Literature search queries

Upper airway surgery

Primary:

(sleep apnea syndromes [Majr] or "sleep disordered breathing" or (sleep [Majr] or apnea [Majr])) and (uvulo* or palato* or pharyngo* or tracheo* or septo* or genio* or gloss* or naso* or epiglot* or tonsil* or maxillo* or mandibulo* or turbin* or lingual* or hyo* or *surg or *mandibular or maxillary or mandibular or nasal or turbinate or septum or "nasal valve" or polyp or ethmoid or sphenoid or alar or lingual or tongue or tonsil or palate or pharynx) and (surg* or orthognathic or ESP or multilevel or septoplasty or "turbinate reduction" or "nasal obstruction" or "nasal polypectomy" or sinus or rhinoplasty) Limits: English language, human subjects, adults

Secondary:

("obstructive sleep apnea") and ("hypoglossal nerve stimulation" or "airway stimulation" or "upper airway stimulation") Limits: English language, human subjects, adults

Bariatric Surgery

(sleep apnea syndromes[Majr] OR "sleep disordered breathing" OR sleep[Majr] OR apnea[Majr]) and ("bariatric surgery" OR "metabolic surgery" OR "stomach stapling" OR "gastric bypass" OR "sleeve gastrectomy" OR lagb OR lrygb OR "gastric banding" OR "vertical banded gastroplasty" OR "duodenal switch" OR "biliopancreatic diversion" OR "weight loss surgery" OR LSG OR "obesity surgery" OR Roux*) Limits: English language, human subjects, adults

Inclusion/exclusion criteria

Inclusion criteria

Publication date (i.e., 1970 – present), appropriate control (no surgery or best medical care), outcome data present, surgical procedures including palatal modifications, base of tongue reduction, skeletal modifications, tracheostomy, hypoglossal nerve stimulation (HNS), nasal surgeries, relevant outcomes reported, RCTs with ≥30 participants, and observational studies with ≥10 participants, patients >18 years of age, study length ≥3 months after surgery (except for HNS) but data reported within 1 year for PICOs 1, 3, 4 and data present in a format suitable for meta-analysis

Exclusion criteria

Publication type (i.e., book and book chapters, conference abstracts, dissertations, editorials, letters to the editor, method papers, and guidelines), animal studies, participants <18 years of age, sample size <10 participants, not related to sleep apnea or sleep disordered breathing, study does not address any PICO question, laser-assisted uvulopalatopharyngoplasty procedure.

Box 1. Abbreviations

AHI – apnea-hypopnea index
BMI – body mass index
BP – blood pressure
CPAP – continuous positive airway pressure
DBP – diastolic blood pressure
ESS – Epworth Sleepiness Scale
FOSQ – Functional Outcomes of Sleep Questionnaire
ODI – oxygen desaturation index
PAP -positive airway pressure
PICO – patient, intervention, comparator, outcome
PSQI – Pittsburgh Sleep Quality Index
QOL – quality of life
RDI – respiratory disturbance index
SBP – systolic blood pressure
OR – operating room
LSAT – lowest oxygen saturation
SAQLI – sleep apnea quality of life index
SBP – systolic blood pressure
SF-36 PCS – Short Form 36 Physical Component Score
SF-36 MCS – Short Form 36 Mental Component Score
SF-36 VS – Short Form 36 Vitality Score

Surgical treatment of OSA in adults who were intolerant or unaccepting of PAP or other conservative treatment prior to surgery (PICO 1)

Figure S1. OR-based surgery vs. control (ESS) [RCTs]

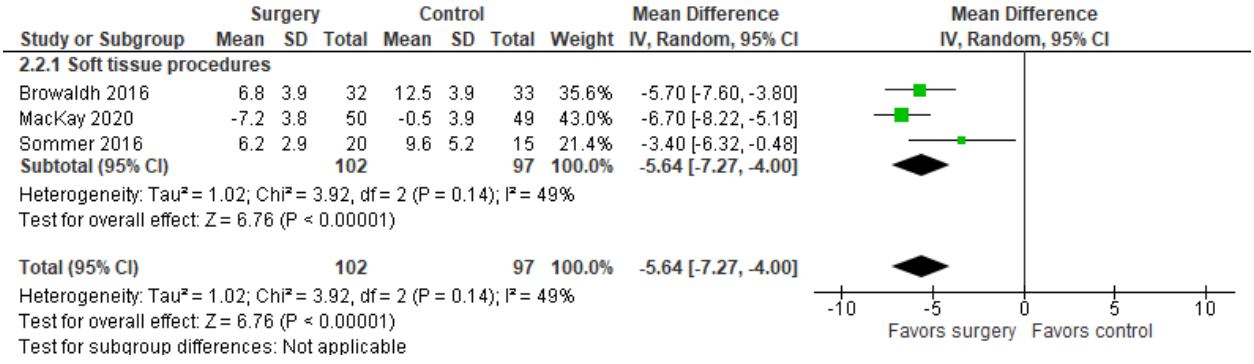
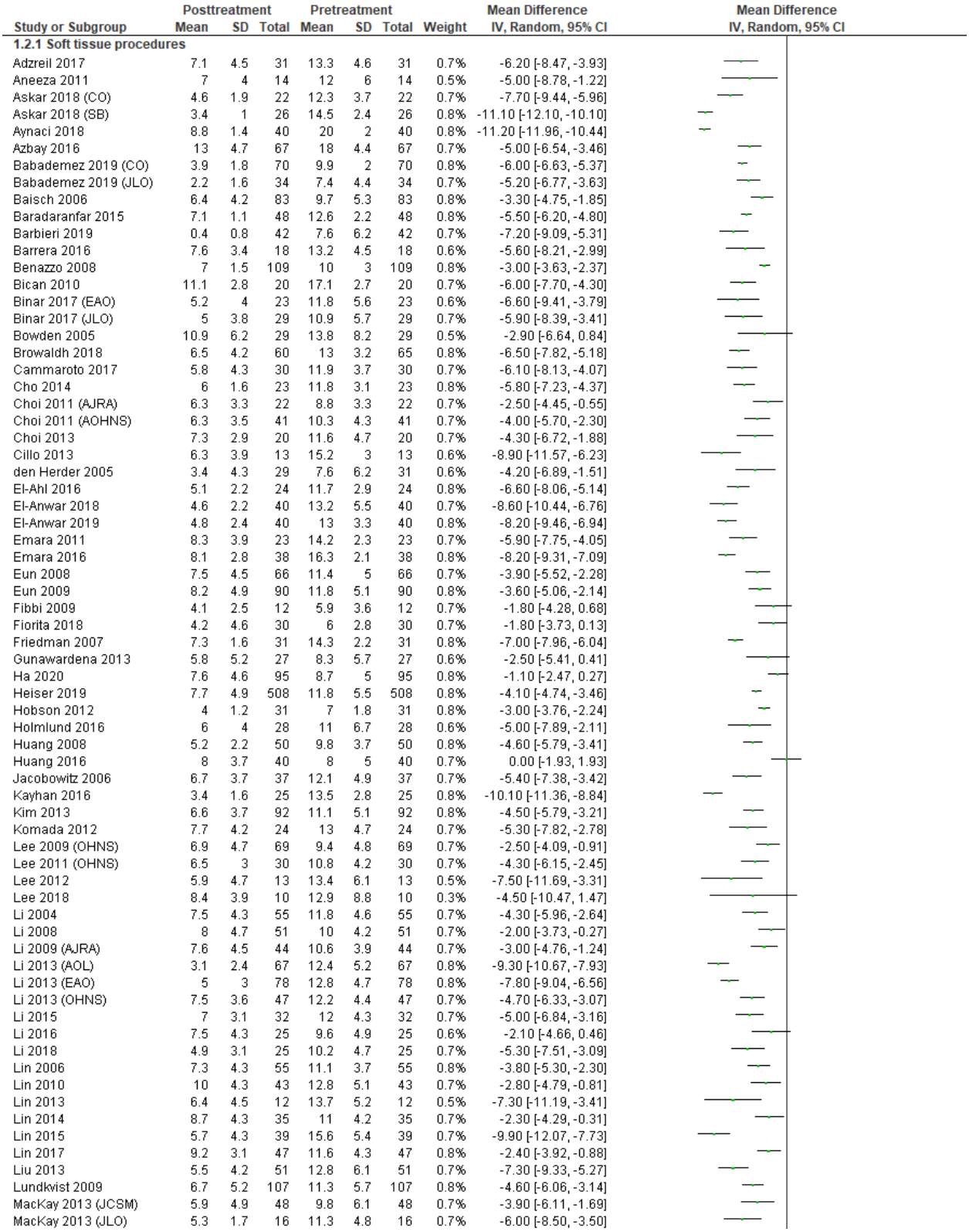


