23.8 (5.5); Placebo: 20.1 (5.4)
Extra comments	-
Indirectness of population	No
Interventions	(n=31) Intervention 1: Surgery. Palatal implant placement: palatal stiffening procedure using the Pillar implant technique (PIT). The oral cavity was prepped with chlorhexidine gluconate rinse. Implantation sites were marked just in front of the hard palate–soft palate junction (Approximately 0.5 mL of 1% xylocaine with adrenaline 1:100,000 was injected into each of the marked sites. The Pillar implant system included an applicator and a Dacron implant. The applicator has 3 markings. The applicator tip was introduced into the soft palate until the third mark; care was taken not to bypass the soft palate. The device was then withdrawn until the second mark and the palatal implant was delivered into the soft palate. In same fashion, the other 2 implants were applied on each side of midline. Post-treatment antibiotics were used for 5 days, but no steroids were used. The key technical points are insertion of the implants as close to the hard palate junction as possible and keeping the 3 implants as close together as possible. Duration NA. Concurrent medication/care: not stated. Indirectness: No indirectness Comments: All patients underwent a preoperative mouth rinse with chlorhexidine and received a 5-day postoperative course of prophylactic antibiotics.

Study	Friedman 2008[27]
	(n=31) Intervention 2: No intervention - Inactive control therapy. The palatal implant insertion tools provided by the manufacturer for the placebo control group did not include the palatal implants, but they were in all other aspects identical to the implant insertion tools used in the treatment group receiving the implant. Duration NA. Concurrent medication/care: not stated. Indirectness: No indirectness Comments: All patients underwent a preoperative mouth rinse with chlorhexidine and received a 5-day postoperative course of prophylactic antibiotics.
Funding	Equipment / drugs provided by industry

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SURGERY versus INACTIVE CONTROL THERAPY

Protocol outcome 1: Quality of life at >1 month
- Actual outcome for Moderate: vitality sf-36 at 3 months; Group 1: mean 23.6 (SD 19.3); n=31, Group 2: mean -3.8 (SD 13.2); n=31
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: ; Group 2 Number missing:
- Actual outcome for Moderate: physical sf-36 at 3 months; Group 1: mean 10.3 (SD 8.6); n=31, Group 2: mean -1.5 (SD 4.8); n=31
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: ; Group 2 Number missing:
- Actual outcome for Moderate: mental sf-36 at 3 months; Group 1: mean 18.7 (SD 15.4); n=31, Group 2: mean -0.16 (SD 7.7); n=31
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: ; Group 2 Number missing:

Protocol outcome 2: Sleepiness score at >1 month
- Actual outcome for Moderate: ESS at 3 months; Group 1: mean 10.2 (SD 3.1); n=29, Group 2: mean 11.1 (SD 2.7); n=26
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness ; Group 1

Study	Friedman 2008[27]
Number missing: ; Group 2 Number missing: Protocol outcome 3: AHI/RDI at >1 month - Actual outcome for Moderate: AHI at 3 months; Group 1: mean 15.9 (SD 7.6); n=29, Group 2: mean 21 (SD 4.8); n=26 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: ; Group 2 Number missing:	
Protocol outcomes not reported by the study	Mortality at >1 month; ODI at >1 month; CO2 control at >1 month; Adverse effects of treatment at >1 month; Disruption of partners sleep at >1 month; Driving outcomes at >1 month; Neurocognitive outcomes at >1 month; Adherence in hours of use at >1 month; Patient preference at >1 month; HbA1c for diabetes at >1 month; Systolic blood pressure for hypertension at >1 month; Cardiovascular events at >1 month

Study	**Joar 2018[38] and Browaldh 2013[9]**
Study type	RCT (Patient randomised; Crossover)
Number of studies (number of participants)	2 (n=71)
Countries and setting	Conducted in Sweden; Setting: Ear, Nose, and Throat Department of the Karolinska University Hospital, Stockholm
Line of therapy	Mixed line
Duration of study	Intervention + follow up: 24 months

Study	Joar 2018[38] and Browaldh 2013[9]
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Severe
Subgroup analysis within study	Not applicable
Inclusion criteria	The following inclusion criteria were used: males and females > 18 years of age; AHI ≥ 15 events/hour of sleep (from PSG); ESS score ≥ 8; excessive daytime sleepiness three times a week or more; body mass index (BMI) of less than 36 kg/m^2; Friedman stage I or II6; and nonadherence with CPAP and MRD treatments, with the exception of patients with Friedman stage I and BMI of less than 30 kg/m^2.
Exclusion criteria	The following exclusion criteria were used: serious psychiatric, cardiopulmonary, or neurological disease or an American Society of Anaesthesiologists (ASA) classification of >3; patients who decline surgery; insufficient knowledge of Swedish language to complete questionnaires; nightshift workers; patients who could be dangerous in traffic according to responses in our non-standardised questionnaire; severe nasal congestion (could be included after topical nasal treatment); previous tonsillectomy (as such patients were considered partially treated); Friedman stage III; and severe clinical worsening of OSA during the study.
Recruitment/selection of patients	All OSA patients referred to the Ear, Nose, and Throat Department of the Karolinska University Hospital, Stockholm, Sweden from June 2007 to May 2011 for UPPP were eligible for this single-centre study

Study	Joar 2018[38] and Browaldh 2013[9]
Age, gender and ethnicity	Age - Mean (SD): surgery group= 41.5 (11.5), control group = 42.9 (11.7). Gender (M:F): 59/6. Ethnicity: unclear
Further population details	1. BMI: BMI of less than 30 2 kg/m^2. Co-existing conditions: Not stated. Sleepiness: ESS >9 ; Apnoea/hypopnoea index (events/h sleep) : surgery: 53.3 (19.7) ; control: 52.6 (21.7)
Indirectness of population	Serious indirectness: patients with moderate and severe OSA included
Interventions	(n=32) Intervention 1: Surgery. The surgical procedure UPPP including tonsillectomy involved minor resections of the soft palate and uvula using the cold steel technique and suturing of the tonsillar pillars including the palatopharyngeal muscle within 1 month. All surgeons used the same technique. . Duration 6 months. Concurrent medication/care: The patients underwent all-night in-lab PSG procedures and completed questionnaires at baseline (both groups), preoperatively (control group), and at 6- and 24-month postoperative follow-ups (both groups). The morning after PSG, the patients underwent vigilance testing, a modified OSLER (only performed once at each evaluation point), to measure sleep latency. . Indirectness: No indirectness Further details: 1. Intervention type: Not stated / Unclear 2. Type of surgery: tonsil (n=33) Intervention 2: No intervention - Usual care (lifestyle advice etc). no treatment for 6 months after baseline measurements recorded. Duration 6 months. Concurrent medication/care: The patients underwent all-night in-lab PSG procedures and completed questionnaires at baseline (both groups), preoperatively (control group), and at 6- and 24-month postoperative follow-ups (both groups). The morning after PSG, the patients underwent vigilance testing, a modified OSLER (only performed once at each evaluation point), to measure sleep latency. Indirectness: No indirectness
Funding	Academic or government funding

Study	Joar 2018[38] and Browaldh 2013[9]
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SURGERY versus USUAL CARE (LIFESTYLE ADVICE ETC) Protocol outcome 1: Quality of life at >1 month - Actual outcome for Severe: FOSQ - total change score at 6 months; Group 1: mean 1.53 (SD 2.64); n=32, Group 2: mean -0.2 (SD 1.22); n=33; functional outcomes of sleep questionnaire 5-20 Top=High is good outcome Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: patients with a mix of moderate and severe OSA included; Group 1 Number missing: 2, Reason: did not report FOSQ score; Group 2 Number missing: 0 Protocol outcome 2: Mortality Actual outcome: mortality: Group 1: 0/32; Group 2: 0/33 Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: patients with a mix of moderate and severe OSA included; Group 1 Number missing: 2, Reason: did not report FOSQ score; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Mortality at >1 month; Sleepiness score at >1 month; AHI/RDI at >1 month; ODI at >1 month; CO2 control at >1 month; Adverse effects of treatment at >1 month; Disruption of partners sleep at >1 month; Driving outcomes at >1 month; Neurocognitive outcomes at >1 month; Adherence in hours of use at >1 month; Patient preference at >1 month; HbA1c for diabetes at >1 month; Systolic blood pressure for hypertension at >1 month; Cardiovascular events at >1 month

Study	Koutsourelakis 2008[42]
Study type	RCT (Patient randomised; Crossover)
Number of studies (number of participants)	(n=49)

Study	Koutsourelakis 2008[42]
Countries and setting	Conducted in Greece; Setting: Centre of Sleep Disorders of the ‘‘Evangelismos’’ General Hospital of Athens, Greece.
Line of therapy	Mixed line
Duration of study	Follow up (post intervention): 3-4 months post-surgery
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Severe
Subgroup analysis within study	Not applicable:
Inclusion criteria	Enrolment criteria were: 1) nasal septum deviation with or without inferior turbinal hypertrophy, as assessed by clinical examination and flexible fibreoptic nasopharyngoscopy along with nasal resistance values exceeding normal limits at baseline (symptomatic fixed nasal obstruction); 2) AHI 5 events/h-1 at baseline; 3) no upper or lower respiratory tract disease, including a history of nasal allergy; 4) no recent surgery involving the upper airways; 5) no use of medications known to influence nasal resistance (antihistamine, decongestants, etc.); and 6) no history of neuromuscular or cardiovascular disease.
Exclusion criteria	Exclusion criterion was the treatment of OSA with continuous positive airway pressure (CPAP) during the course of the study.
Recruitment/selection of patients	Consecutive subjects who referred to the Centre of Sleep Disorders of the ‘‘Evangelismos’’ General Hospital of Athens, Greece for suspected sleep-disordered breathing were recruited.

Study	Koutsourelakis 2008[42]
Age, gender and ethnicity	Age - Mean (SD): surgery group 39 (7.5), placebo group 37.6 (8.8) . Gender (M:F): surgery group = 17/10, placebo group = 13/9. Ethnicity: unclear
Further population details	1. BMI: BMI of 30.2 kg/m^2 or more. Co-existing conditions: Not stated / Unclear 3. Gender: 4. High risk occupation group: Not stated / Unclear 5. Race: Not stated / Unclear 6. Sleepiness: ESS >9 ; Mean AHI: surgery - 31.5 (16.7); control - 30.6 (13.8)
Extra comments	patients with nasal septum deviation .
Indirectness of population	Serious indirectness: patients with mild, moderate and severe OSA included
Interventions	(n=27) Intervention 1: Surgery. All patients underwent sub mucous resection of the deviated nasal septum. In 18 out of 27 patients, sub mucous resection of the bilateral inferior turbinates was also performed. Nasal packing was removed on the second post-operative day, and routine saline nasal irrigation and debridement were performed. Post-operatively, none of the patients experienced any complication. . Duration of surgery. Concurrent medication/care: Sleep studies were performed <1 month before (baseline study) and 3–4 months after surgery. Each subject reported to the sleep laboratory at 21:00–22:00 h. Nasal resistance was measured in upright-seated and supine positions. A full-night diagnostic polysomnography with concomitant monitoring of the breathing route during sleep was then performed, usually 00:00–07:00 h. . Indirectness: No indirectness Further details: 1. Intervention type: Not stated / Unclear 2. Type of surgery: nasal (n=22) Intervention 2: No intervention - Placebo. To ensure blinding, a standard submucosal resection of the nasal septum was simulated. After the infiltration of the nasal septum with 10 mL lidocaine 1% containing epinephrine 1:200,000, the surgeon asked for all instruments and manipulated the nose as if submucosal resection was being performed. Patients remained in the operating room for the same amount of time required for the surgery group. Patients spent the night after the procedure in the hospital and were cared for by nurses who were unaware of the treatment group assignment. Nasal packing was removed on the second post-operative day and routine saline nasal irrigation and debridement were performed.

Study	Koutsourelakis 2008[42]
	. Duration of surgery. Concurrent medication/care: Sleep studies were performed <1 month before (baseline study) and 3–4 months after surgery. Each subject reported to the sleep laboratory at 21:00–22:00 h. Nasal resistance was measured in upright-seated and supine positions. A full-night diagnostic polysomnography with concomitant monitoring of the breathing route during sleep was then performed, usually 00:00–07:00 h. . Indirectness: No indirectness Further details: 1. Intervention type: Not stated / Unclear 2. type of surgery: nasal
Funding	Academic or government funding
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SURGERY versus PLACEBO Protocol outcome 1: Sleepiness score at >1 month - Actual outcome for Severe: ESS at 3-4 months post op; Group 1: mean 11.7 (3.4) n=27, Group 2: mean 12.5 (3.7)n=22 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: patients with mild, moderate and severe OSA included; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: AHI/RDI at >1 month - Actual outcome for Severe: AHI at 3-4 months post op; Group 1: mean 31.5 (SD 18.2); n=27, Group 2: mean 32.1 (SD 14.3); n=22 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: patients with mild, moderate and severe OSA included; Group 1 Number missing: 0; Group 2 Number missing: 0	
Protocol outcomes not reported by the study	Quality of life at >1 month; Mortality at >1 month; ODI at >1 month; CO2 control at >1 month; Adverse effects of treatment at >1 month; Disruption of partners sleep at >1 month; Driving outcomes at >1 month; Neurocognitive outcomes at >1 month; Adherence in hours of use at >1 month; Patient preference at >1 month; HbA1c for diabetes at >1 month; Systolic blood pressure for hypertension at >1 month; Cardiovascular events at >1 month

Study	Lojander 1996[45]
Study type	RCT (Patient randomised; Crossover)
Number of studies (number of participants)	1 (n=76)
Countries and setting	Conducted in Finland; Setting: hospital
Line of therapy	1st line
Duration of study	Intervention + follow up:
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Severe
Subgroup analysis within study	Not applicable
Inclusion criteria	Patients between the ages of 18-65 years with previously untreated OSAHS.
Exclusion criteria	Patients with asthma and other COPD, periodic leg movements syndrome, hypothyroidism, or other serious concomitant illness were excluded. Patients with BMI over 40 kg/m^2 were excluded.
Recruitment/selection of patients	Patients recruited between 1987 to 1992
Age, gender and ethnicity	Age - Median (range): conservative: 48 (36-61); 47 (27-62). Gender (M: F): Conservative- 14/0; surgery 16/1. Ethnicity: not stated
Further population details	ODI4 in control -median (range):34 (20-68); ODI4 in surgery group- median (range):45 (21-72)

Study	Lojander 1996[45]
Indirectness of population	No indirectness
Interventions	(n=18) Intervention 1: Surgery. UPPP + mandibular surgery. UPPP was performed according to the method of Fujita. Mandibular osteotomy was performed according to the method of Powell et al. For 13 patients UPPP alone was performed whereas 5 patients underwent both UPPP and mandibular osteotomy. . Duration NA. Concurrent medication/care: not stated. . Indirectness: No indirectness Further details: 1. Intervention type: 2. type of surgery: Comments: On each visit all affected patients were advised about smoking cessation and avoidance of alcohol as well as weight reduction but no specific treatments were programmed. (n=14) Intervention 2: No intervention - Usual care (lifestyle advice etc). Conservative management- no further details. Duration NA. Concurrent medication/care: not stated. Indirectness: No indirectness Further details: 1. Intervention type: 2. type of surgery: Comments: On each visit all affected patients were advised about smoking cessation and avoidance of alcohol as well as weight reduction but no specific treatments were programmed.
Funding	Funding not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SURGERY versus USUAL CARE (LIFESTYLE ADVICE ETC)

Protocol outcome 1: ODI at >1 month
- Actual outcome for Moderate-severe: ODI14 at 12 months; Group 1: mean 14 (SD 0); n=16, Group 2: mean 23 (SD 0); n=10
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness

- Actual outcome for Moderate-severe: ODI10 at 12 months; Group 1: mean 3 (SD 0); n=16, Group 2: mean 6 (SD 0); n=10
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness
Protocol outcome 2: Adverse effects of treatment at >1 month
- Actual outcome for Moderate-severe: velopharyngeal insufficiency at 12 months; Group 1: 2/18, Group 2: 0/14
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement -

Study	**Lojander 1996[45]**
Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness Protocol outcome 3: Cardiovascular events at >1 month - Actual outcome for Moderate-severe: CV events (non Q myocardial infarction and transient ischemic cerebral attack at 12 months; Group 1: 2/16, Group 2: 0/10 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: No indirectness	
Protocol outcomes not reported by the study	Quality of life at >1 month; Mortality at >1 month; Sleepiness score at >1 month; AHI/RDI at >1 month; CO2 control at >1 month; Disruption of partners sleep at >1 month; Driving outcomes at >1 month; Neurocognitive outcomes at >1 month; Adherence in hours of use at >1 month; Patient preference at >1 month; Systolic blood pressure for hypertension at >1 month; HbA1c for diabetes at >1 month

Study	MacKay, 2020 [47]
Study type	RCT (multicentre, parallel-group, open-label)
Number of studies (number of participants)	1 (n=102) Multilevel surgery (modified uvulopalatopharyngoplasty and minimally invasive tongue volume reduction; n = 51) or ongoing medical management (e.g, advice on sleep positioning, weight loss; n = 51).
Countries and setting	Conducted in Australia; Setting: across 6 Australian academic centres
Line of therapy	Unclear
Duration of study	Follow up (post intervention): 6 months