Figure S2. OR-based surgery pretreatment vs. posttreatment (ESS)



Montevecchi 2018	6.1	3.6	111	10.2	4.5	111	0.8%	-4.10 [-5.17, -3.03]
Mora 2012	4.7	2.2	21	14	3.7	21	0.7%	-9.30 [-11.14, -7.46]
Moxness 2014	8.9	3.8	59	10.7	3.7	59	0.8%	-1.80 [-3.15, -0.45]
Mure 2019	8.1	4.9	11	13	3.4	11	0.5%	-4.90 [-8.42, -1.38]
Nakata 2006	4.8	3.7	30	12.1	4.3	30	0.7%	-7.30 [-9.33, -5.27]
Neruntarat 2003	8.2	1.7	31	14.9	2.3	31	0.8%	-6.70 [-7.71, -5.69]
Neruntarat 2011	6.4	2.7	83	16.4	2.8	83	0.8%	-10.00 [-10.84, -9.16]
Omur 2005	5.4	4.3	22	13.9	2.2	22	0.7%	-8.50 [-10.52, -6.48]
Pang 2016	2.9	2.1	73	11.5	2.2	73	0.8%	-8.60 [-9.30, -7.90]
Pang 2017	4.9	1.4	112	11.5	4.1	112	0.8%	-6.60 [-7.40, -5.80]
Park 2018	7.9	4.4	29	10.1	4.9	29	0.7%	-2.20 [-4.60, 0.20]
Peng 2016	5	5	52	11.3	5.4	52	0.7%	-6.30 [-8.30, -4.30]
Peng 2019	6.8	3.4	48	11.3	5	48	0.7%	-4.50 [-6.21, -2.79]
Plaza 2019	4.6	6.6	75	11.5	4.7	75	0.7%	-6.90 [-8.73, -5.07]
Rotenberg 2014	4.1	2.5	126	14.3	3.5	126	0.8%	-10.20 [-10.95, -9.45]
Roustan 2018	3.1	2.5	20	16.5	4.3	20	0.7%	-13.40 [-15.58, -11.22]
Salapatas 2016	6.8	2.9	582	12.1	4.8	582	0.8%	-5.30 [-5.76, -4.84]
Sezen 2011	6.6	3.1	20	13.9	2.5	20	0.7%	-7.30 [-9.05, -5.55]
Shah 2018	7	3.4	16	11	4.9	16	0.6%	-4.00 [-6.92, -1.08]
Shin 2013	9.1	4.1	30	12.3	4.9	30	0.7%	-3.20 [-5.49, -0.91]
Shuaib 2015	7.1	7.1	26	11.1	7.8	26	0.5%	-4.00 [-8.05, 0.05]
Sorrenti 2003	6.6	1.4	15	11.3	2	15	0.8%	-4.70 [-5.94, -3.46]
Sun 2008	8.9	4.9	31	17.1	4.1	31	0.7%	-8.20 [-10.45, -5.95]
Sundman 2021	6.5	3.7	60	13	3.7	65	0.8%	-6.50 [-7.80, -5.20]
Toh 2014	5.6	4.4	20	13	2.8	20	0.7%	-7.40 [-9.69, -5.11]
Turhan 2020	4.1	2.7	64	10.5	5.2	64	0.8%	-6.40 [-7.84, -4.96]
Verse 2002	7.7	5	26	11.9	4.7	26	0.6%	-4.20 [-6.84, -1.56]
Vicini 2010 (ORL)	6.9	2.8	10	12.4	3.5	10	0.6%	-5.50 [-8.28, -2.72]
Vicini 2014 (ORL)	5.7	3.5	243	12.3	5.2	243	0.8%	-6.60 [-7.39, -5.81]
Vicini 2015	4.3	2	10	11.6	4.9	10	0.6%	-7.30 [-10.58, -4.02]
Vilaseca 2002	7.9	5	20	12	5.8	20	0.6%	-4.10 [-7.46, -0.74]
Wang 2013	5.5	3.6	36	12.2	5.8	36	0.7%	-6.70 [-8.93, -4.47]
Weaver 2011	6.9	4.2	34	12.9	5.5	34	0.7%	-6.00 [-8.33, -3.67]
Yaremchuk 2011	5.5	3.5	40	11.1	4.8	40	0.7%	-5.60 [-7.44, -3.76]
Yi 2009	4.8	3.4	34	11.9	6.2	34	0.7%	-7.10 [-9.48, -4.72]
Yi 2011	6.8	5.2	26	13.5	5.9	26	0.6%	-6.70 [-9.72, -3.68]
Yin 2007	9.6	5.6	18	16.1	4.3	18	0.6%	-6.50 [-9.76, -3.24]
Yuksel 2018	3.4	3.2	63	8.2	4.9	63	0.8%	-4.80 [-6.25, -3.35]
Subtotal (95% CI)			**5620**			**5632**	**75.4%**	**-5.63 [-6.13, -5.13]**

Heterogeneity: Tau² = 5.84; Chi² = 1353.02, df = 106 (P < 0.00001); I² = 92%
Test for overall effect: Z = 22.20 (P < 0.00001)

1.2.2 Skeletal procedures

Brevi 2011 (JOMS)	1.4	1.9	33	15.5	6.5	33	0.7%	-14.10 [-16.41, -11.79]
Giarda 2013	2.6	2	16	12.9	1.7	16	0.8%	-10.30 [-11.59, -9.01]
Islam 2015 (BJOMS v1)	5.4	3.4	50	14.7	7.1	50	0.7%	-9.30 [-11.48, -7.12]
Islam 2015 (BJOMS v3)	5.4	3.8	39	13.7	4.3	39	0.7%	-8.30 [-10.10, -6.50]
Liu 2012	2.9	2	11	20.5	3.7	11	0.7%	-17.60 [-20.09, -15.11]
Liu 2015	7.1	2.6	16	19.5	2.9	16	0.7%	-12.40 [-14.31, -10.49]
Liu 2016	5.7	2.7	20	17	4.8	20	0.7%	-11.30 [-13.71, -8.89]
Makovey 2017	7.3	4.9	11	12.6	6.8	11	0.4%	-5.30 [-10.25, -0.35]
Shin 2013	7.3	3.2	10	10.1	3.3	10	0.6%	-2.80 [-5.65, 0.05]
Tsui 2020	6.2	3.7	18	16.4	2	18	0.7%	-10.20 [-12.14, -8.26]
Varghese 2012	8.8	3.3	14	13.6	5.4	14	0.6%	-4.80 [-8.12, -1.48]
Subtotal (95% CI)			**238**			**238**	**7.2%**	**-9.88 [-11.94, -7.81]**