Study	MacKay, 2020 [47]
Method of assessment of guideline condition	Adequate method of assessment/diagnosis.
Stratum	Moderate-severe
Subgroup analysis within study	Not applicable
Inclusion criteria	Adults with symptomatic moderate or severe OSA in whom conventional treatments had failed were enrolled from August 2014 to November 2017, with follow-up until August 2018. Eligible adults were aged 18 to 70 years with moderate or severe OSA (defined as apnoea-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
Exclusion criteria	Patients were excluded if they had significant medical or psychiatric comorbidities, were judged to be a high anaesthetic risk, were pregnant, or had specific anatomical contraindications to the intended surgery (eg, severe palatal scarring from previous surgery or severe retrognathia).
Age, gender and ethnicity	Age - Mean (SD): surgery group: 42.7 (12.8); control group - 46.4 (12.6) Gender (Men): surgery group: 41 (80%); control group : 43 (84%)
Further population details	BMI: Men: surgery: 30.1 (4.0); control: 30.0 (3.6); Women: surgery: 33.3 (2.8); control: 26.6 (2.9) mean AHI- surgery: 47.9 (23.1); Control: 45.3 (23.9) Epworth Sleepiness Scale, mean (SD): surgery: 12.4 (3.6); control: 11.1 (4.7) Previous OSA treatment, No. (%) Tried CPAP: surgery: 38 (75); control: 37 (73)

Study	MacKay, 2020 [47]
	Refused CPAP: surgery: 13 (25); control:14 (27) Tried mandibular advancement device: surgery:16 (31); control: 12 (24)
Extra comments	Participants were predominantly middle-aged men with overweight or obesity and severe OSA. Friedman stages were reasonably evenly distributed within each group. 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 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
Indirectness of population	Serious indirectness: patients included with moderate or severe OSA (defined as apnoea-hypopnea index [AHI] of 15-30 and >30 events/h of sleep)
Interventions	(n=51) Intervention 1: Surgery. The surgery intervention 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. A training workshop was conducted to standardise the surgical techniques among the 7 participating surgeons. (n=51) Intervention 2: Medical management. 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.
Funding	Academic or government funding
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SURGERY versus MEDICAL MANAGEMENT	

Study	MacKay, 2020 [47]

Protocol outcome 1: Quality of life at >1 month
- Actual outcome: FOSQ at 6 months

Group 1: mean 18.6 (SD 1.8); n=50, Group 1: mean 16.3 (SD 2.4); n=49,
Risk of bias: All domain - Very high, Selection - Low, Blinding - high, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of moderate and severe OSA included

Protocol outcome 2: Sleepiness score at >1 month
- Actual outcome: Epworth sleepiness scale at 6 months

Group 1: mean 5.3 (SD 3.0); n=50, Group 2: mean 10.5 (SD 4.7); n=49
Risk of bias: All domain – very high, Selection - Low, Blinding - high, Incomplete outcome data - low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of moderate and severe OSA included

Protocol outcome 3: AHI/RDI at >1 month
- Actual outcome: AHI at 6 months

Group 1: mean 20.8 (SD 18.4); n=50, Group 2: mean 34.5 (SD 23.0); n=49
Risk of bias: All domain - High, Selection - Low, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of moderate and severe OSA included

Protocol outcome 4: Adverse effects of treatment at >1 month
- Actual outcome: serious adverse events; Group 1: 2/50, Group 2: 0/49
Risk of bias: All domain - High, Selection - Low, Blinding - low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of moderate and severe OSA included.

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. 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

Study	MacKay, 2020 [47]
medical management group experienced a serious adverse event Protocol outcome 6: Systolic blood pressure for hypertension at >1 month - Actual outcome: 24 h ambulatory systolic blood pressure at 6 months Group 1: mean 120.0 mmhg (SD 12.0); n=50, Group 2: mean 124.7 mmhg (SD 13.6); n=49 Risk of bias: All domain - High, Selection - Low, Blinding - low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of moderate and severe OSA included Protocol outcome 6: Diastolic blood pressure for hypertension at >1 month - Actual outcome: 24 h ambulatory diastolic blood pressure at 6 months Group 1: mean 74.1 mmhg (SD 8.0); n=50, Group 2: mean 77.7 mmhg (SD 10.9); n=49 Risk of bias: All domain - High, Selection - Low, Blinding - low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: mix of moderate and severe OSA included	
Protocol outcomes not reported by the study	Mortality at >1 month; CO2 control at >1 month; Driving outcomes at >1 month; Neurocognitive outcomes at >1 month; Patient preference at >1 month; HbA1c for diabetes at >1 month; Cardiovascular events at >1 month

Study	Sommer 2016[70]
Study type	RCT (Patient randomised; Crossover)
Number of studies (number of participants)	1 (n=42)
Countries and setting	Conducted in Germany; Setting: The trial was conducted at the Department of Otorhinolaryngology, Head and Neck Surgery, Sleep Disorder Centre, University Hospital Mannheim and the Department of Otorhinolaryngology, University Hospital Klinikum rechts der Isar, Technische Universität München, Munich.
Line of therapy	Mixed line
Duration of study	Follow up (post intervention): 3 months
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Severe
Subgroup analysis within study	Not applicable
Inclusion criteria	Inclusion criteria were obstructive sleep apnoea confirmed by polysomnography (PSG) with an Apnoea–Hypopnea Index (AHI) above 15, according to the second edition of the International Classification of Sleep Disorders valid at that time, and tonsillar hypertrophy with velopharyngeal obstruction confirmed by clinical examination. A further very important inclusion criterion was rejection of or poor compliance with ventilation therapy and an explicit wish on the part of the patient for a different approach (second-line therapy). All enrolled patients had tried CPAP treatment without success for at least one night.
Exclusion criteria	The most important exclusion criteria were body mass index above 34 kg/m, increased anaesthetic risk

Study	Sommer 2016[70]
	according to the criteria of the American Society of Anaesthesiologists (ASA), specifically ASA class above III, and other relevant types of obstruction or significant malformations of the facial skeleton confirmed by clinical examination.
Age, gender and ethnicity	Age - Mean (SD): 37.4 (10.7). Gender (M:F): 40/2. Ethnicity: unclear
Further population details	1. BMI: BMI 28.8 ±3.2. 2. Co-existing conditions: Not stated / Unclear 3. Gender: Male 4. High risk occupation group: Not stated / Unclear 5. Race: Not stated / Unclear 6. Sleepiness: Not stated / Unclear ; AHI : control 35.7 ± 19.4; 33.7 ± 14.5
Indirectness of population	Serious indirectness: patients with moderate and severe OSA included
Interventions	(n=23) Intervention 1: Surgery. After cold steel tonsillectomy using general anaesthesia, uvulopalatopharyngoplasty according to the modifications by Pirsig was performed. Tonsil size was determined immediately following surgery using volume displacement. Complications occurring during inpatient stay, particularly haemorrhages were recorded by type and severity. TE-UPPP was performed by three different surgeons at the Mannheim trial site and by one in Munich. Duration of surgery + 3 months. Concurrent medication/care: patients received EEGs at baseline less than 1 month after randomisation and at 3 months follow up. Indirectness: No indirectness Further details: 1. Intervention type: Not stated / Unclear 2. Type of surgery: tonsil (n=19) Intervention 2: No intervention - Usual care (lifestyle advice etc). Patients in the control arm initially received no treatment and underwent repeat polysomnography again after three months, then underwent TE-UPPP. Duration 3 months. Concurrent medication/care: patients received EEGs at baseline less than 1 month after randomisation and at 3 months follow up. Indirectness: No indirectness
Funding	Principal author funded by industry
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SURGERY versus USUAL CARE (LIFESTYLE ADVICE ETC)	

Study	Sommer 2016[70]
Protocol outcome 1: Sleepiness score at >1 month - Actual outcome: ESS at 3 months; Group 1: mean 6.2 (SD 2.9); n=20, Group 2: mean 9.6 (SD 5.2); n=15 Risk of bias: All domain - Very high, Selection - high Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: patients with moderate and severe OSA included; Group 1 Number missing: 3, Reason: none provided; Group 2 Number missing: 4, Reason: none provided Protocol outcome 2: AHI/RDI at >1 month - Actual outcome: AHI at 3 months; Group 1: mean 15.4 (SD 14.1); n=18, Group 2: mean 28.6 (SD 19.3); n=16 Risk of bias: All domain - High, Selection – high, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: serious indirectness, Comments: patients with moderate and severe OSA included; Group 1 Number missing: 5, Reason: none provided; Group 2 Number missing: 3, Reason: none provided Protocol outcome 1: Spo2 - Actual outcome: Spo2; Group 1: mean 92.9 (SD 5.5); n=17, Group 2: mean 91.7 (SD 4.6); n=16 Risk of bias: All domain - Very high, Selection - high Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: Serious indirectness, Comments: patients with moderate and severe OSA included; Group 1 Number missing: 3, Reason: none provided; Group 2 Number missing: 4, Reason: none provided	
Protocol outcomes not reported by the study	Quality of life at >1 month; Mortality at >1 month; ODI at >1 month; CO2 control at >1 month; Adverse effects of treatment at >1 month; Driving outcomes at >1 month; Neurocognitive outcomes at >1 month; HbA1c for diabetes at >1 month; Systolic blood pressure for hypertension at >1 month; Cardiovascular events at >1 month

Study	Vicini 2010[82]
Study type	RCT (Patient randomised; Crossover)
Number of studies (number of participants)	1 (n=50)

Study	Vicini 2010[82]
Countries and setting	Conducted in Italy; Setting: The study was carried out at the Department of Special Surgery, ENT and Oral Surgery Unit, Morgagni-Pierantoni Hospital, University of Pavia in Forlì, Italy.
Line of therapy	Mixed line
Duration of study	Intervention + follow up: 1 year
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Severe
Subgroup analysis within study	Not applicable
Inclusion criteria	Inclusion criteria included the presence of severe OSAHS (AHI >30), regardless of body mass index (BMI; which was usually abnormally high) and no formal contraindication for surgery according to the Stanford protocol (pre-existing local and general medical conditions that could increase the risk of surgery or might compromise the final outcome, fear of surgery, concern over pain and discomfort, loss of work or income during convalescence, advancing age) and no formal contraindication for APAP Chronic Obstructive Pulmonary Disease (COPD), heart dysrhythmia, heart failure, restrictive lung disease, neuromuscular disease, previous surgery for SDB).
Exclusion criteria	Contraindication for surgery according to the Stanford protocol (pre-existing local and general medical conditions that could increase the risk of surgery or might compromise the final outcome, fear of surgery, concern over pain and discomfort, loss of work or income during convalescence, advancing age. contraindication for APAP Chronic Obstructive Pulmonary Disease (COPD), heart dysrhythmia,

Study	Vicini 2010[82]
	heart failure, restrictive lung disease, neuromuscular disease, previous surgery for SDB.
Age, gender and ethnicity	Age - Mean (SD): treatment group = 49.1 (9.1) control group = 48.7 (10.7). Gender (M:F): 43/7. Ethnicity: unclear
Further population details	1.BMI of 30 2 kg/m^2 or more. Co-existing conditions: Not applicable 3. Gender: Male 4. High risk occupation group: Not stated / Unclear 5. Race: Not stated / Unclear 6. Sleepiness: ESS >9 ; Mean AHI was 56. 8 ± 16.5 SD,
Indirectness of population	No indirectness
Interventions	(n=25) Intervention 1: Surgery. Surgical procedures were performed under general anaesthesia after routine fibre optic orotracheal intubation. Temporary tracheotomy was routinely carried out to avoid possible episodes of dyspnoea in the first 24 to 48 hours after surgery and to facilitate the possibility of having to suction mucous secretion. As a first step, a sagittal split ramus mandibulae osteotomy according to Obwegeser-Dal Pont was performed with a powered reciprocating saw and a Lindemann cutting burr (in the ramus inner cortex area). The fixed amount of advancement, 11 mm for all the cases, was checked by means of a customized intermediated splint. To stabilize the achieved advancement, the surgeon inserted 3 to 4 bi cortical screws. In 3 cases, a titanium plate was added on each side to improve stabilization. As a second step, a low Le Fort I maxillary osteotomy was carried out step by step using a powered reciprocating saw and different kinds of special osteotomes. The final position of the maxilla was stabilized by 4 titanium screwed plates. The rigid inter maxillary fixation was removed after 24 hours, and oral intake of food was immediately encouraged. The tracheotomy was removed usually on the fourth/fifth day. Discharge was possible within 1 week for all the patients. Duration 1 year. Concurrent medication/care: Irrespective of the randomization result, all patients were previously informed of all the alternative medical and/or surgical therapies available. Fort I maxillary osteotomy was carried out step by step using a powered reciprocating saw and different kinds of special osteotomes. The final position of the maxilla was stabilized by 4 titanium screwed plates. The rigid inter maxillary fixation was removed after 24 hours, and oral intake of food was

Study	Vicini 2010[82]
	immediately encouraged. The tracheotomy was removed usually on the fourth/fifth day. Discharge was possible within 1 week for all the patients. Indirectness: No indirectness Further details: 1. Intervention type: Not stated / Unclear 2. type of surgery: (mandibular). (n=25) Intervention 2: CPAP. The patients enrolled in the conservative section of the present study were submitted to automatic APAP application (Vivisol SOMNO smart2, BREAS, MedicAir) with a nasal mask, held in position with an elastic headgear and attached to a flow generator by elephant tubing. This APAP is able to detect 3 different parameters for pressure auto titration: (1) forced oscillation, (2) flow limitation, (3) snoring. The patients were requested to continue using the Auto-CPAP only after a successful test period, usually 1 week, checked by interview and smart card evaluation (the smart card records the true time of utilization and different operative parameters). Duration 1 year. Concurrent medication/care: Irrespective of the randomization result, all patients were previously informed of all the alternative medical and/or surgical therapies available. Patient counselling and management were both handled by the same physician. Indirectness: No indirectness
Funding	Funding not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SURGERY versus CPAP Protocol outcome 1: Mortality at >1 month - Actual outcome for Severe: mortality at 1 year; Group 1: 0/25, Group 2: 0/22 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 3, Reason: 3 dropped out and submitted to MMA surgery Protocol outcome 2: Sleepiness score at >1 month - Actual outcome for Severe: ESS at 1 year; Group 1: mean 7.7 (SD 1.3); n=25, Group 2: mean 5.9 (SD 1.6); n=22 Risk of bias: All domain - Very high, Selection - Low, Blinding - High, Incomplete outcome data - High, Outcome reporting - Low, Measurement	

Study	Vicini 2010[82]
- Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 3, Reason: 3 dropped out and submitted to MMA surgery Protocol outcome 3: AHI/RDI at >1 month - Actual outcome for Severe: AHI at 1 year; Group 1: mean 8.1 (SD 7); n=25, Group 2: mean 6.3 (SD 1.9); n=22 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 3, Reason: 3 dropped out and submitted to MMA surgery Protocol outcome 4: Adverse effects of treatment at >1 month - Actual outcome for Severe: parasthesia around the chin at 1 year; Group 1: 7/25, Group 2: 0/22 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 3, Reason: 3 dropped out and submitted to MMA surgery - Actual outcome for Severe: minimal malocclusion at 1 year; Group 1: 6/25, Group 2: 0/22 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 3, Reason: 3 dropped out and submitted to MMA surgery - Actual outcome for Severe: facial lesion at 1 year; Group 1: 0/25, Group 2: 1/22 Risk of bias: All domain - High, Selection - Low, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 3, Reason: 3 dropped out and submitted to MMA surgery	
Protocol outcomes not reported by the study	Quality of life at >1 month; ODI at >1 month; CO2 control at >1 month; Disruption of partners sleep at >1 month; Driving outcomes at >1 month; Neurocognitive outcomes at >1 month; Adherence in hours of use at >1 month; Patient preference at >1 month; HbA1c for diabetes at >1 month; Systolic blood pressure for hypertension at >1 month; Cardiovascular events at >1 month

Study	Vicini, 2020 [83]
Study type	RCT (randomised; Parallel)
Number of studies (number of participants)	1 (n=50); barbed repositioning pharyngoplasty (BRP) (n = 25), Observation (n= 25)
Countries and setting	Conducted in Italy; Setting: hospital
Line of therapy	1st line
Duration of study	Intervention + 6 months follow up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	severe OSA
Subgroup analysis within study	Not applicable
Inclusion criteria	Patients suffering from moderate to severe OSA (AHI ≥15 events/h) with certain degree of nasal obstruction planned for BRP and tonsillectomy, with nasal surgery (septoturbinoplasty), Grades 1–2 tonsillar hypertrophy, aged between 18 and 65 years old, BMI ≤ 35, failure of CPAP or low adherence to this treatment during the last 3 months (< 4 h per night), mainly palatal/pharyngeal collapses at DISE (severe circular palatal collapses and severe transversal pharyngeal collapses with none or mild tongue collapses).
Exclusion criteria	Serious psychiatric, cardiopulmonary, or neurological disease, American Society of Anesthesiologists (ASA) classification > 3, patients negative to surgery, previous tonsillectomy and OSA surgery, significant craniofacial anomalies, pregnant women, Grades 3–4 tonsillar hypertrophy, mainly lingual/base of tongue collapses at DISE and follow-up < 6 months.

Study	Vicini, 2020 [83]
Recruitment/selection of patients	All patients with OSA referred consecutively to Otolaryngology and Head Neck Department, Hospital Morgagni Pierantoni, Forlì, Italy, from February 2015 to February 2018 for palatal surgery were possible candidates for inclusion in the study. Patients who met the criteria for the study were invited to participate and were enrolled in the study by different physicians.
Age, gender and ethnicity	Age – Mean yrs: surgery-44.64; control- 50. Gender (M:F): surgery- 22/3; control- 20/1. Ethnicity: not stated
Further population details	mean AHI-: surgery 25.58 ± 14.60; control: 36.83 ± 23.82 ESS: surgery: 9.28 ± 3.10 ; control: 10.4 ± 23.68
Extra comments	The definition of surgical response and success were a reduction from the preoperative AHI of at least 50% (response) and less than 20 events per hour (success).
Indirectness of population	No
Interventions	(n=25) Intervention 1: Surgery. Barbed repositioning pharyngoplasty (BRP) After bilateral tonsillectomy meticulous sparing of the palatoglossus and palatopharyngeus muscles was performed. Two weakening or releasing partial incisions were done by a pinpoint bowie (Colorado) at the inferior (caudal) part of the palatopharyngeal muscle. The centre of the palate was marked at palatal spine, and the pterygomandibular raphe in both sides were located by digital palpation and marked. Single barbed suture, bidirectional polydioxanone absorbable monofilament, size 2.0, with transition zone in the middle was generally used. One needle was introduced at the centre point then passed laterally within the palate, turning around pterygomandibular raphe till it came out at the most superior part of the raphe at one side; the thread was pulled until it hung at the central transition. The needle again was re-introduced close to point of exit, passing around the pterygomandibular raphe, till it came out into the tonsillectomy bed, then through the upper part of the

Study	Vicini, 2020 [83]
	palatopharyngeus muscle and came out near to mucosa of posterior pillar not through it. The posterior pillar was entered at the junction between the upper third and the lower two-thirds. Then, again, the needle was passed back through the tonsillectomy, bed and then this suture would be suspended around the raphe again. The opposite side was done by the same way. Finally, each thread came out at the raphe of the same side, for locking of the stitches and looseness prevention; a superficial stitch in the opposite direction was taken, and then, the thread was cut while pushing the tissue downward for more traction (n=25) Intervention 2: No intervention Observation. No further information.
Funding	Not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SURGERY versus CONTROL

Protocol outcome 1: AHI
- Actual outcome: AHI at 6 months; Group 1: mean (SD) 9.82 (9.88), n=25; Group 2: mean (SD) 31.93 (21.89), n= 25
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: no

Protocol outcome 1: ESS
- Actual outcome: ESS at 6 months; Group 1: mean (SD) 3.76 (4.42), n=25; Group 2: mean (SD) 10.85 (3.91), n= 25
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: no

Protocol outcomes not reported by the study	Mortality at >1 month; Quality of life at >1 month; CO2 control at >1 month; Adverse effects of treatment at >1 month; Disruption of partners sleep at >1 month; Driving outcomes at >1 month;

Study	Vicini, 2020 [83]
	Neurocognitive outcomes at >1 month; Adherence in hours of use at >1 month; Patient preference at >1 month; HbA1c for diabetes at >1 month; Systolic blood pressure for hypertension at >1 month; Cardiovascular events at >1 month

Study (subsidiary papers)	**Tegelberg 1999[78], Walker-Engstrom 2002[85], Walker-Engstrom 2000[87] and Wilhelmsson 1999[88]**
Study type	RCT (Patient randomised; Crossover)
Number of studies (number of participants)	4 (n=95)
Countries and setting	Conducted in Sweden; Setting:
Line of therapy	1st line
Duration of study	Intervention + follow up: 4 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Moderate
Subgroup analysis within study	Not applicable:
Inclusion criteria	Ninety-five male patients with mild to moderate OSA (AI > 5 or AHI >10) were included.

Study (subsidiary papers)	Tegelberg 1999[78], Walker-Engstrom 2002[85], Walker-Engstrom 2000[87] and Wilhelmsson 1999[88]
Exclusion criteria	Patients with an AI > 25, mental illness, drug abuse, significant nasal obstruction, insufficient teeth to anchor an appliance, pronounced dental malocclusion, severe cardiovascular disease or neurological and respiratory disease were not studied. An additional exclusion criterion was patients who were aged < 20 or > 65 y.
Recruitment/selection of patients	To participate in the study, patients needed to have sufficient dental support for anchorage of the dental appliance, i.e. at least one premolar or molar tooth in both upper and lower jaws on both sides and not suffer from severe periodontal and cariogenic disease.
Age, gender and ethnicity	Age - Mean (SD): surgery group = 51.0 95% CI = 49.1-52.9, oral device group = 49.3 95% CI = 46.8-51.9. Gender (M:F): 95/0. Ethnicity: unclear
Further population details	1.BMI of less than 30 2 kg/m^2. Co-existing conditions: Not stated / Unclear 3. Gender: Male 4. High risk occupation group: Not stated / Unclear 5. Race: Not stated / Unclear 6. Sleepiness: Not stated / Unclear ; baseline AHI: oral device: 18.2 (15.7-20.8); surgery: 20.4 (17.44-23.3).
Extra comments	-
Indirectness of population	Serious indirectness: patients with mild and moderate OSA included
Interventions	(n=46) Intervention 1: Surgery. The procedure involved tonsillectomy regardless of the size of the tonsils, and resection of excess fat and mucosa of the soft palate and uvula. The palpable musculature was saved, and several sutures approximated the anterior and posterior tonsil pillars. The UPPP surgery was performed under general anaesthesia. . Duration 4 years. Concurrent medication/care: Sleep studies were performed at baseline and 6 and 12 months after intervention in patients' homes with a portable unit by a blinded technician. Further details: 1. Intervention type: Not stated / Unclear 2. type of surgery: tonsil

Study (subsidiary papers)	Tegelberg 1999[78], Walker-Engstrom 2002[85], Walker-Engstrom 2000[87] and Wilhelmsson 1999[88]
	(n=49) Intervention 2: oral devices - MAS. Duration 4 years. Concurrent medication/care: sleep studies were performed at baseline and 6 and 12 months after intervention in patients' homes with a portable unit by a blinded technician. Indirectness: No indirectness Further details: 1. Intervention type: Not applicable 2. type of surgery:
Funding	Academic or government funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SURGERY versus MAS

Protocol outcome 1: Quality of life at >1 month
- Actual outcome for Moderate: quality of life- MSE vitality at 1 year; Group 1: mean 26.4 (SD 11.67); n=43, Group 2: mean 31.6 (SD 13.63); n=37
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness

- Actual outcome for Moderate: quality of life- MSE -sleep at 1 year; Group 1: mean 25.2 (SD 15.08); n=43, Group 2: mean 29.2 (SD 17.03); n=37
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness

- Actual outcome for Moderate: quality of life- MSE -contentment at 1 year; Group 1: mean 27.4 (SD 12.07); n=43, Group 2: mean 33.7 (SD 13.32); n=37
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness
Protocol outcome 2: AHI/RDI at >1 month
- Actual outcome for Moderate: AHI at 6 months; Group 1: mean 8.6 (SD 7.1); n=43, Group 2: mean 6.6 (SD 8.8); n=37
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness

Study (subsidiary papers)	Tegelberg 1999[78], Walker-Engstrom 2002[85], Walker-Engstrom 2000[87] and Wilhelmsson 1999[88]
- Actual outcome for Moderate: AHI at 1 year; Group 1: mean 10.4 (SD 9.1); n=43, Group 2: mean 5.9 (SD 9); n=37 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness - Actual outcome for Moderate: AHI at 4 year; Group 1: mean 14.2 (SD 3.4); n=40, Group 2: mean 7.2 (SD 2.6); n=32 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness Protocol outcome 3: ODI at >1 month - Actual outcome for Moderate: ODI at 6 months; Group 1: mean 8 (SD 8); n=43, Group 2: mean 6.4 (SD 8.5); n=37 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness - Actual outcome for Moderate: ODI at 1 year; Group 1: mean 9.3 (SD 9.9); n=43, Group 2: mean 6.1 (SD 9.8); n=37 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness - Actual outcome for Moderate: ODI at 4 year; Group 1: mean 13.1 (SD 10.7); n=40, Group 2: mean 6.7 (SD 6.8); n=32 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: No indirectness Protocol outcome 4: Adverse effects of treatment at >1 month - Actual outcome for Moderate: dysphagia at 4 year; Group 1: 4/40, Group 2: 0/32 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness - Actual outcome for Moderate: nasopharyngeal regurgitation at 4 year; Group 1: 3/40, Group 2: 0/32 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness	
Protocol outcomes not reported by the study	Mortality at >1 month; Sleepiness score at >1 month; CO2 control at >1 month; Disruption of partners sleep at >1 month; Driving outcomes at >1 month; Neurocognitive outcomes at >1 month; Adherence

Study (subsidiary papers)	Tegelberg 1999[78], Walker-Engstrom 2002[85], Walker-Engstrom 2000[87] and Wilhelmsson 1999[88]
	in hours of use at >1 month; Patient preference at >1 month; HbA1c for diabetes at >1 month; Systolic blood pressure for hypertension at >1 month; Cardiovascular events at >1 month

Study	Woodson 2003[89]
Study type	RCT (Patient randomised; Crossover)
Number of studies (number of participants)	1 (n=90)
Countries and setting	Conducted in USA; Setting: hospital
Line of therapy	1st line
Duration of study	Intervention + follow up:
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Moderate
Subgroup analysis within study	Not applicable
Inclusion criteria	(1) age 18 to 65 years, (2) self-reports of daytime somnolence, (3) body mass index (BMI) _34 kg/m^2, (4) no prior surgical or CPAP treatment for OSAS, and (5) mild to moderate OSAS defined by an AHI of 10 to 30 on screening sleep study
Exclusion criteria	(1) another significant sleep disorder (e.g., insomnia, periodic limb movement), (2) tonsillar hypertrophy, (3) nasal or supraglottic obstruction on examination, (4) ASA class IV/V, (5) claustrophobia, (6) Latex allergy, (7)

Study	Woodson 2003[89]
	pregnancy or plans to become pregnant, (8) major depression or non-stabilised psychiatric disorder, (9) drug or alcohol abuse, (10) history of an accident secondary to sleepiness, or (11) participation in another study
Recruitment/selection of patients	Subjects were recruited directly from the academic otolaryngology practices and from poster and newspaper advertisements
Age, gender and ethnicity	Age - Mean (SD): Placebo: 46.0 (8.1); TCRFTA : 49.4 (9.2); CPAP : 51.7 (8.6). Gender (M:F): MALE (%): placebo: 70.0; TCRFTA : 89.7 ; CPAP: 75.0. Ethnicity: not stated
Further population details	AHI: TCRFTA: 21.3 (11.1); Placebo: 15.4 (7.8)
Extra comments	-
Indirectness of population	No indirectness
Interventions	(n=30) Intervention 1: Surgery. TCRFTA (radiofrequency energy delivered to create non-overlapping lesions in two/three tongue sites, occurring at 4-week intervals. Data recorded after last treatment session. Palate sessions also included) Active temperature-controlled radiofrequency tissue ablation (TCRFTA) was performed with the Somnoplasty radiofrequency generator (Gyrus- ENT, Memphis, TN). Five tongue and 2 palate sessions were planned for each active subject. Subjects were treated peri-operatively with oral antibiotics, prednisone, antiseptic oral rinse, analgesic (as needed), and nonsteroidal anti-inflammatory medication (as needed). A local anaesthetic mixture (2.5 mL of 2% lidocaine with 1:100,000 epinephrine, 2.0 mL of normal saline, and 0.5 mL of 8.4% sodium bicarbonate) was injected into each tongue treatment site, and 1% lidocaine with 1:100,000

Study	Woodson 2003[89]
	epinephrine (1 to 2 mL) was injected into each palate site. Radiofrequency energy was delivered to create non-overlapping lesions in 2 or 3 tongue sites (1000 or 750 J, respectively, per site; target temperature 85° C; maximum power 10 W) per tongue treatment session, which occurred at 4-week intervals. Radiofrequency energy was delivered to create 1 midline and 2 lateral lesions (non-overlapping) to the soft palate (650 J and 325 J, respectively) in each palate treatment session. Investigators were instructed to adjust lesion numbers per treatment session based on clinical judgment and patient tolerance. When tongue and palate sessions were combined, the subject was offered overnight hospital admission. Investigators were instructed to perform sequential and not simultaneous tongue and palate treatments if there were concerns about airway oedema or patient tolerance. Attempts were made to apply similar levels of energy in all patients irrespective of the timing of sessions. Duration NA. Concurrent medication/care: NA Further details: 1. Intervention type: 2. type of surgery: (n=30) Intervention 2: No intervention - Inactive control therapy. Sham TRCFTA. Sham-placebo TCRFTA was performed as described above for tongue TCRFTA except that a blocking control box on the radiofrequency generator was set to “off” to prevent delivery of energy. Three tongue sessions were planned for each sham-placebo subject at 4-week intervals with 3 tongue lesions created per session. Subjects were anesthetised and medicated as described for active tongue TCRFTA. The sham treatment sessions were limited to 3 to balance the risk of hematoma, oedema, or abscess formation at the site of anaesthetic injection or TCRFTA probe insertion versus the goal of providing a realistic placebo

Study	Woodson 2003[89]
	Duration NA. Concurrent medication/care: NA Further details: 1. Intervention type: 2. type of surgery: (n=30) Intervention 3: CPAP. CPAP Nasal CPAP therapy was titrated unattended over 3 or more nights with the AutoSet T device. Final constant CPAP pressure was set as the 95-percentile pressure and was continued for 8 weeks. Subjects were seen at 1, 2, and 4 weeks to troubleshoot and optimise compliance. Duration NA. Concurrent medication/care: NA
Funding	No funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SURGERY versus INACTIVE CONTROL THERAPY

Protocol outcome 1: Quality of life at >1 month
- Actual outcome for Moderate: SF 36 physical at 8 weeks; Group 1: mean 0.5 (SD 6.8); n=24, Group 2: mean 1.5 (SD 7.8); n=27
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness

- Actual outcome for Moderate: SF 36 mental at 8 weeks; Group 1: mean 2.9 (SD 7.3); n=24, Group 2: mean 0.4 (SD 6.4); n=27
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness

- Actual outcome for Moderate: FOSQ at 8 weeks; Group 1: mean 1.2 (SD 1.6); n=26, Group 2: mean 0.4 (SD 2); n=28
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness
Protocol outcome 2: Sleepiness score at >1 month
- Actual outcome for Moderate: ESS at 8 weeks; Group 1: mean -2.1 (SD 3.9); n=26, Group 2: mean -1 (SD 3.1); n=28
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness
Protocol outcome 3: Adverse effects of treatment at >1 month
- Actual outcome for Moderate: Bleeding at 8 weeks; Group 1: 3/26, Group 2: 3/28

Study	Woodson 2003[89]
Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness - Actual outcome for Moderate: ulcerations at 8 weeks; Group 1: 1/28, Group 2: 0/26 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness - Actual outcome for Moderate: infections at 8 weeks; Group 1: 0/26, Group 2: 0/28 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SURGERY versus CPAP Protocol outcome 1: Quality of life at >1 month - Actual outcome for Moderate: FOSQ at 8 weeks; Group 1: mean 1.2 (SD 1.6); n=26, Group 2: mean 1.5 (SD 2.1); n=25 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness - Actual outcome for Moderate: sf 36 physical at 8 weeks; Group 1: mean 0.5 (SD 6.8); n=24, Group 2: mean 0.1 (SD 7.7); n=24 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness - Actual outcome for Moderate: sf 36 mental at 8 weeks; Group 1: mean 2.9 (SD 7.3); n=24, Group 2: mean 2 (SD 6.1); n=24 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness Protocol outcome 2: Sleepiness score at >1 month - Actual outcome for Moderate: ESS at 8 weeks; Group 1: mean -2.1 (SD 3.9); n=26, Group 2: mean -2.3 (SD 5.2); n=25 Risk of bias: All domain - Low, Selection - Low, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Subgroups - Low, Other 1 - Low, Other 2 - Low, Other 3 - Low; Indirectness of outcome: Serious indirectness	
Protocol outcomes not reported by the study	Mortality at >1 month; AHI/RDI at >1 month; ODI at >1 month; CO2 control at >1 month; Disruption of partners sleep at >1 month; Driving outcomes at >1 month; Neurocognitive outcomes at >1 month; Adherence in hours of use at >1 month; Patient preference at >1 month; HbA1c for diabetes at >1 month; Systolic blood pressure for hypertension at >1 month; Cardiovascular events at >1 month

Appendix E: Forest plots

E.1 Surgery vs conservative management/no surgery/sham surgery- Moderate OSAHS

Figure 2: AHI (lower is better)