Heterogeneity: Tau² = 10.47; Chi² = 95.46, df = 10 (P < 0.00001); I² = 90%
Test for overall effect: Z = 9.39 (P < 0.00001)

1.2.3 Upper airway stimulation

Boon 2018	7.5	4.7	239	11.9	5.5	261	0.8%	-4.40 [-5.29, -3.51]
Catella 2019	10.9	6.1	10	15.9	3.5	10	0.5%	-5.00 [-9.36, -0.64]
Choi 2015	7.3	3.6	62	10.8	3.8	62	0.8%	-3.50 [-4.80, -2.20]
Eastwood 2011	8.1	4.4	19	12.1	4.7	21	0.6%	-4.00 [-6.82, -1.18]
Eastwood 2020	8	5.4	22	11	5.3	22	0.6%	-3.00 [-6.16, 0.16]
Friedman 2016	8.3	4.4	43	12	4.8	43	0.7%	-3.70 [-5.65, -1.75]
Hasselbacher 2018	6.5	4.5	56	12.8	5.3	56	0.7%	-6.30 [-8.12, -4.48]
Heiser 2017	5.9	5.2	31	12.6	5.6	31	0.6%	-6.70 [-9.39, -4.01]
Huntley 2018 (AORL)	5.9	4.1	109	11	4.3	108	0.8%	-5.10 [-6.22, -3.98]
Huntley 2018 (OHNS)	6.4	4.2	164	11.5	5	164	0.8%	-5.10 [-6.10, -4.10]
Huntley 2019 (ENT)	5.2	4	27	10.3	5.7	27	0.6%	-5.10 [-7.73, -2.47]
Kent 2016	6	4.4	20	10.3	5.2	20	0.6%	-4.30 [-7.29, -1.31]
Lai 2018	8.7	4.3	121	9.8	4.5	121	0.8%	-1.10 [-2.21, 0.01]
Mahmoud 2018	7	6.8	47	10	8.2	47	0.6%	-3.00 [-6.05, 0.05]
Parikh 2018	4.7	1.5	14	15.1	5.5	14	0.6%	-10.40 [-13.39, -7.41]
Philip 2018	10	6.1	10	15.9	3.5	10	0.5%	-5.90 [-10.26, -1.54]
Sarber 2020	6.5	3.9	31	10.5	1.9	31	0.7%	-4.00 [-5.53, -2.47]

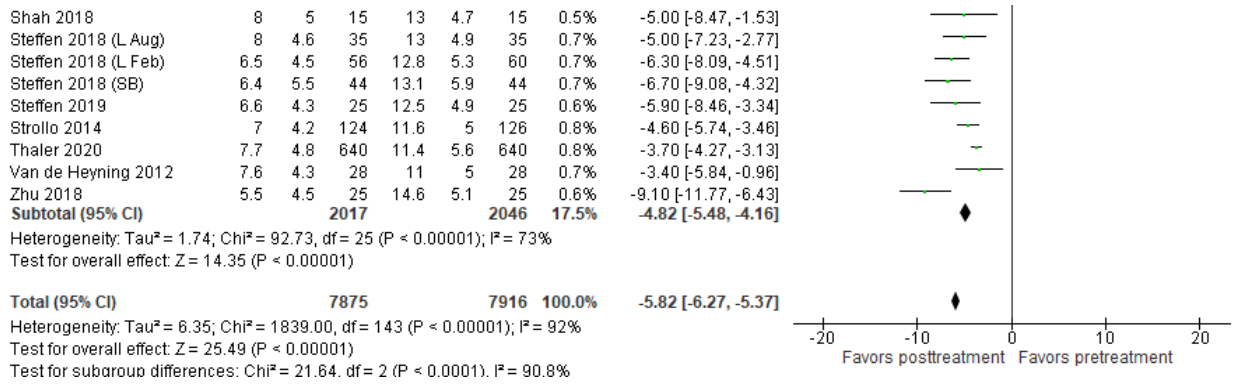


Figure S3. OR-based surgery vs. control (QOL, FOSQ) [RCT]

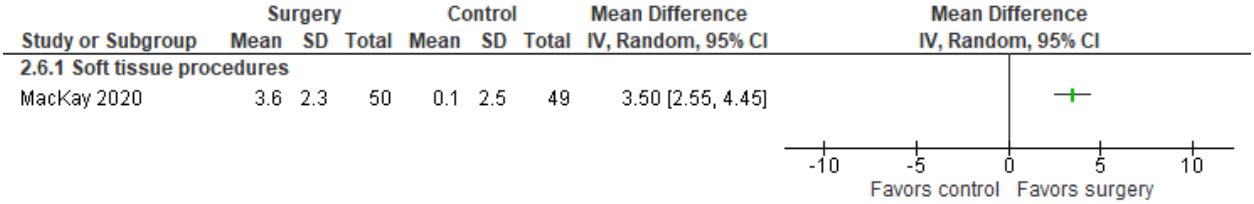


Figure S4. OR-based surgery pretreatment vs. posttreatment (QOL, FOSQ)