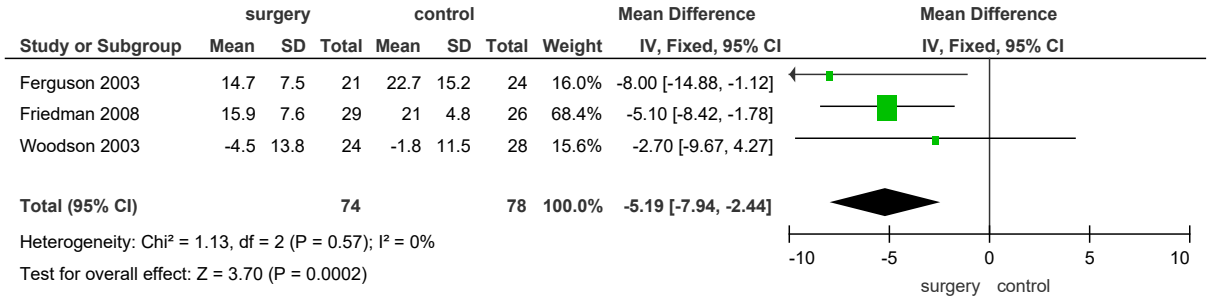


Figure 3: ESS (Epworth Sleepiness Scale) (0 to 24,lower is better)

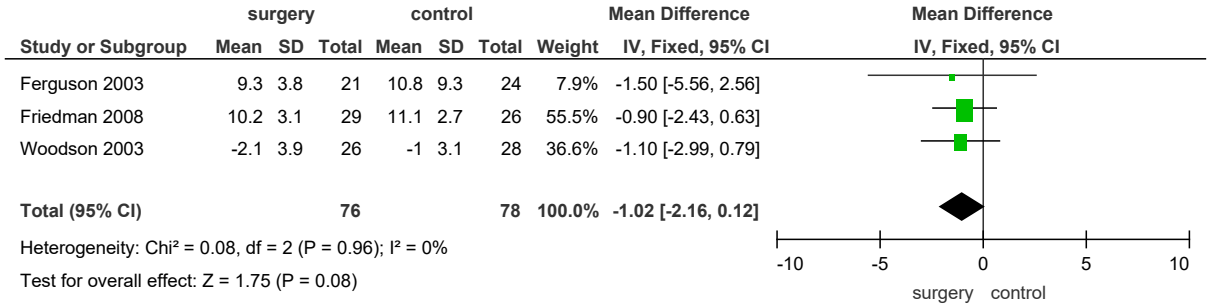


Figure 4: Vitality (SF-36) (0-100, higher is better)

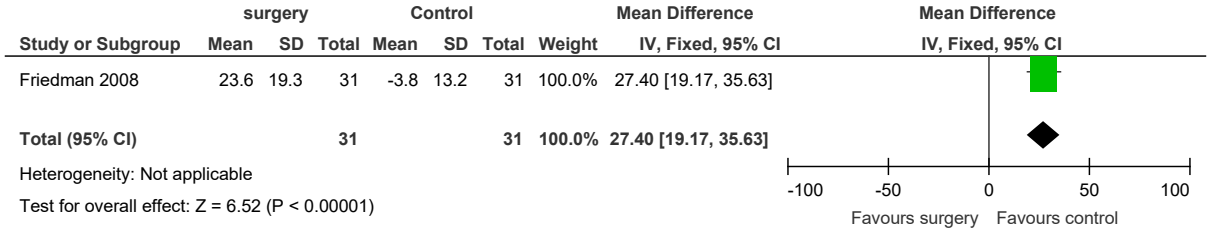


Figure 5: SF- 36- Physical health(0-100, higher is better)

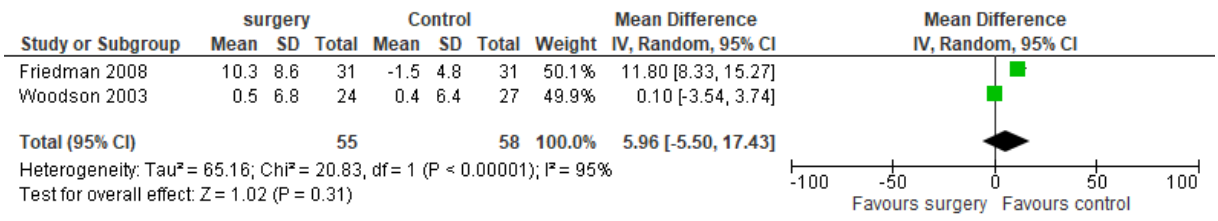


Figure 6: SF-36- mental health(0-100, higher is better)

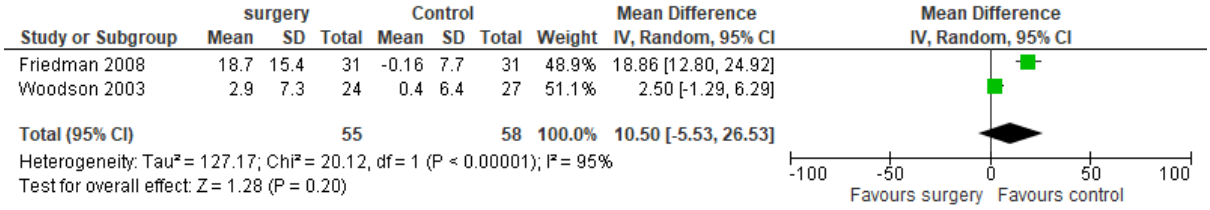


Figure 7: FOSQ (Functional Outcome of Sleep Questionnaire) (5-20, higher is better)

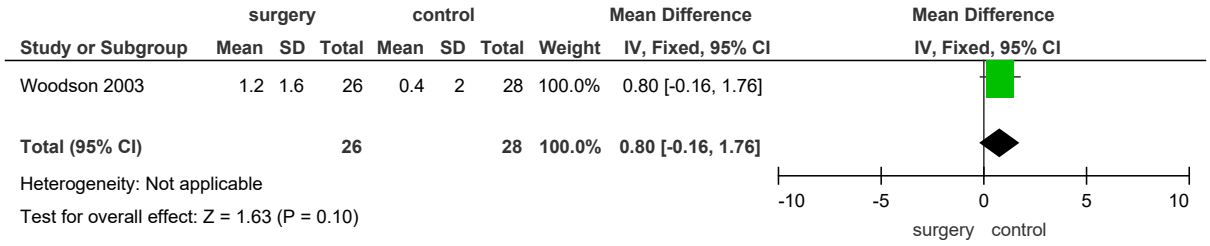


Figure 8: Quality of life (sleep apnoea quality of life index) (0-7, higher is better)

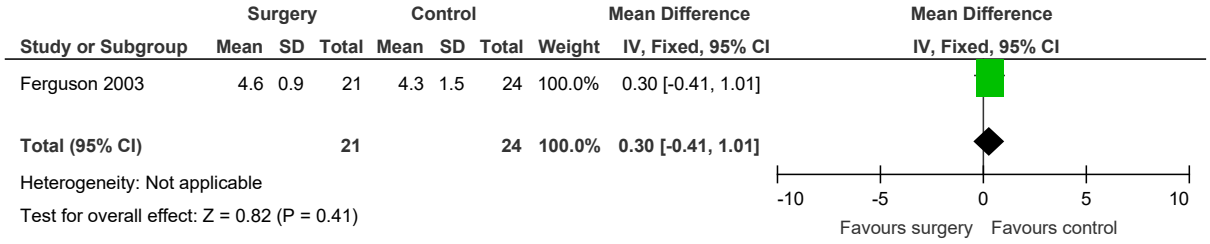


Figure 9: Dysphagia

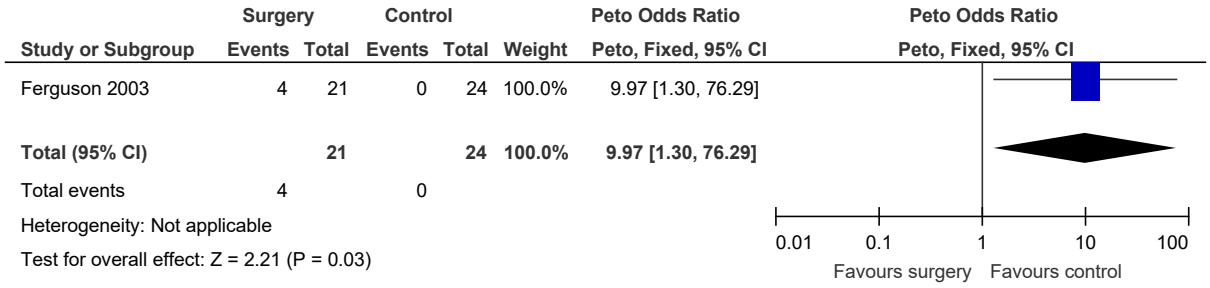


Figure 10: Infection

Study or Subgroup	Surgery Events	Surgery Total	Control Events	Control Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Ferguson 2003	4	21	0	24	100.0%	9.97 [1.30, 76.29]
Woodson 2003	0	26	0	28		Not estimable
Total (95% CI)		**47**		**52**	**100.0%**	**9.97 [1.30, 76.29]**
Total events	4		0			

Heterogeneity: Not applicable

Test for overall effect: Z = 2.21 (P = 0.03)

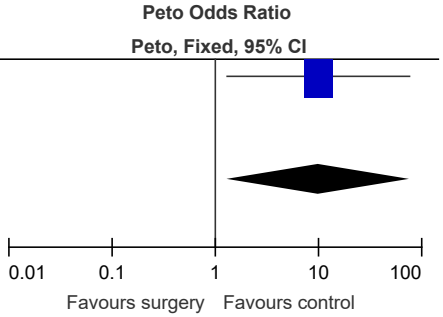


Figure 11: Bleeding (mild-severe)

Study or Subgroup	Surgery Events	Surgery Total	Control Events	Control Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Ferguson 2003	9	21	0	24	100.0%	13.72 [3.23, 58.37]
Woodson 2003	0	26	0	28		Not estimable
Total (95% CI)		**47**		**52**	**100.0%**	**13.72 [3.23, 58.37]**
Total events	9		0			

Heterogeneity: Not applicable

Test for overall effect: Z = 3.55 (P = 0.0004)

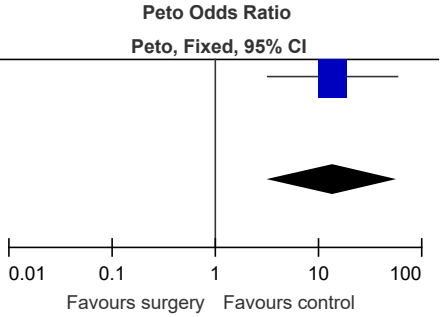


Figure 12: Pain

Study or Subgroup	Surgery Events	Surgery Total	Control Events	Control Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Ferguson 2003	17	21	0	24	100.0%	29.00 [8.78, 95.76]
Total (95% CI)		**21**		**24**	**100.0%**	**29.00 [8.78, 95.76]**
Total events	17		0			

Heterogeneity: Not applicable

Test for overall effect: Z = 5.53 (P < 0.00001)

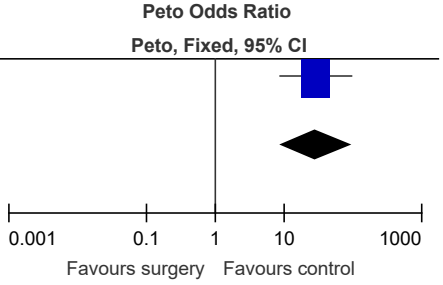


Figure 13: Nasal regurgitation

Study or Subgroup	Surgery Events	Surgery Total	Control Events	Control Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Ferguson 2003	5	21	0	24	100.0%	10.56 [1.67, 66.68]
Total (95% CI)		**21**		**24**	**100.0%**	**10.56 [1.67, 66.68]**
Total events	5		0			

Heterogeneity: Not applicable

Test for overall effect: Z = 2.51 (P = 0.01)

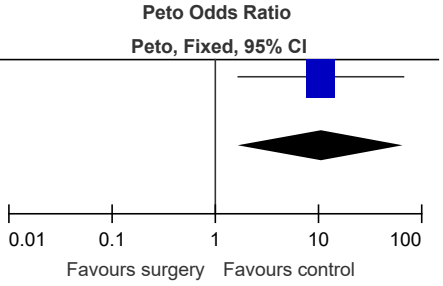


Figure 14: Ulcerations

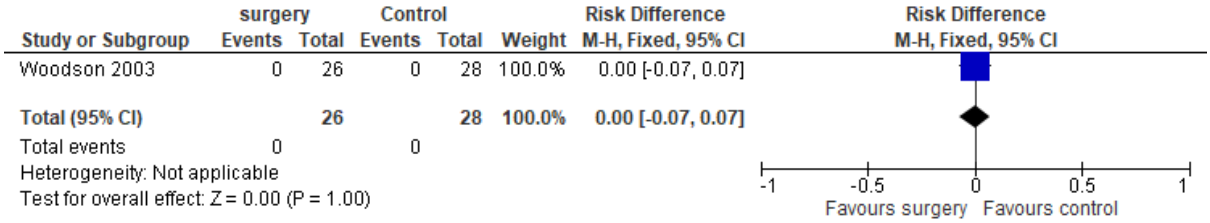


Figure 15: Pain at 1 week (10 cm visual analog scale range 0-10)

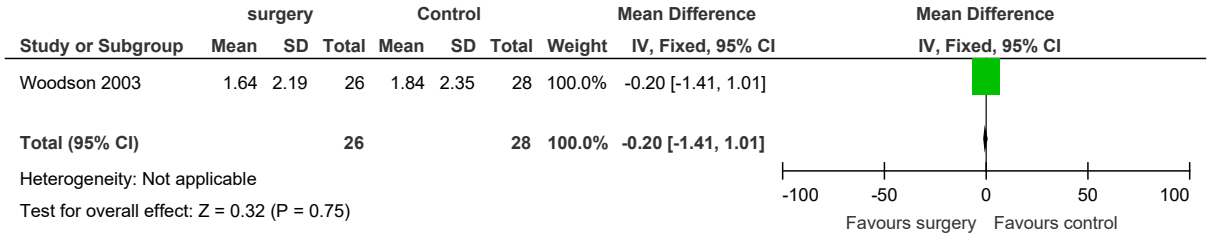


Figure 16: Pain at 3 weeks (10 cm visual analog scale range 0-10)

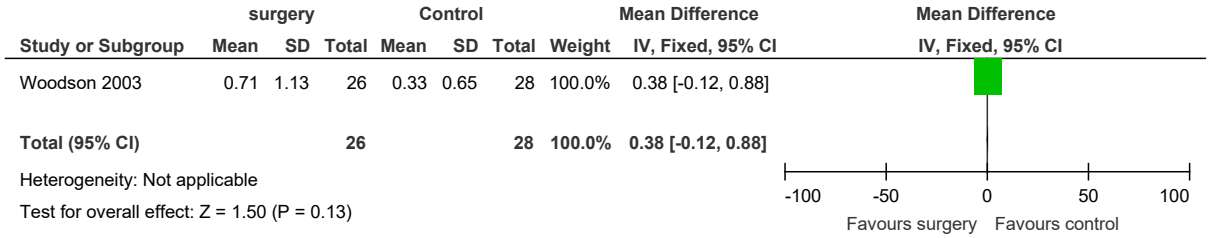


Figure 17: swallowing at 1 week (10 cm visual analog scale range 0-10)

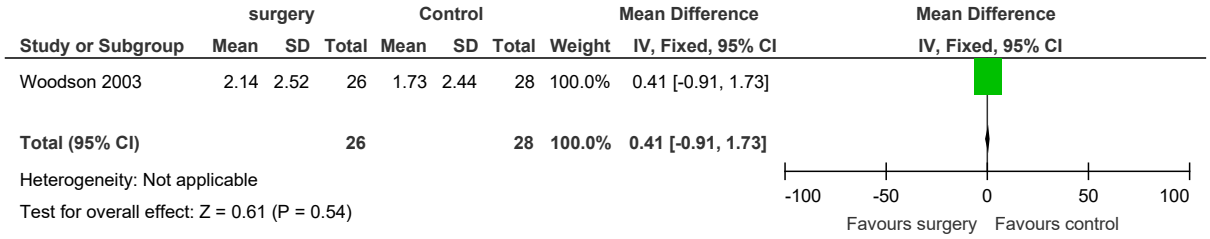
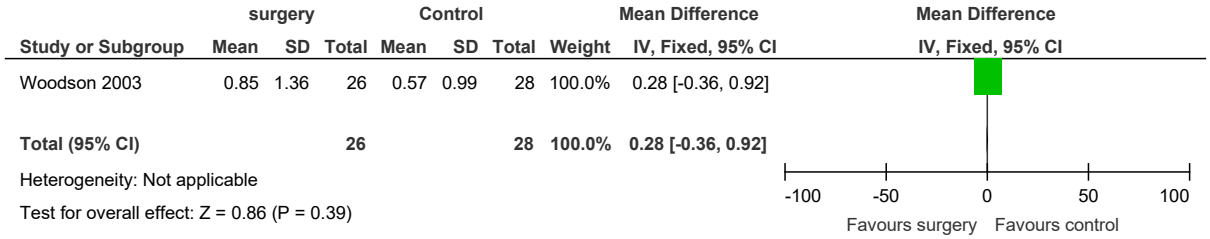


Figure 18: swallowing at 3 weeks (10 cm visual analog scale range 0-10)



E.2 Surgery vs conservative management/no surgery/sham surgery- Severe OSAHS

Figure 19: AHI (lower better) (su-group analysis- oro- pharyngeal vs nasal surgery)

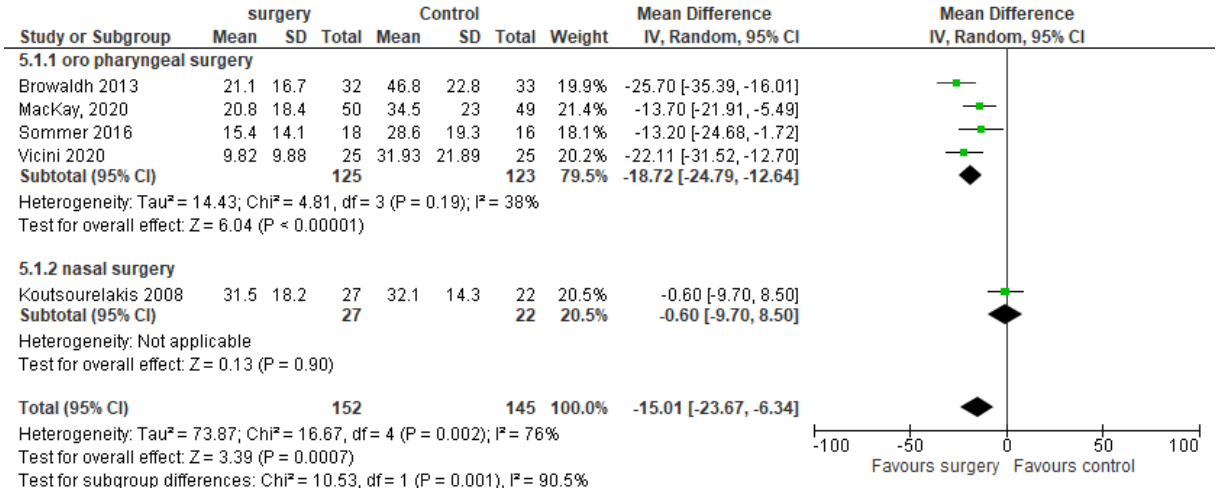


Figure 20: AHI (lower better) (su-group analysis- all people with OSAHS (including people who are tolerant and not tolerant to CPAP)vs not tolerant/adherent to CPAP)

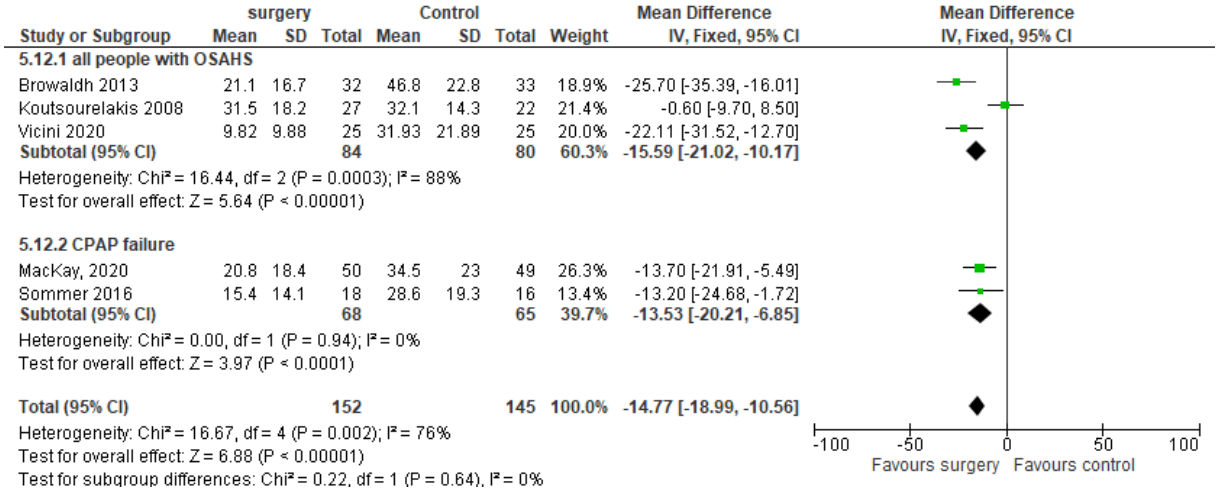


Figure 21: FOSQ (5-20, higher is better)

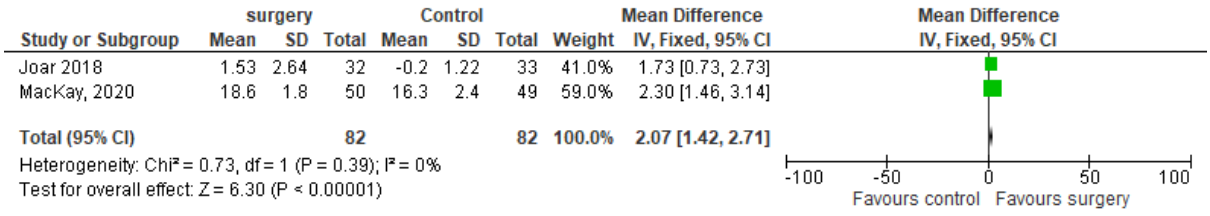


Figure 22: ESS (Epworth Sleepiness Scale) (0 to 24, lower is better) (sub-group analysis- oro-pharyngeal vs nasal surgery)

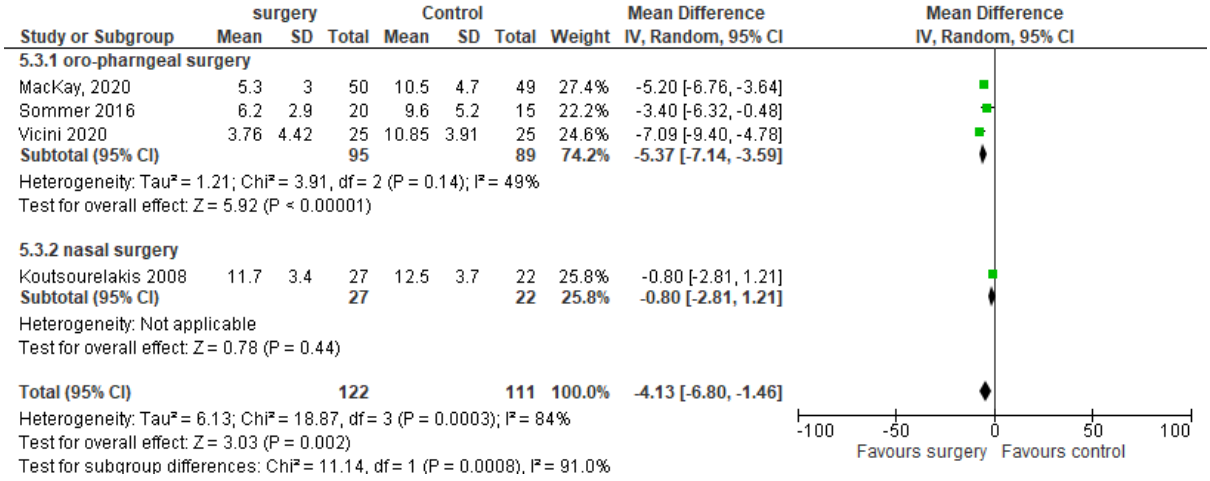


Figure 23: ESS (Epworth Sleepiness Scale) (0 to 24, lower is better) (sub-group analysis- all people with OSAHS (including people who are tolerant and not tolerant to CPAP)vs not tolerant/adherent to CPAP)

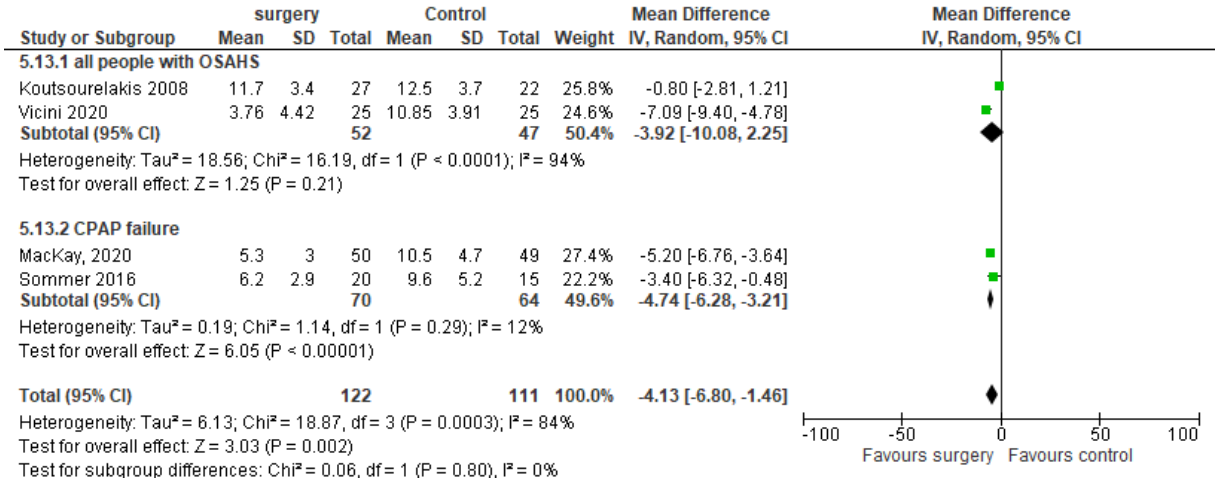


Figure 24: Velopharyngeal insufficiency (speech abnormalities)

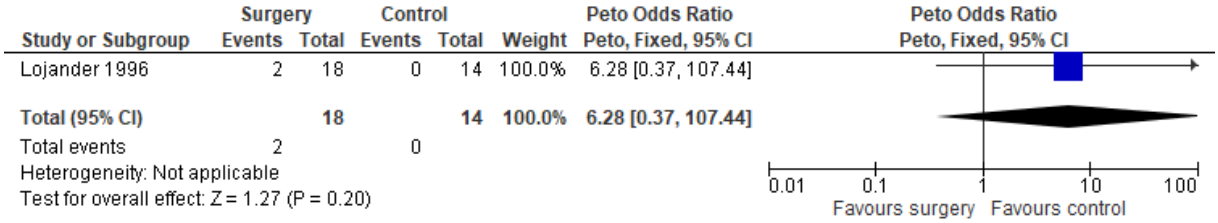


Figure 25: SpO2 (peripheral capillary oxygen saturation) (higher is better)

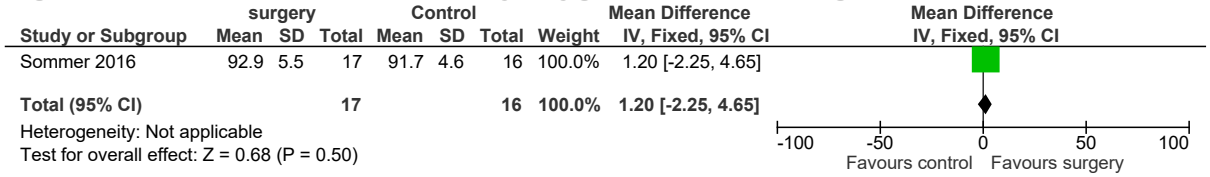


Figure 26: Serious adverse events [Lojander- CV events -non Q myocardial infarction and transient ischemic cerebral attack and Mackay 2020 – myocardial infarction and hematemesis]

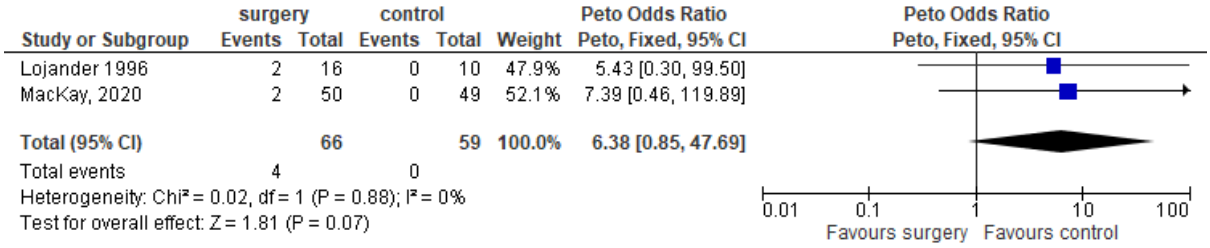


Figure 27: Tracheotomy

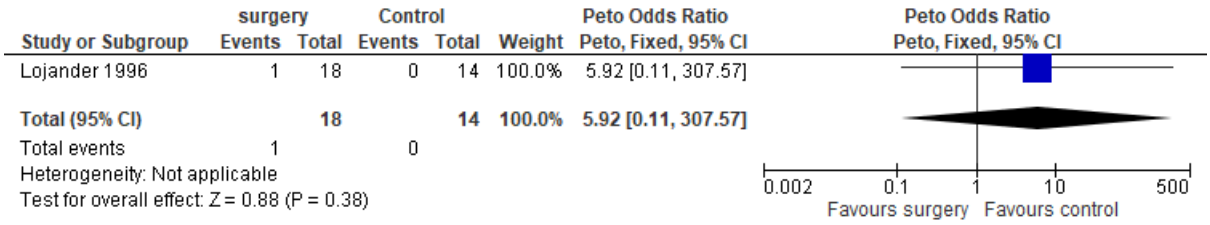


Figure 28: Re-operations

Study or Subgroup	surgery Events	surgery Total	Control Events	Control Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Lojander 1996	2	18	0	14	100.0%	6.28 [0.37, 107.44]
Total (95% CI)		**18**		**14**	**100.0%**	**6.28 [0.37, 107.44]**
Total events	2		0			

Heterogeneity: Not applicable

Test for overall effect: Z = 1.27 (P = 0.20)

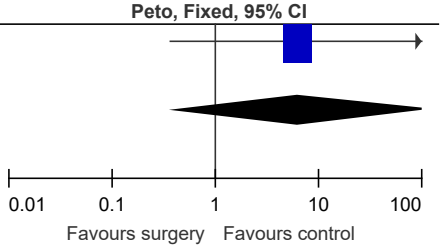


Figure 29: Mortality

Study or Subgroup	surgery Events	surgery Total	Control Events	Control Total	Weight	Risk Difference M-H, Fixed, 95% CI
Browaldh 2013	0	32	0	0		Not estimable
Total (95% CI)		**32**		**0**		**Not estimable**
Total events	0		0			

Heterogeneity: Not applicable

Test for overall effect: Not applicable

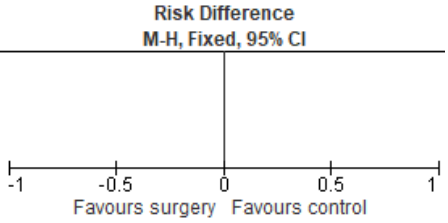


Figure 30: 24 h ambulatory systolic blood pressure

Study or Subgroup	surgery Mean	surgery SD	surgery Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Fixed, 95% CI
MacKay, 2020	120	12	50	124.7	13.6	49	100.0%	-4.70 [-9.76, 0.36]
Total (95% CI)			**50**			**49**	**100.0%**	**-4.70 [-9.76, 0.36]**

Heterogeneity: Not applicable

Test for overall effect: Z = 1.82 (P = 0.07)

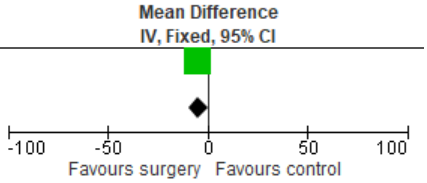
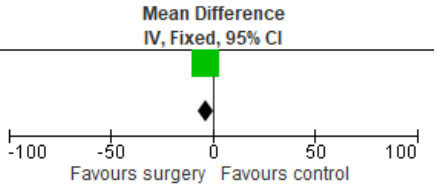


Figure 31: 24 hr ambulatory diastolic blood pressure

Study or Subgroup	surgery Mean	surgery SD	surgery Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Fixed, 95% CI
MacKay, 2020	74.1	8	50	77.7	10.9	49	100.0%	-3.60 [-7.37, 0.17]
Total (95% CI)			**50**			**49**	**100.0%**	**-3.60 [-7.37, 0.17]**

Heterogeneity: Not applicable

Test for overall effect: Z = 1.87 (P = 0.06)



E.3 Surgery versus APAP- Severe OSAHS

Figure 32: AHI (lower is better)

Study or Subgroup	surgery Mean	surgery SD	surgery Total	APAP Mean	APAP SD	APAP Total	Weight	Mean Difference IV, Fixed, 95% CI
Vincini 2010	8.1	7	25	6.3	1.9	25	100.0%	1.80 [-1.04, 4.64]
Total (95% CI)			**25**			**25**	**100.0%**	**1.80 [-1.04, 4.64]**

Heterogeneity: Not applicable

Test for overall effect: Z = 1.24 (P = 0.21)

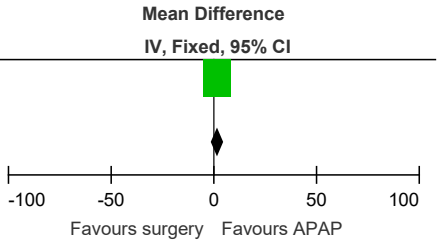


Figure 33: ESS (Epworth Sleepiness Scale) (0 to 24, higher is worse)