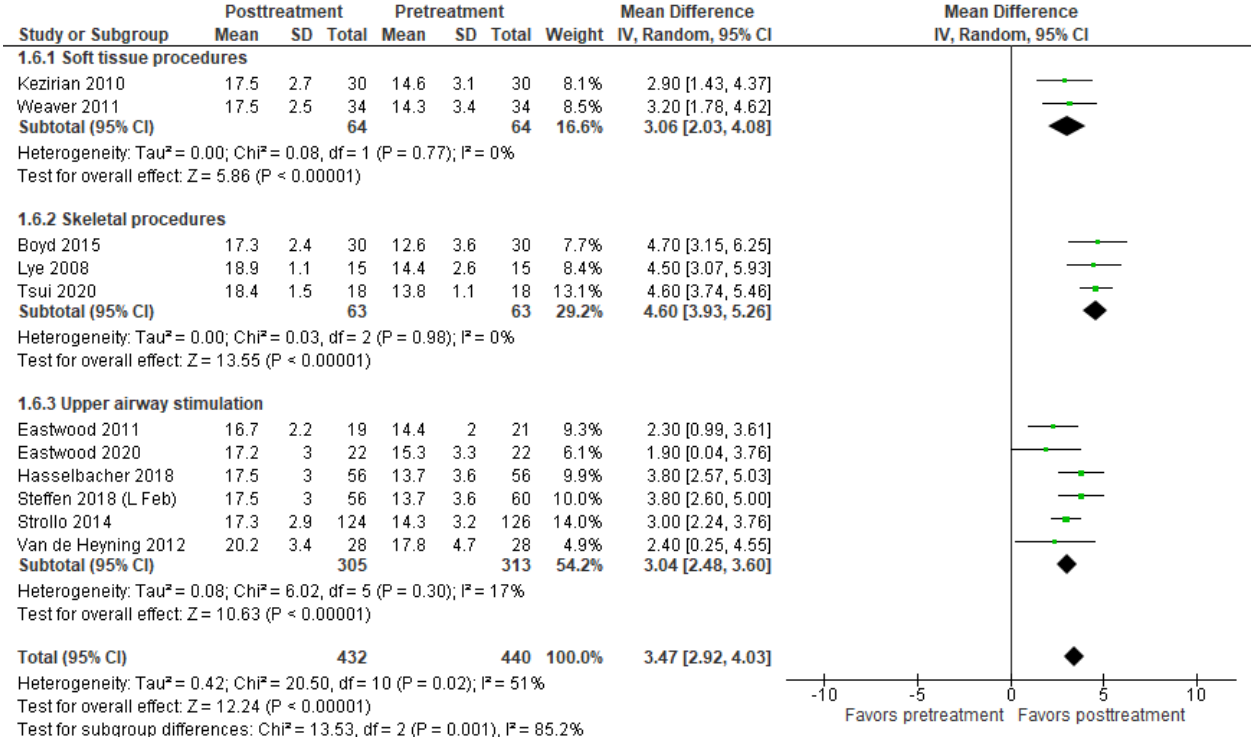


Figure S5. OR-based surgery pretreatment vs. posttreatment (QOL, SAQLI)

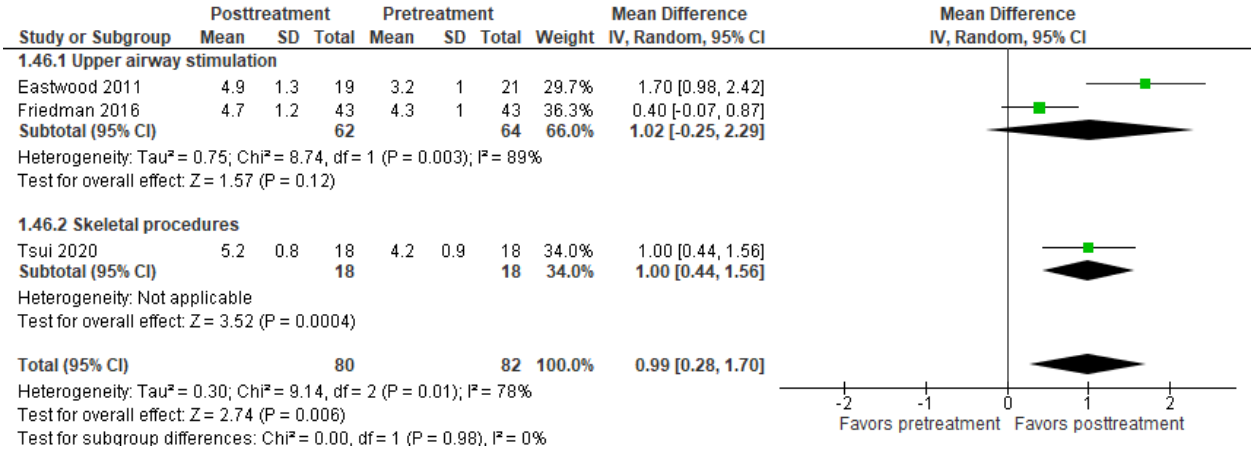


Figure S6. OR-based surgery vs. control (QOL, SF-36 PCS) [RCT]

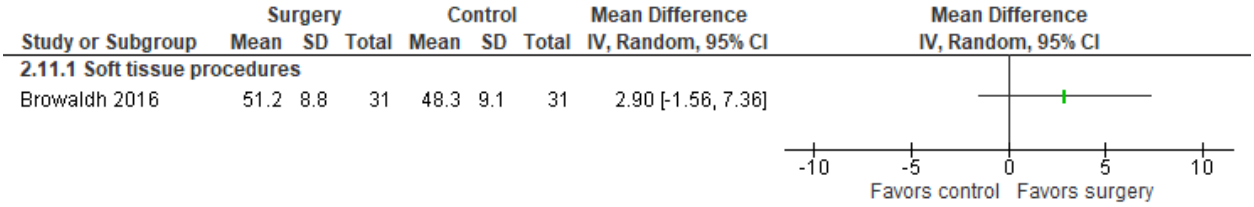


Figure S7. OR-based surgery pretreatment vs. posttreatment (QOL, SF-36 PCS)

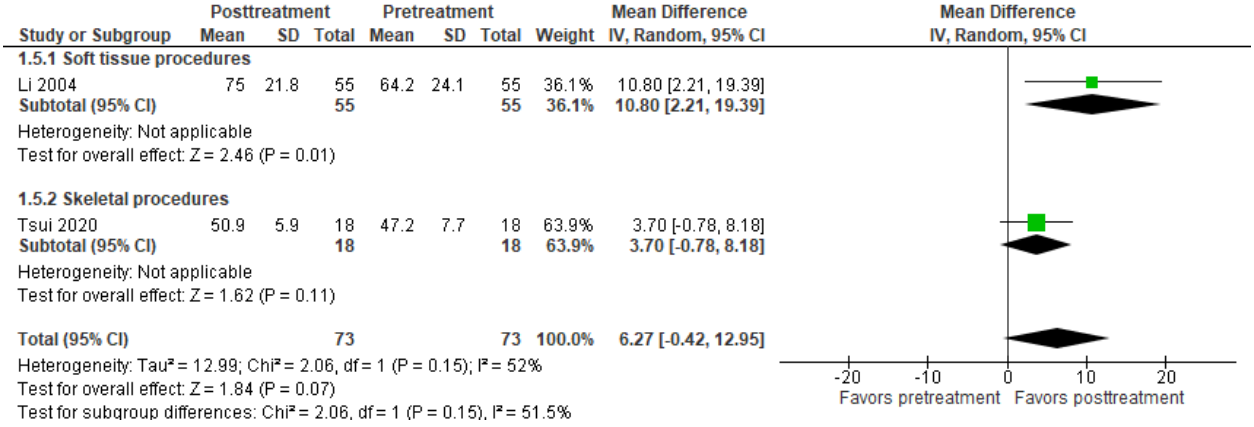


Figure S8. OR-based surgery vs. control (QOL, SF-36 MCS) [RCT]

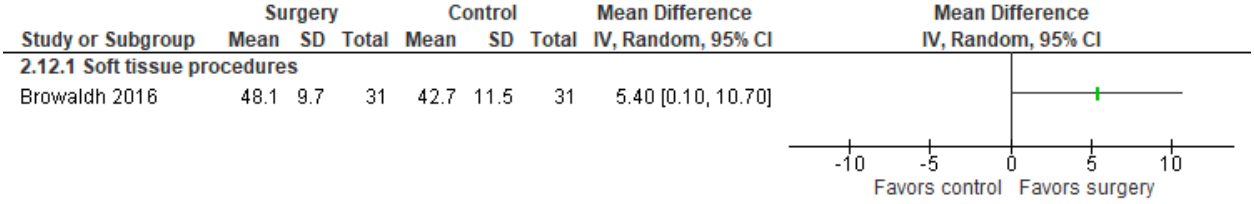


Figure S9. OR-based surgery pretreatment vs. posttreatment (QOL, SF-36 MCS)