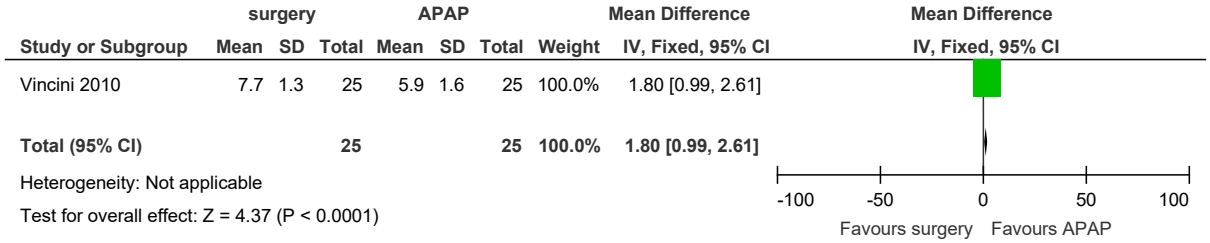


Figure 34: Bleeding

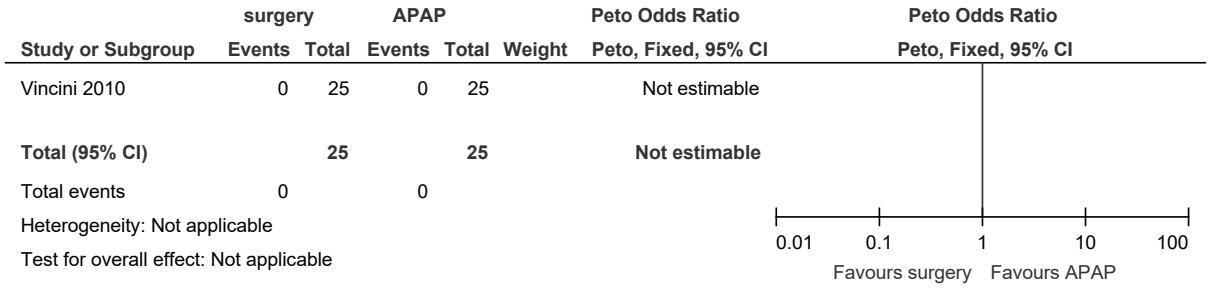


Figure 35: Dyspnoea

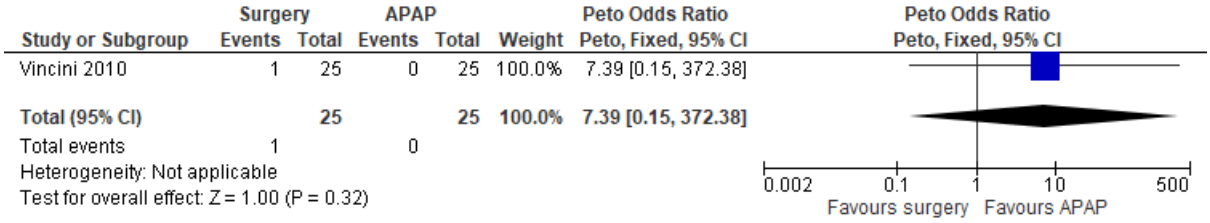


Figure 36: Persistent paraesthesia

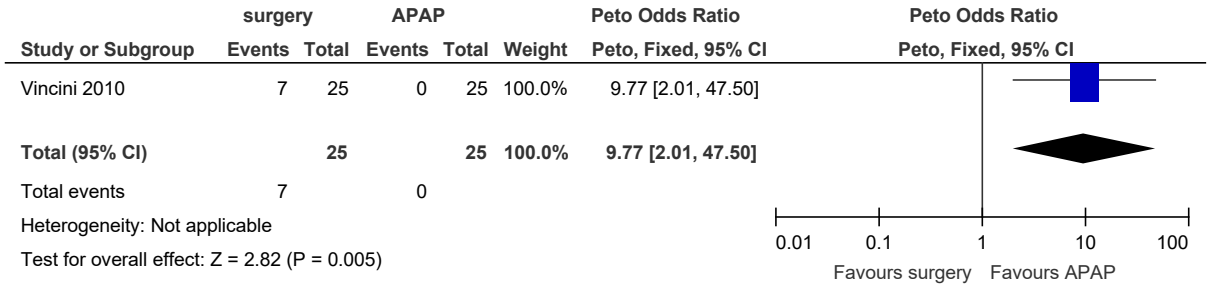
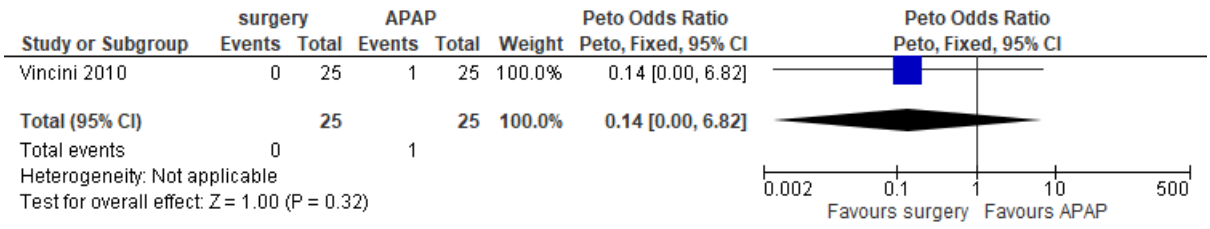


Figure 37: Lesion in facial skin (with APAP)



E.4 Surgery versus CPAP- Moderate OSAHS

Figure 38: SF-36 physical (0-100, higher is better)

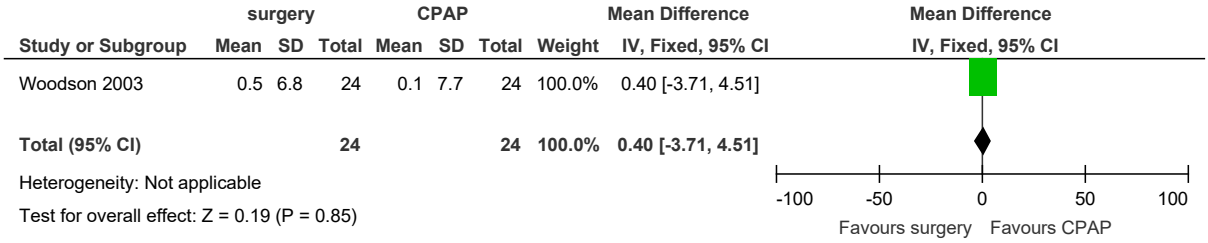


Figure 39: SF-36 mental (0-100, higher is better)

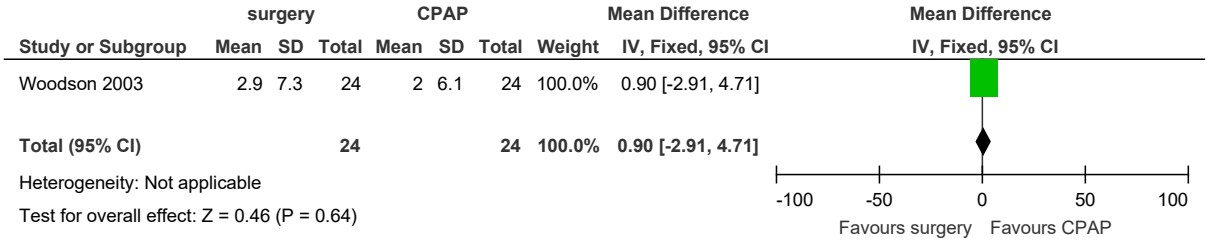


Figure 40: FOSQ (Functional Outcomes of Sleep Questionnaire) (5-20, higher is better)

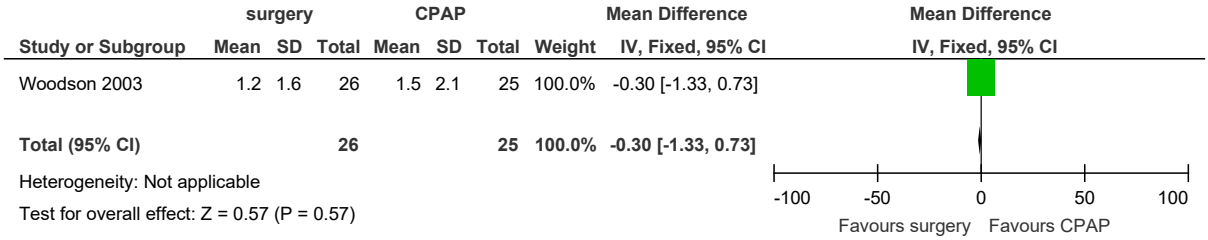
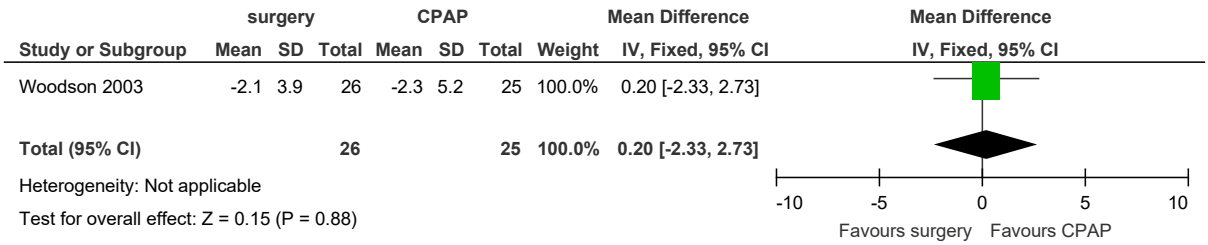


Figure 41: ESS (Epworth Sleepiness Score) (0 to 24, higher is worse)



E.5 Surgery versus oral devices- Moderate OSAHS

Figure 42: AHI (Apnoea Hypopnea Index) (lower is better)

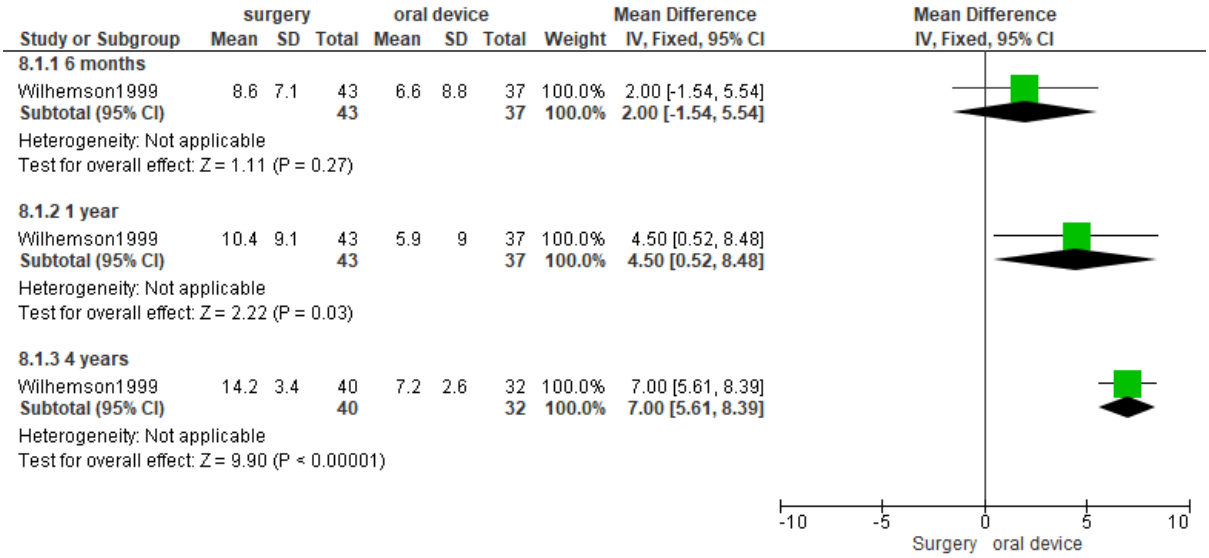


Figure 43: Oxygen desaturation index (ODI) (lower is better)

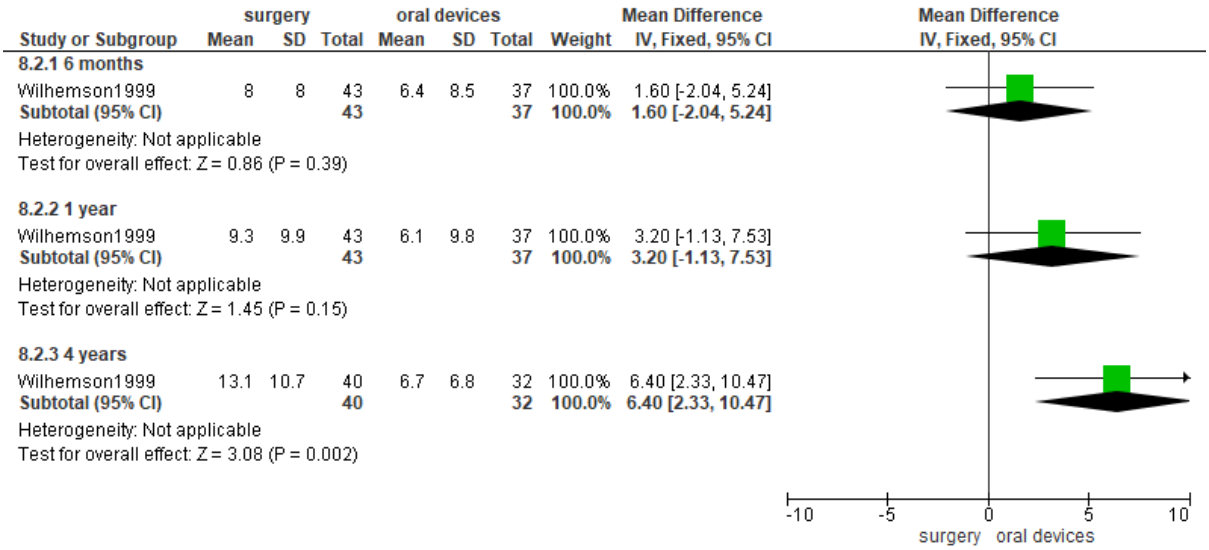


Figure 44: Quality of life: Vitality [minor symptoms evaluation profile MSE) at 1 year Scale 0-100 (lower is better)

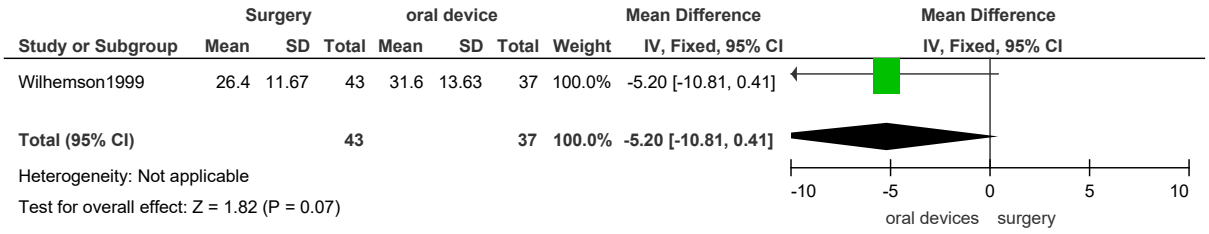


Figure 45: Quality of life: Sleep[minor symptoms evaluation profile MSE) at 1 year Scale 0-100 (lower is better)

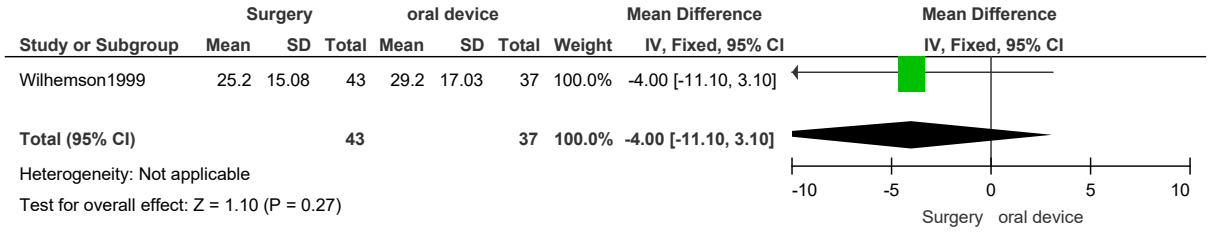


Figure 46: Quality of life: Contentment [minor symptoms evaluation profile MSE) at 1 year Scale 0-100 (lower is better)

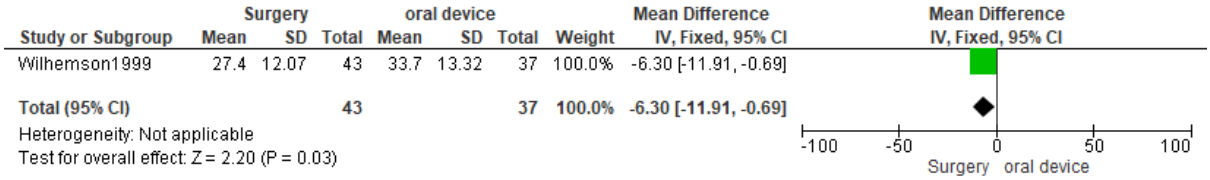


Figure 47: Dysphagia at 4 years

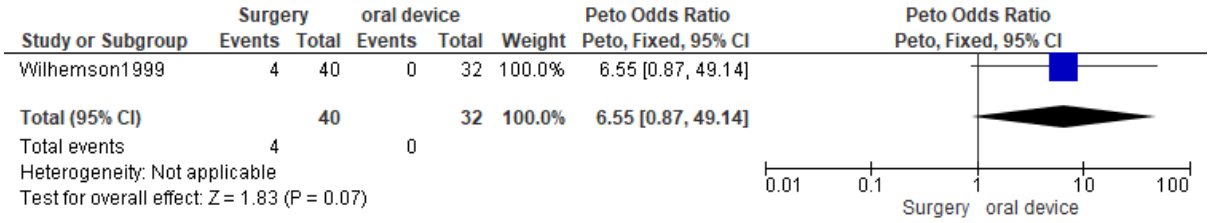
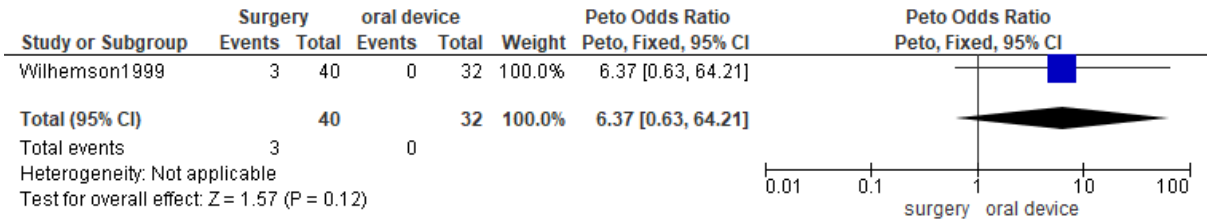


Figure 48: Nasopharyngeal regurgitation at 4 years



Appendix F: GRADE tables

Table 15: Clinical evidence profile: Surgery vs conservative management/no surgery/sham surgery- Moderate OSAHS [Category of surgery: oro-pharyngeal surgery]

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Surgery	Conservative management/no surgery/sham surgery- MODERATE	Relative (95% CI)	Absolute		
AHI (Better indicated by lower values)												
3	randomised trials	serious[1]	no serious inconsistency	serious indirectness[6]	serious[2]	None	74	78	-	MD 5.19 lower (7.94 to 2.44 lower)	⊕OOO VERY LOW	IMPORTANT
Epworth sleepiness score (Better indicated by lower values)												
3	randomised trials	serious[1]	no serious inconsistency	serious indirectness[6]	no serious imprecision	None	76	78	-	MD 1.02 lower (2.16 lower to 0.12 higher)	⊕⊕OO LOW	IMPORTANT
vitality (SF-36) (Better indicated by higher values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[6]	no serious imprecision	None	31	31	-	MD 27.4 higher (19.17 to 35.63 higher)	⊕⊕⊕O MODERATE	CRITICAL
SF- 36- Physical health (Better indicated by higher values)												
2	randomised trials	no serious risk of bias	very serious[3]	serious indirectness[6]	serious imprecision[2]	none	55	58	-	MD 5.96 higher (-5.50 lower to	⊕OOO VERY LOW	CRITICAL

										17.43 to 8.74 higher)		
SF-36- mental health (Better indicated by higher values)												
2	randomised trials	no serious risk of bias	very serious[3]	serious indirectness[6]	serious imprecision[2]	None	55	58	-	MD 10.50 higher (-5.53 lower to 26.53 higher)	⊕OOO VERY LOW	CRITICAL
FOSQ (Better indicated by higher values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[6]	no serious imprecision	None	26	28	-	MD 0.8 higher (0.16 lower to 1.76 higher)	⊕⊕⊕O MODERATE	CRITICAL
Quality of life (sleep apnoea quality of life index) (Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious indirectness[6]	no serious imprecision	None	21	24	-	MD 0.3 higher (0.41 lower to 1.01 higher)	⊕OOO VERY LOW	CRITICAL
Dysphagia												
1	randomised trials	very serious[1]	no serious inconsistency	serious indirectness[6]	no serious imprecision	None	4/21 (19%)	0%	OR 9.97 (1.3 to 76.29)[4]	190 more per 1000 (from 13 more to 360 more)[5]	⊕OOO VERY LOW	IMPORTANT
Infection												
2	randomised trials	serious[1]	no serious inconsistency	serious indirectness[6]	no serious imprecision	None	4/47 (8.5%)	0%	OR 9.97 (1.3 to 76.29)[4]	80 more per 1000 (from 1 fewer to 170 more)[5]	⊕⊕OO LOW	IMPORTANT
Bleeding (mild-severe)												
2	randomised trials	serious[1]	no serious inconsistency	serious indirectness[6]	no serious imprecision	None	9/47 (19.1%)	0%	OR 13.72 (3.23 to 58.37)[4]	190 more per 1000 (from 75 more to 300 more)[5]	⊕⊕OO LOW	IMPORTANT
Pain												

1	randomised trials	very serious[1]	no serious inconsistency	serious indirectness[6]	no serious imprecision	None	17/21 (81%)	0%	OR 29 (8.78 to 95.76)[4]	800 more per 1000 (from 630 more to 980 more)[5]	⊕OOO VERY LOW	IMPORTANT
Nasal regurgitation												
1	randomised trials	very serious[1]	no serious inconsistency	serious indirectness[6]	no serious imprecision	None	5/21 (23.8%)	0%	OR 10.56 (1.67 to 66.68)[4]	230 more per 1000 (from 40 more to 420 more)[5]	⊕OOO VERY LOW	IMPORTANT
Ulcerations												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[6]	no serious imprecision	None	0/26 (0%)	0%	not pooled	not pooled	⊕⊕⊕O MODERATE	IMPORTANT
Pain at 1 week (10 cm visual analog scale range 0-10) (Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[6]	very serious[2]	None	26	28	-	MD 0.2 lower (1.41 lower to 1.01 higher)	⊕OOO VERY LOW	IMPORTANT
Pain at 3 weeks (10 cm visual analog scale range 0-10) (Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[6]	serious[2]	None	26	28	-	MD 0.38 higher (0.12 lower to 0.88 higher)	⊕⊕OO LOW	IMPORTANT
swallowing at 1 week (10 cm visual analog scale range 0-10) (Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[6]	very serious[2]	None	26	28	-	MD 0.41 higher (0.91 lower to 1.73 higher)	⊕OOO VERY LOW	IMPORTANT
swallowing at 3 weeks (10 cm visual analog scale range 0-10) (Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[6]	very serious[2]	None	26	28	-	MD 0.28 higher (0.36 lower to 0.92 higher)	⊕OOO VERY LOW	IMPORTANT
Mortality												

Not reported												CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] Downgraded by one increment if the confidence interval crossed one MID and downgraded by two increments if the confidence interval crossed both MIDs. Established MIDs for SF-36 physical/mental- 2/3 ; FOSQ- 2 ; SAQLI – 2ESS -2.5 . GRADE default MID (0.5XSD) used for all other continous outcomes.

[3] I2=95%. Downgraded by 1 or 2 increments for heterogeneity, unexplained by subgroup analysis. Random effects analysis used
[4] Peto odds ratio used as zero events in one/both groups.
[5] Risk difference calculated in Revman

6 Downgraded by 1 or 2 increments because the majority of the evidence included an indirect or very indirect population respectively. The population was deemed to be indirect when the outcome included evidence from studies with mixed severity OSAHS populations.

Table 16: Clinical evidence profile: Surgery vs conservative management/no surgery/sham surgery- Severe OSAHS [Category of surgery: oro-pharyngeal surgery and nasal surgery]

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Surgery**	**Conservative management/no surgery/sham surgery- SEVERE**	**Relative (95% CI)**	**Absolute**		
AHI (all studies) (Better indicated by lower values)												
5	randomised trials	serious[1]	very serious[2]	serious indirectness[6]	No serious imprecision	none	152	145	-	MD 15.01 lower (23.67 to 6.34 lower)	⊕OOO VERY LOW	IMPORTANT
AHI – sub-group analysis oro pharyngeal surgery (Better indicated by lower values)												
4	randomised trials	serious[1]	no serious inconsistency	Serious[6]	no serious imprecision	none	125	123	-	MD 18.724 lower (24.79 to 12.64 lower)	⊕⊕OO LOW	IMPORTANT

AHI - sub-group analysis nasal surgery (Better indicated by lower values)											
1	randomised trials	serious[1]	no serious inconsistency	Serious[6]	no serious imprecision	none	27	22	-	MD 0.6 lower (9.7 lower to 8.5 higher)	⊕⊕OO LOW	IMPORTANT
AHI sub-group analysis- all people with OSAHS (including people who are tolerant and not tolerant to CPAP) (Better indicated by lower values)											
3	randomised trials	serious[1]	serious[2]	Serious[6]	no serious imprecision	none	84	80	-	MD 15.59 lower (21.02 to 10.17 lower)	⊕OOO VERY LOW	IMPORTANT
AHI sub-group analysis- not tolerant/adherent to CPAP (Better indicated by lower values)											
2	randomised trials	serious[1]	no serious inconsistency	Serious[6]	Serious3	none	68	65	-	MD 13.53 lower (20.21 to 6.85 lower)	⊕OOO VERY LOW	IMPORTANT
FOSQ (Better indicated by higher values)											
2	randomised trials	serious[1]	no serious inconsistency	serious indirectness[6]	serious[3]	none	82	82	-	MD 2.07 higher (1.42 to 2.71 higher)	⊕OOO VERY LOW	CRITICAL
ESS (all studies) (Better indicated by lower values)											
4	randomised trials	serious[1]	Very serious [2]	serious indirectness[6]	no serious imprecision	none	122	111	-	MD 4.13 lower (6.80 lower to 1.46 lower)	⊕OOO VERY LOW	IMPORTANT
ESS - sub-group analysis oro-pharngeal surgery (Better indicated by lower values)											
2	randomised trials	serious[1]	very serious [2]	Serious[6]	no serious imprecision	none	95	89	-	MD 5.37 lower (7.14 to 3.59 lower)	⊕OOO VERY LOW	IMPORTANT
ESS - sub-group analysis nasal surgery (Better indicated by lower values)											
1	randomised trials	serious[1]	no serious inconsistency	Serious[6]	Serious[3]	none	27	22	-	MD 0.8 lower (2.81 lower to 1.21 higher)	⊕⊕⊕O VERY LOW	IMPORTANT
ESS sub-group analysis- all people with OSAHS (including people who are tolerant and not tolerant to CPAP) (Better indicated by lower values)											

2	randomised trials	serious[1]	very serious[2]	Serious[6]	no serious imprecision	none	52	47	-	MD 3.92 lower (10.08 lowerto 2.25 higher)	⊕OOO VERY LOW	IMPORTANT
ESS sub-group analysis- not tolerant/adherent to CPAP (Better indicated by lower values)												
2	randomised trials	serious[1]	no serious inconsistency	Serious[6]	no serious imprecision	none	70	64	-	MD 4.74 lower (6.28 to 3.21 lower)	⊕⊕OO LOW	IMPORTANT
velopharyngeal insufficiency (speech abnormalities)												
1	randomised trials	serious[1]	no serious inconsistency	serious indirectness[6]	very serious[3]	none	2/18 (11.1%)	0%	OR 6.28 (0.37 to 107.44)[4]	111 more per 1000 (from 60 fewer to 280 more)[5]	⊕OOO VERY LOW	IMPORTANT
SpO2 (Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious indirectness[6]	very serious[3]	none	17	16	-	MD 1.2 higher (2.25 lower to 4.65 higher)	⊕OOO VERY LOW	IMPORTANT
Serious adverse events [Lojander 1996- CV events (non Q myocardial infarction and transient ischemic cerebral attack; MacKay 2020- myocardial infraction and hematemesis												
2	randomised trials	serious[1]	no serious inconsistency	serious indirectness[6]	serious[3]	none	4/66 (6.1%)	0/59 (0%)	OR 6.38 (0.85 to 47.69)[4]	60 more per 1000 (from3 fewer to 120 more)[5]	⊕OOO VERY LOW	IMPORTANT
Tracheotomy												
1	randomised trials	serious[1]	no serious inconsistency	serious indirectness[6]	very serious[3]	none	1/18 (5.6%)	0%	OR 5.92 (0.11 to 307.57)[4]	50 more per 1000 (from 90 fewer to 200 more)[5]	⊕OOO VERY LOW	IMPORTANT
re-operations												
1	randomised trials	serious[1]	no serious inconsistency	serious indirectness[6]	very serious[3]	none	2/18 (11.1%)	0%	OR 6.28 (0.37 to 107.44)[4]	111 more per 1000 (from 60 fewer to 280 more)[5]	⊕OOO VERY LOW	IMPORTANT
Mortality												

1	randomised trials	serious[1]	no serious inconsistency	serious indirectness[6]	Not estimable	none	32	33	-	Not estimable Zero events in both groups	⊕⊕OO LOW	CRITICAL
24 h ambulatory systolic blood pressure (Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[3]	serious[4]	None	50	49	-	MD 4.7 lower (9.76 lower to 0.36 higher)	⊕OOO VERY LOW	IMPORTANT
24 h ambulatory diastolic blood pressure (Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[3]	serious[4]	None	50	49	-	MD 3.6 lower (7.37 lower to 0.17 higher)	1	IMPORTANT

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] Downgraded by 1 or 2 increments because: for heterogeneity, unexplained by subgroup analysis. Random effects analysis used
[3] Downgraded by one increment if the confidence interval crossed one MID and downgraded by two increments if the confidence interval crossed both MIDs. Established MIDs for SF-36 physical/mental- 2/3 ; FOSQ- 2 ; ESS -2.5 . GRADE default MID(0.5XSD) used for all other continuous outcomes.
[4] Peto odds ratio used when zero events in one/both groups.
[5] Risk difference calculated in Revman

6 Downgraded by 1 or 2 increments because the majority of the evidence included an indirect or very indirect population respectively. The population was deemed to be indirect when the outcome included evidence from studies with mixed severity OSAHS populations.

Table 17: Clinical evidence profile: surgery versus APAP- Severe OSAHS [category of surgery- Skeletal Framework Surgery]

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Surgery versus APAP- SEVERE**	**Control**	**Relative (95% CI)**	**Absolute**		
AHI (Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	None	25	25	-	MD 1.8 higher (1.04 lower to 4.64 higher)	⊕OOO VERY LOW	IMPORTANT
ESS (Better indicated by lower values)												

1	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	serious[2]	None	25	25	-	MD 1.8 higher (0.99 to 2.61 higher)	⊕OOO VERY LOW	IMPORTANT
Bleeding												
1	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	None	0/25 (0%)	0%	not pooled	not pooled	⊕⊕OO LOW	IMPORTANT
Dyspnoea												
1	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	None	1/25 (4%)	0%	OR 7.39 (0.15 to 372.38)[3]	40 more per 1000 (from 60 fewer to 140 more)[4]	⊕OOO VERY LOW	IMPORTANT
Persistent paresthesia												
1	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	None	7/25 (28%)	0%	OR 9.77 (2.01 to 47.5)[3]	280 more per 1000 (from 90 more to 460 more)[4]	⊕⊕OO LOW	IMPORTANT
lesion in facial skin (with APAP)												
1	randomised trials	very serious[1]	no serious inconsistency	no serious indirectness	very serious[2]	None	0/25 (0%)	4%	OR 0.14 (0 to 6.82)[3]	40 fewer per 1000 (from 14 fewer to 60 more)[4]	⊕OOO VERY LOW	IMPORTANT
Mortality												
Not reported												CRITICAL

[1] Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias
[2] Downgraded by one increment if the confidence interval crossed one MID and downgraded by two increments if the confidence interval crossed both MIDs. Established MIDs for SF-36 physical/mental- 2/3 ; FOSQ- 2 ; ESS -2.5 . GRADE default MID(0.5XSD) used for all other continuous outcomes.
[3] Peto odds ratio used when zero events in one/both groups.
[4] risk difference calculated in Revman

Table 18: Clinical evidence profile: Surgery versus CPAP- Moderate OSAHS [category of surgery- oro-pharyngeal surgery]

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Surgery versus CPAP- MODERATE	Control	Relative (95% CI)	Absolute		
SF-36 physical (Better indicated by higher values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[2]	very serious[1]	none	24	24	-	MD 0.4 higher (3.71 lower to 4.51 higher)	⊕OOO VERY LOW	CRITICAL
SF-36 mental (Better indicated by higher values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[2]	serious[1]	none	24	24	-	MD 0.9 higher (2.91 lower to 4.71 higher)	⊕⊕OO LOW	CRITICAL
FOSQ (Functional Outcomes of Sleep Questionnaire) (Better indicated by higher values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[2]	no serious imprecision	none	26	25	-	MD 0.3 lower (1.33 lower to 0.73 higher)	⊕⊕⊕O MODERATE	CRITICAL
ESS (Epworth Sleepiness Score) (Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[2]	serious[1]	none	26	25	-	MD 0.2 higher (2.33 lower to 2.73 higher)	⊕⊕OO LOW	IMPORTANT
Mortality												
Not reported												CRITICAL

1 Downgraded by one increment if the confidence interval crossed one MID and downgraded by two increments if the confidence interval crossed both MIDs. Established MIDs for SF-36 physical/mental- 2/3 ; FOSQ- 2 ; ESS -2.5 .. GRADE default MID(0.5XSD) used for all other continuous outcomes.
2Downgraded by 1 or 2 increments because the majority of the evidence included an indirect or very indirect population respectively. The population was deemed to be indirect when the outcome included evidence from studies with mixed severity OSAHS populations.