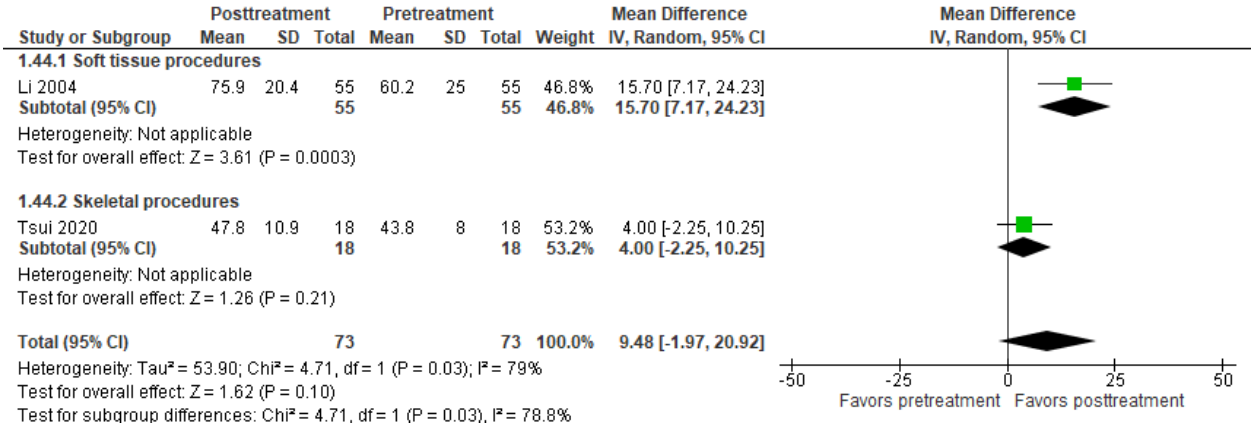


Figure S10. OR-based surgery vs. control (QOL, SF-36 VS) [RCT]

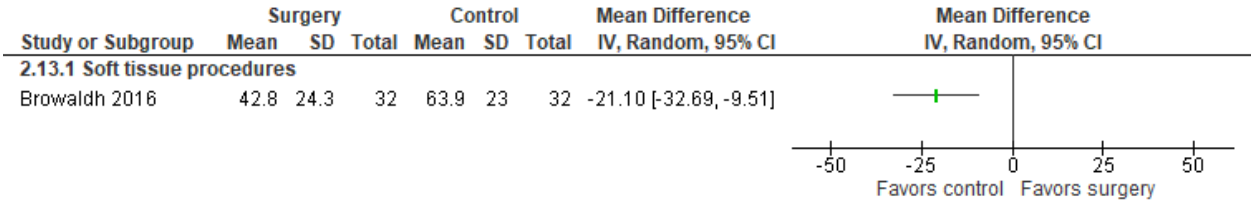


Figure S11. OR-based surgery pretreatment vs. posttreatment (QOL, SF-36 VS)

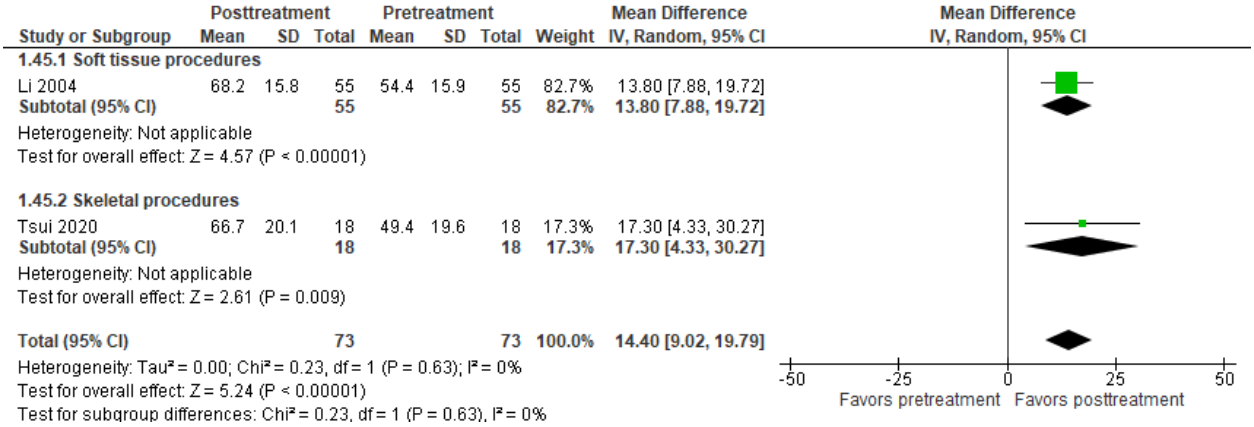


Figure S12. OR-based surgery pretreatment vs. posttreatment (QOL, PSQI)

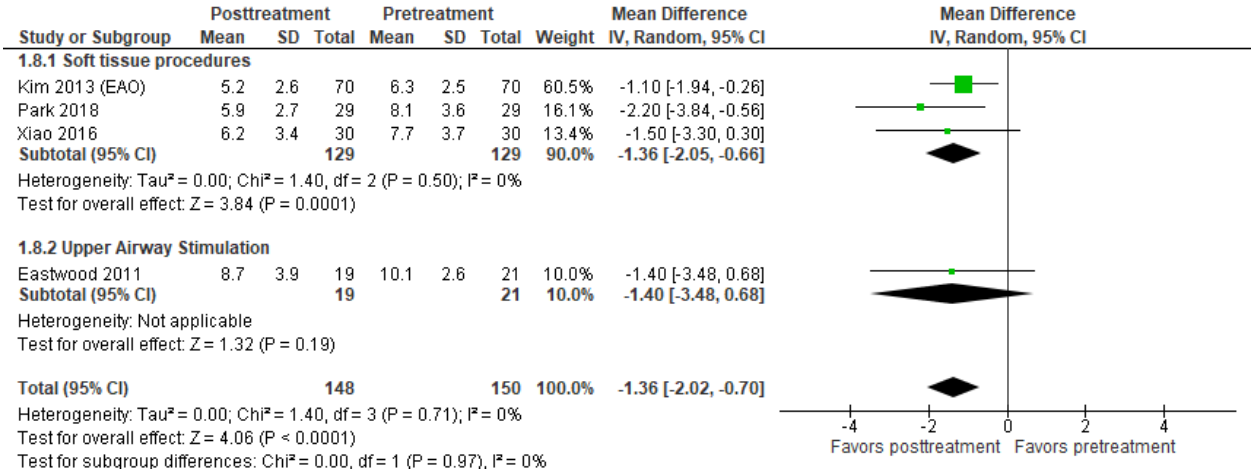


Figure S13. OR-based surgery vs. control (Snoring, VAS) [RCT]

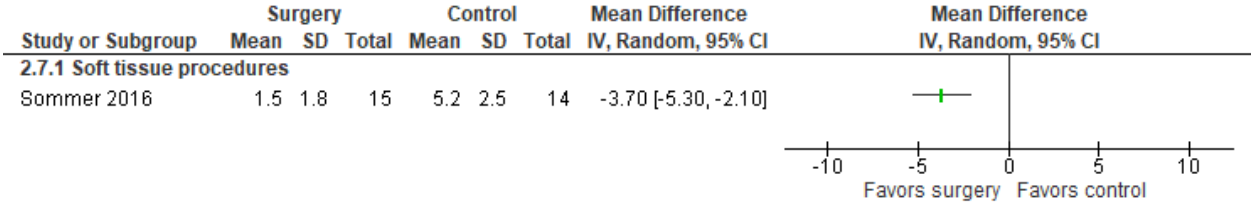


Figure S14. OR-based surgery pretreatment vs. posttreatment (Snoring, VAS)

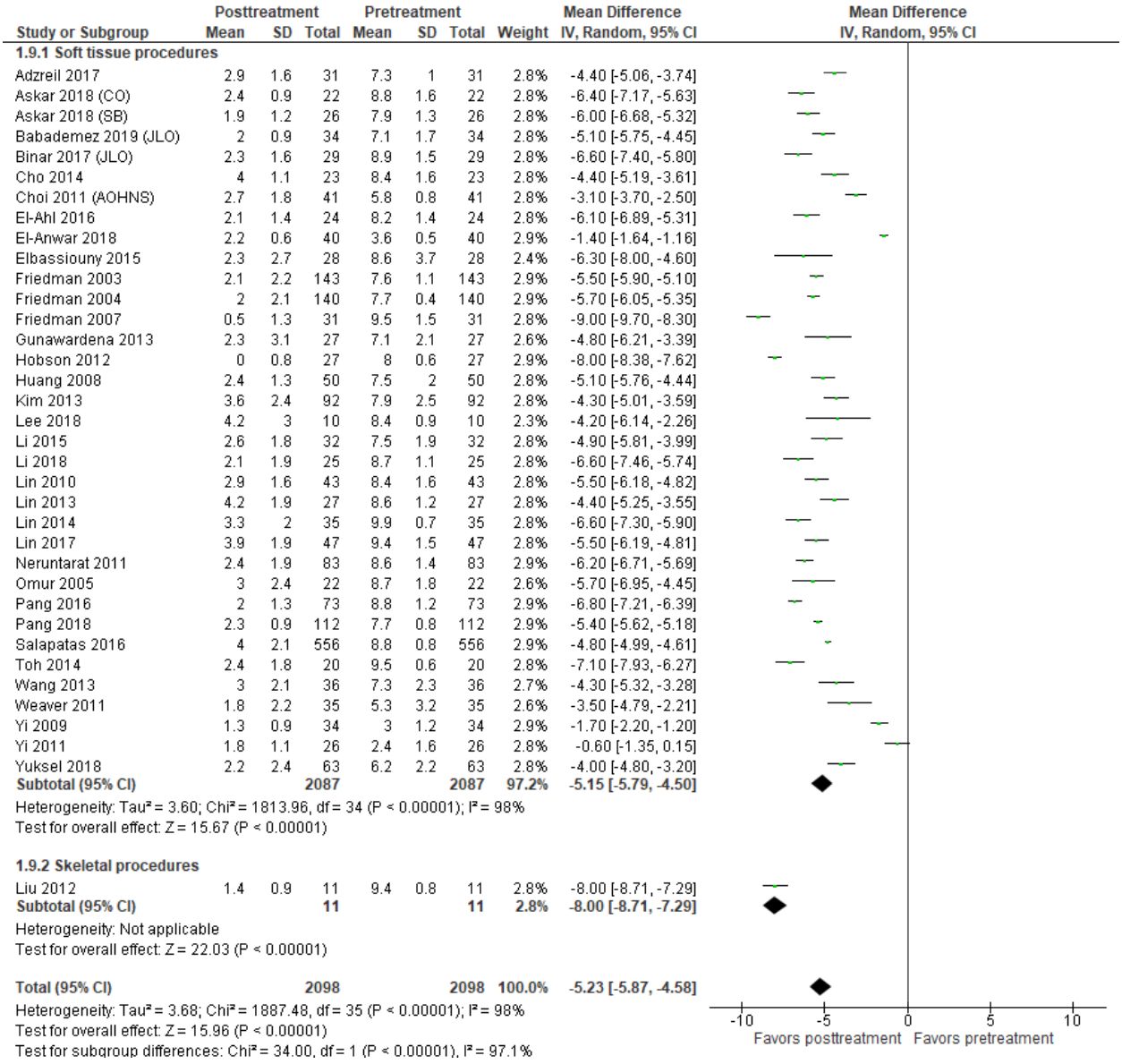


Figure S15. OR-based surgery vs. control (Mean SBP, mmHg) [RCT]

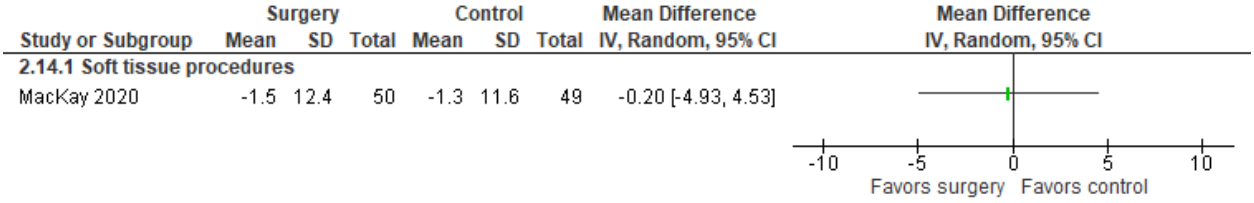


Figure S16. OR-based surgery pretreatment vs. posttreatment (Mean SBP, mmHg)

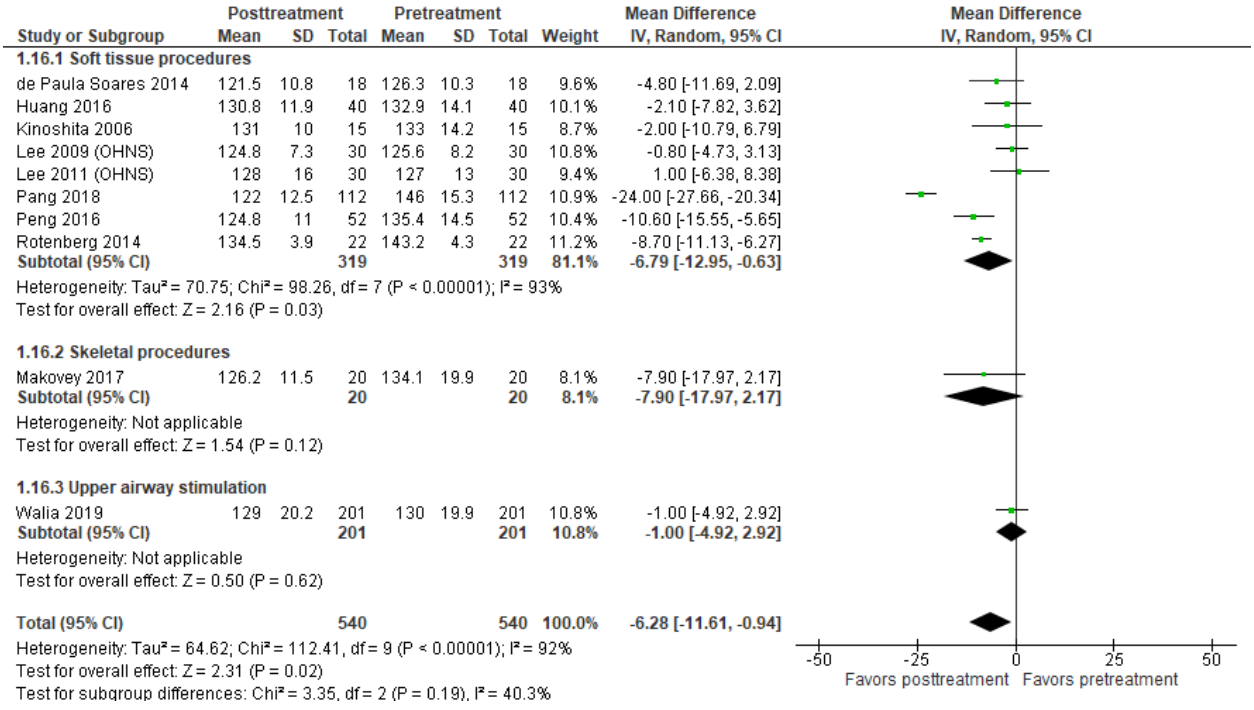


Figure S17. OR-based surgery vs. control (Mean DBP, mmHg) [RCT]