Table 19: Clinical evidence profile: Surgery versus oral devices- Moderate OSAHS

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Surgery versus oral devices- MODERATE	Control	Relative (95% CI)	Absolute		
AHI (Apnoea Hypopnea Index) - 6 months (Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[4]	very serious[1]	none	43	37	-	MD 2 higher (1.54 lower to 5.54 higher)	⊕OOO VERY LOW	IMPORTANT
AHI (Apnoea Hypopnea Index) - 1 year (Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[4]	serious[1]	none	43	37	-	MD 4.5 higher (0.52 to 8.48 higher)	⊕⊕OO LOW	IMPORTANT
AHI (Apnoea Hypopnea Index) - 4 years (Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[4]	no serious imprecision	none	40	32	-	MD 7 higher (5.61 to 8.39 higher)	⊕⊕⊕O MODERATE	IMPORTANT
Oxygen desaturation index (ODI) - 6 months (Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[4]	very serious[1]	none	43	37	-	MD 1.6 higher (2.04 lower to 5.24 higher)	⊕OOO VERY LOW	IMPORTANT
Oxygen desaturation index (ODI) - 1 year (Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[4]	serious[1]	none	43	37	-	MD 3.2 higher (1.13 lower to 7.53 higher)	⊕⊕OO LOW	IMPORTANT
Oxygen desaturation index (ODI) - 4 years (Better indicated by lower values)												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[4]	serious[1]	none	40	32	-	MD 6.4 higher (2.33 to 10.47 higher)	⊕⊕OO LOW	IMPORTANT
Quality of life: Vitality [minor symptoms evaluation profile MSE) at 1 year (Better indicated by lower values) Scale 0-100												

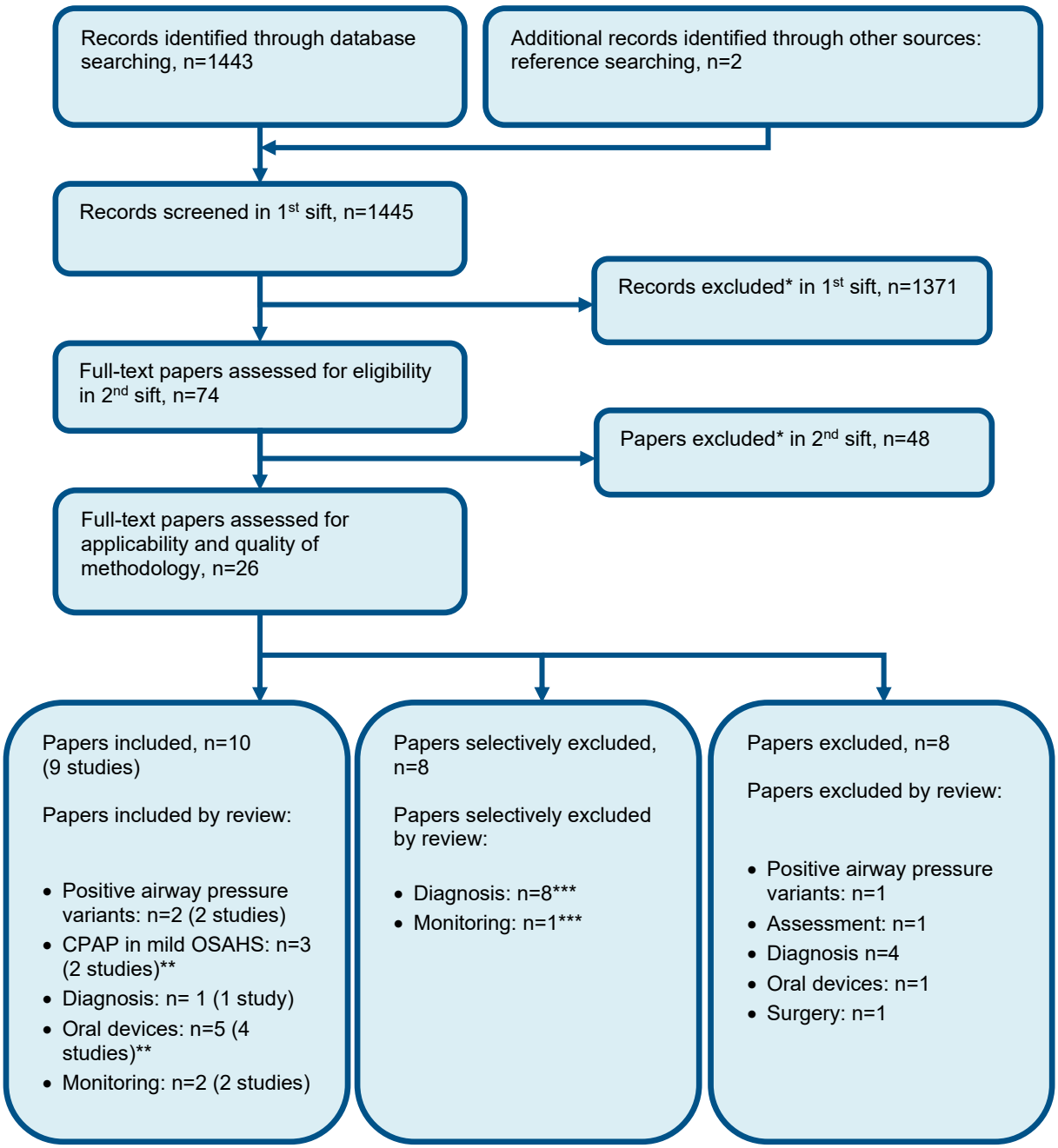
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[4]	serious[1]	none	43	37	-	MD 5.2 lower (10.81 lower to 0.41 higher)	⊕⊕OO LOW	CRITICAL
Quality of life: Sleep[minor symptoms evaluation profile MSE) at 1 year (Better indicated by lower values) Scale 0-100												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[4]	very serious[1]	none	43	37	-	MD 4 lower (11.1 lower to 3.1 higher)	⊕OOO VERY LOW	CRITICAL
Quality of life: Contentment [minor symptoms evaluation profile MSE) at 1 year (Better indicated by lower values) Scale 0-100												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[4]	serious[1]	none	43	37	-	MD 6.3 lower (11.91 to 0.69 lower)	⊕⊕OO LOW	CRITICAL
Dysphagia at 4 years												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[4]	no serious imprecision	none	4/40 (10%)	0%	OR 6.55 (0.87 to 49.14)[2]	100 more per 1000 (from 4 fewer to 200 more)[3]	⊕⊕⊕O MODERATE	IMPORTANT
Nasopharyngeal regurgitation at 4 years												
1	randomised trials	no serious risk of bias	no serious inconsistency	serious indirectness[4]	serious[1]	none	3/40 (7.5%)	0%	OR 6.37 (0.63 to 64.21)[2]	75 more per 1000 (from 200 fewer to 170 more)[3]	⊕⊕OO LOW	IMPORTANT
Mortality												
Not reported												CRITICAL

[1] Downgraded by one increment if the confidence interval crossed one MID and downgraded by two increments if the confidence interval crossed both MIDs. Established MIDs for SF-36 physical/mental- 2/3 ; FOSQ- 2 ; ESS -2.5 .. GRADE default MID(0.5XSD) used for all other continous outcomes.
[2] Peto odds ratio used when zero events in one/both groups.
[3] Risk difference calculated in Revman

4 Downgraded by 1 or 2 increments because the majority of the evidence included an indirect or very indirect population respectively. The population was deemed to be indirect when the outcome included evidence from studies with mixed severity OSAHS populations.

Appendix G: Health economic evidence selection

Figure 49: Flow chart of health economic study selection for the guideline



* Non-relevant population, intervention, comparison, design or setting; non-English language
** Two studies (in three papers) were included for two different questions
*** One study was considered for two different questions

Appendix H: Health economic evidence tables

None.

Appendix I: Excluded studies

I.1 Excluded clinical studies

Table 20: Studies excluded from the clinical review

Study	Exclusion reason
Ayers 2016[1]	Systematic review. Screened for relevant references
Back 2009[2]	No useable outcomes
Balcerzak 2005[3]	Not in English
Baradaranfar 2015[4]	Not RCT
Bayir 2016[5]	No outcomes of interest. Study assessed effects of anterior palatoplasty on floppy eyelid syndrome patients with obstructive sleep apnoea.
Binar 2018[6]	Systematic review. Screened for relevant references.
Bostanci 2016[7]	Systematic review- screened for relevant references
Bridgman 2000[8]	Systematic review- screened for relevant references
Camacho 2014[10]	Systematic review- screened for relevant references.
Camacho 2015[14]	Systematic review. Screened for relevant references.
Camacho 2015[15]	Systematic review. Screened for relevant references.
Camacho 2016[11]	Systematic review. Screened for relevant references.
Camacho 2017[12]	Systematic review. Screened for relevant references.
Camacho 2019[13]	Systematic review. Screened for relevant references
Caples 2010[16]	Systematic review. Screened for relevant references.
Certal 2015[17]	Systematic review- screened for relevant references
Chisholm 2007[18]	Inappropriate study design- case control study

Choi 2016[19]	Systematic review. Screened for relevant references.
Couch 2005[20]	Article
DRKS 2010[22]	Trial registration. Not a study.
Fernandez-Ferrer 2015[24]	Systematic review- screened for relevant references.
Franklin 2009[25]	Systematic review. Screened for relevant references.
Friedman 2006[26]	Abstract
Gaddam 2014[28]	Screened for relevant references
Gakwaya 2011[29]	abstract only
Gao 2019[30]	Systematic review. Screened for relevant references.
Gerek 2005[31]	Not in English
Halle 2017[32]	Literature review- screened for relevant references
Holmlund 2016[34]	Not RCT
Ho 2020[33]	Review- screened for relevant references
Holty 2010[35]	Systematic review- screened for relevant references.
Iftikhar 2017[36]	NMA- does not include surgery.
Ishii 2015[37]	Systematic review. Screened for relevant references.
Justin 2016[39]	Systematic review. Screened for relevant references.
Kompelli 2019[40]	Systematic review. Screened for relevant references.
Kotecha 2014[41]	Article
Koutsourelakis 2008[43]	Conference Abstract
Li 2008[44]	Not in English
Lu 2017[46]	Incorrect interventions. Assessment of alternative therapy for obstructive sleep apnoea.
Marzetti 2013[48]	Inappropriate comparison. Anterior palatoplasty (AP) compared with uvulopalatal flap (UPF).
Meccariello 2017[50]	Systematic review- screened for relevant references
Miller 2017[51]	Systematic review- screened for relevant references
Mulholland 2019[52]	Systematic review- screened for relevant references

Murphey 2015[54]	Systematic review. Screened for relevant references.
Murphey 2016[53]	Retrospective review
Mwenge 2015[55]	Literature review- screened for relevant references
NCT 2012[57]	Citation only
Neruntarat 2010[58]	comparison of two types of surgery
Noller 2017[59]	Systematic review- screened for relevant references.
Pang 2016[61]	Systematic review- screened for relevant references
Pang 2018[60]	Systematic review. Screened for relevant references.
Pang 2018[62]	Systematic review- screened for relevant references
Pengo 2016[63]	Incorrect interventions
Pietzsch 2015[64]	Only cost-effectiveness data.
Reckley 2018[66]	Overview of systematic reviews. Screened for relevant references.
Rojo-Sanchis 2018[67]	Screened for relevant references.
Sharma 2019[68]	Systematic review- screened for relevant references
Sommer 2012[69]	Not in English
Song 2016[71]	Systematic review- screened for relevant references
Steward 2013[72]	Abstract
Strollo 2013[73]	Conference Abstract
Strollo 2014[75]	Inappropriate study design. Cohort study.
Strollo 2015[74]	Not RCT
Sundaram 2005[76]	Screened for relevant references
Tang 2013[77]	Not in English
Terris 2002[79]	Not RCT
Thomas 2002[80]	Conference Abstract
Veer 2014[81]	Meta-analysis - screened for relevant references
Volner 2017[84]	Systematic review- screened for relevant references
Walker-Engstrom 2001[86]	abstract only

Wu 2017[90]	Systematic review. Screened for relevant references.
Yang 2015[91]	Inappropriate comparison - surgery vs surgery+oral device
Zhao 2013[93]	Not in English
Zhang 2019[92]	Meta-analysis- not appropriate comparison –oral devices vs CPAP.
Zhu 2018[94]	Not RCT

I.2 Excluded health economic studies

Published health economic studies that met the inclusion criteria (relevant population, comparators, economic study design, published 2003 or later and not from non-OECD country or USA) but that were excluded following appraisal of applicability and methodological quality are listed below:

Table 21: Studies excluded from the health economic review

Reference	Reason for exclusion
Pietzsch 2019[65]	This study was rated as having very serious limitations because it did not use randomised evidence.

Appendix J: Research recommendations

J.1 Upper airway surgery in people unable to tolerate or adhere to CPAP

Research question: What is the clinical and cost effectiveness of upper airway surgical interventions for people with obstructive sleep apnoea/hypopnoea syndrome who are unable to tolerate or adhere to CPAP?

Why this is important

Upper airway surgery has a definitive role in “selected” patients who are unable to tolerate or adhere to CPAP therapy. The aim of upper airway surgery is two-fold: firstly, in some cases the surgery may completely resolve the upper airway obstruction and alleviate the need of CPAP therapy but more commonly the surgery may be useful in facilitating CPAP therapy and thus improving compliance and adherence. The aim of the upper airway surgery is to improve the anatomical dimensions of the pharyngeal lumen and by doing so it would potentially reduce the pressure requirement of CPAP treatment and make this more tolerant. Traditionally there has been a paucity of formal randomised controlled studies, although one recently published has shown that this type of multi-centre trial is possible in using a standardised surgical technique in appropriately selected patients. Further studies of this type in a European population are therefore needed to determine efficacy of these specified surgical techniques and to help with patient selection. Patient selection is crucial and this is both in terms of general patient factor (e.g. BMI and co-morbidities) and specific anatomical issues that need addressing Longer term post-operative follow up is essential, ideally beyond one year The newer techniques such as the hypoglossal nerve stimulation is not currently available in UK and one limiting factor in this is the exceptionally high cost of the treatment compared to CPAP therapy.

Criteria for selecting high-priority research recommendations:

PICO question	Population: People (16 and older) with obstructive sleep apnoea/hypopnoea syndrome (OSAHS) who are unable to tolerate or adhere to CPAP. Population to be to be stratified based on severity: mild, moderate and severe OSAHS based on AHI/ODI. Intervention(s): Upper airway surgery: • nasal surgery (to include septal surgery and turbinate surgery as well as surgery for nasal polyps). • oro-pharyngeal surgery - will include tonsillectomy on its own or combined with any palatal surgery, UPPP, any form of palatoplasty, expansion sphincter palatoplasty, laser or radiofrequency palate surgery. • multi-level surgery addressing both soft palate and tongue base simultaneously and in some cases combined with nasal surgery • trans-oral robotic surgery for tongue and/or epiglottis • hypoglossal nerve stimulation/upper airway stimulation • skeletal framework surgery - mainly maxillo-mandibular advancement

	Comparison: • non-surgical intervention (positional modifiers, oral devices) • no intervention/usual care Outcome(s): Critical • generic or disease specific validated quality of life measures (continuous) • mortality (dichotomous) Important • sleepiness scores (continuous, e.g. Epworth) • apnoea-hypopnoea index (continuous) • oxygen desaturation index (continuous) • CO2 control (continuous) • if CPAP still required, pressure reduction following surgery • permanent adverse effects (e.g. neural dysfunction, nasopharyngeal incompetence, globus sensation, dichotomous) • reversible adverse effects (e.g. pain, infection, taste disturbance, secondary bleeding, dichotomous) • driving outcomes (continuous) • neurocognitive outcomes (continuous) • impact on co-existing conditions: ○ HbA1c for diabetes (continuous) ○ cardiovascular events for cardiovascular disease (dichotomous) ○ systolic blood pressure for hypertension (continuous) Sub-group analysis: • clinical, anatomical and physiological phenotypes of OSAHS including gender
Importance to patients or the population	The research will allow a consistentevidence based approach to the management of people with OSAHS who are unable to tolerate or adhere to CPAP.
Relevance to NICE guidance	This research will enable future guidelines to clearly recommend if any surgery can be offered as a treatment in people unable to tolerate or adhere to CPAP, which surgery and which people will benefit. It would also mean that BMI cut offs for surgery to be effective would be clearer.
Relevance to the NHS	A clear recommendation will offer clinicians clearer guidance if surgery can be offered in people with OSAHS who are unable to tolerate or adhere to CPAP and in which selected patients.

National priorities	No
Current evidence base	The current evidence is reviewed in Evidence report J of the full guideline. Eleven studies were included in the review- but only two studies included people who were unable to tolerate or adhere to CPAP. The committee agreed that there was insufficient evidence to make a recommendation for this population.
Equality	The recommendation is unlikely to impact on equality issues.
Study design	Randomised controlled trial with economic analysis.
Feasibility	The trial is feasible and should be straightforward to carry out. The time scale will need to be at least 6 months to ensure adequate follow-up so that differences in interventions can be seen between the groups, but ideally more than 12 months as there is no longer term post operative data available in current randomised controlled trials.
Other comments	-
Importance	High: the research is essential to inform future updates of key recommendations in the guideline.