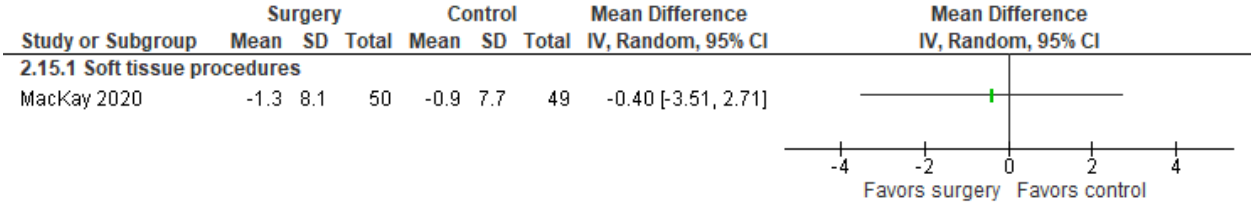


Figure S18. OR-based surgery pretreatment vs. posttreatment (Mean DBP, mmHg)

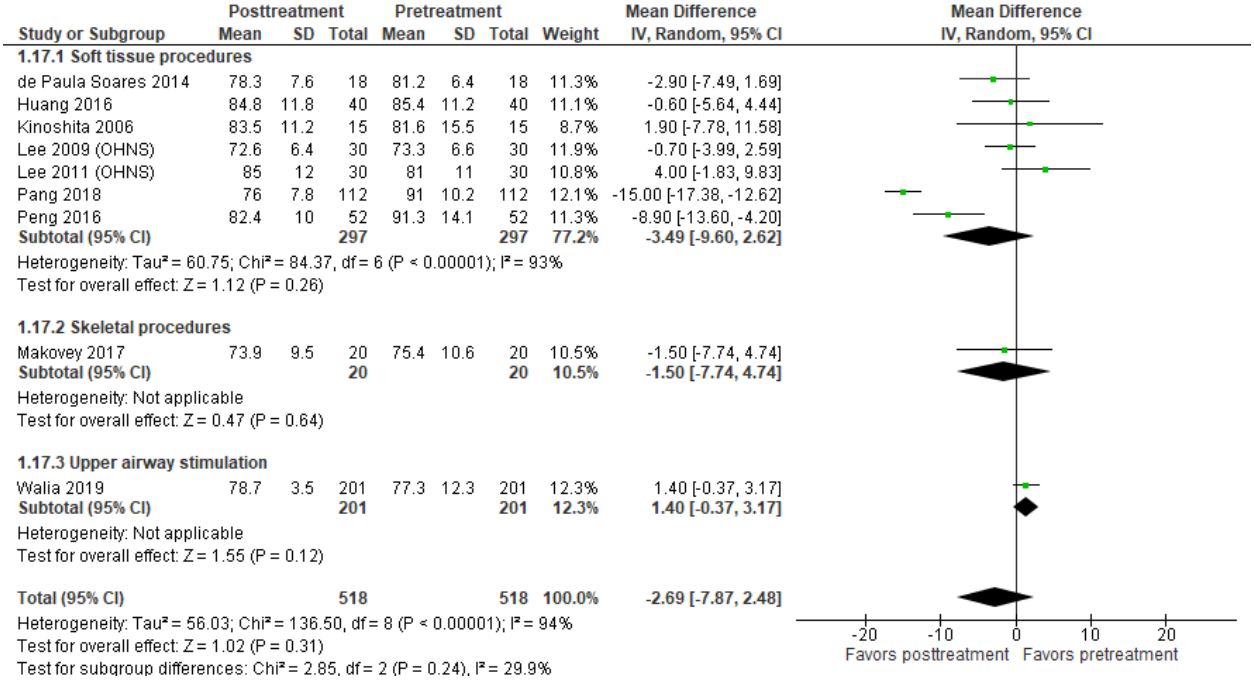


Figure S19. OR-based surgery vs. control (AHI) [RCT]

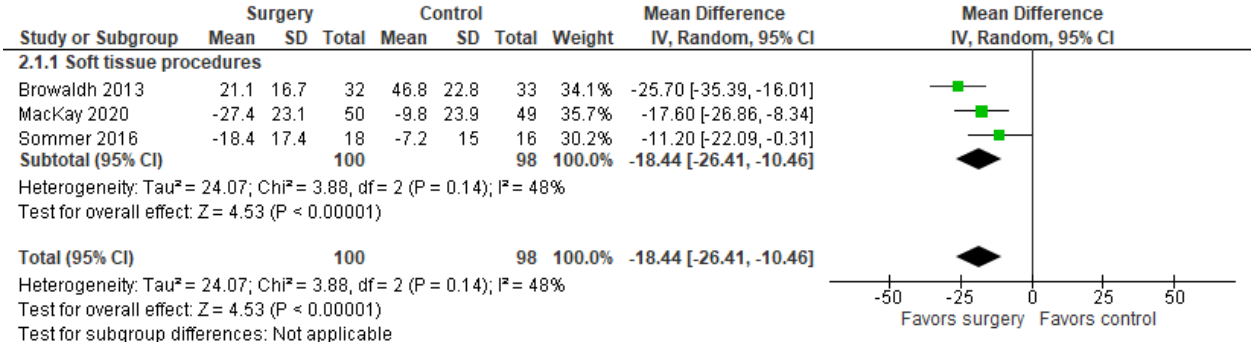
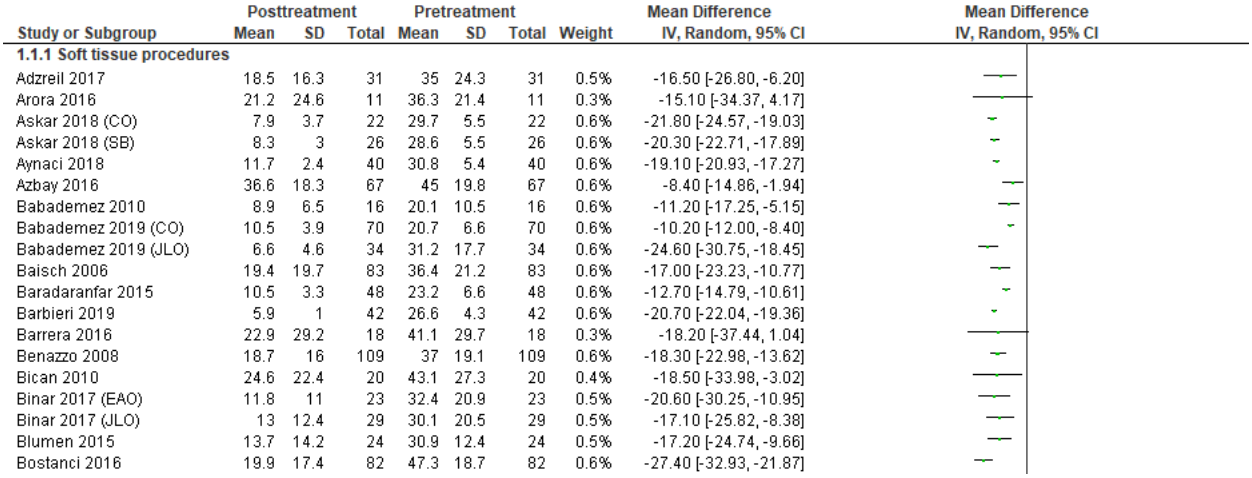


Figure S20. OR-based surgery pretreatment vs. posttreatment (AHI)



Bowden 2005	37.6	28.8	29	36.5	27.6	29	0.4%	1.10 [-13.42, 15.62]
Browaldh 2018	23.6	20.2	60	52.9	20.5	65	0.5%	-29.30 [-36.44, -22.16]
Cammaroto 2017	15.3	10.1	30	35.8	16.5	30	0.5%	-20.50 [-27.42, -13.58]
Chen 2014	27.4	12	77	46.2	15.6	77	0.6%	-18.80 [-23.20, -14.40]
Cho 2014	5.6	3.1	23	32	10.2	23	0.6%	-26.40 [-30.76, -22.04]
Choi 2011 (AJRA)	26.1	21.9	22	28.9	20.4	22	0.4%	-2.80 [-15.31, 9.71]
Choi 2011 (AOHNS)	20.9	22.1	41	45.9	23.4	41	0.5%	-25.00 [-34.85, -15.15]
Choi 2012	20.8	12.5	36	38.3	21	36	0.5%	-17.50 [-25.48, -9.52]
Choi 2013	20.1	20.4	20	37.2	22.9	20	0.4%	-17.10 [-30.54, -3.66]
Choi 2017	17.9	19.7	156	38.3	25	156	0.6%	-20.40 [-25.39, -15.41]
Cillo 2013	12.1	8.2	13	28.3	13.2	13	0.5%	-16.20 [-24.65, -7.75]
Cui 2016	25.3	16.2	12	50.9	30.8	12	0.3%	-25.60 [-45.29, -5.91]
den Herder 2005	22.2	15.2	31	32.1	10.2	31	0.6%	-9.90 [-16.34, -3.46]
de Paula Soares 2014	21	18.3	16	34.9	20.8	18	0.4%	-13.90 [-27.04, -0.76]
El-Ahl 2016	8.9	4.9	24	28.6	4.2	24	0.6%	-19.70 [-22.28, -17.12]
El-Anwar 2018	25	10.2	40	58.6	28.4	40	0.5%	-33.60 [-42.95, -24.25]
El-Anwar 2019	10.5	6.8	40	32	9.8	40	0.6%	-21.50 [-25.20, -17.80]
Elbassiouny 2015	11	2.1	28	46	4.1	28	0.6%	-35.00 [-36.71, -33.29]
Emara 2011	15.4	10.7	23	40.7	17.4	23	0.5%	-25.30 [-33.65, -16.95]
Emara 2016	16.3	10.3	38	42.1	16.3	38	0.6%	-25.80 [-31.93, -19.67]
Eun 2008	13.9	18.7	56	22.9	14.7	56	0.6%	-9.00 [-15.23, -2.77]
Eun 2009	16.6	16.5	90	20.4	8.6	90	0.6%	-3.80 [-7.64, 0.04]
Evans 2020	3.9	12.4	25	37.1	17.1	25	0.5%	-33.20 [-41.48, -24.92]
Fibbi 2009	4.7	3.2	12	14.8	3.1	12	0.6%	-10.10 [-12.62, -7.58]
Friedman 2003	27.3	23.1	277	39.8	24.3	277	0.6%	-12.50 [-16.45, -8.55]
Friedman 2004	31.3	22.6	140	41.7	24.6	140	0.6%	-10.40 [-15.93, -4.87]
Gunawardena 2013	12.5	12	27	44	22.3	27	0.5%	-31.50 [-41.05, -21.95]
Gunbey 2015	15.3	9.8	42	35.8	12.1	42	0.6%	-20.50 [-25.21, -15.79]
Ha 2020	22.2	17	95	43.9	19.8	95	0.6%	-21.70 [-26.95, -16.45]
Hamans 2008	11.8	7.6	10	22.8	11.3	10	0.5%	-11.00 [-19.44, -2.56]
Hasselbacher 2018 (SB)	20.2	13.6	15	34.7	16.2	15	0.5%	-14.50 [-25.20, -3.80]
Hester 1995	9.7	17	15	30.4	21	15	0.4%	-20.70 [-34.37, -7.03]
Hobson 2012	7.5	8.4	31	31.8	27.3	31	0.5%	-24.30 [-34.35, -14.25]
Holmlund 2016	7	9.4	28	40	32.4	28	0.4%	-33.00 [-45.50, -20.50]
Huang 2016	15	29.2	40	35.1	31.5	40	0.4%	-20.10 [-33.41, -6.79]
Jacobowitz 2006	14.9	16.8	37	46.5	24.8	37	0.5%	-31.60 [-41.25, -21.95]
Jung 2020	24.8	19.2	60	46.2	20.6	60	0.5%	-21.40 [-28.53, -14.27]
Karakoc 2018	11.4	13.4	79	21.1	15.6	79	0.6%	-9.70 [-14.23, -5.17]
Karatayli-Ozgursoy 2012	10.6	6.1	20	31	16.4	20	0.5%	-20.40 [-28.07, -12.73]
Katsantonis 1990	46.4	34.8	72	68.5	31.7	72	0.5%	-22.10 [-32.97, -11.23]
Kayhan 2016	9.4	12.4	25	28.7	17.8	25	0.5%	-19.30 [-27.80, -10.80]
Kezirian 2010	27.8	26.4	30	44.9	28.1	30	0.4%	-17.10 [-30.90, -3.30]
Khan 2009	28.3	28.9	63	62	35.4	63	0.5%	-33.70 [-44.98, -22.42]
Kim 2013	7.9	5.6	92	39.1	21.6	92	0.6%	-31.20 [-35.76, -26.64]
Kinoshita 2006	15.6	14.5	15	47.9	23.7	15	0.4%	-32.30 [-46.36, -18.24]
Komada 2012	10.7	10.7	24	50.9	22.3	24	0.5%	-40.20 [-50.10, -30.30]
Lee 2010	22.8	21.2	69	35	17.5	69	0.6%	-12.20 [-18.69, -5.71]
Lee 2011 (OHNS)	17.9	13.6	30	46.2	22.9	30	0.5%	-28.30 [-37.83, -18.77]
Lee 2012	24.1	19.6	20	55.6	26	20	0.4%	-31.50 [-45.77, -17.23]
Lee 2018	37.1	31.9	10	46.6	35	10	0.2%	-9.50 [-38.85, 19.85]
Li 2006	15.1	15.1	110	44.4	26.6	110	0.6%	-29.30 [-35.02, -23.58]
Li 2008	38.1	32.7	51	37.4	28.9	51	0.5%	0.70 [-11.28, 12.68]
Li 2009 (AJRA)	37.5	31.6	44	36.4	29.1	44	0.4%	1.10 [-11.59, 13.79]
Li 2013 (AOL)	12	7.2	67	48.4	21.7	67	0.6%	-36.40 [-41.87, -30.93]
Li 2013 (EAO)	14.9	8.8	78	40.9	19.9	78	0.6%	-26.00 [-30.83, -21.17]
Li 2013 (OHNS)	22.6	17.4	47	59.5	18.2	47	0.5%	-36.90 [-44.10, -29.70]
Li 2014	22	21.4	73	41.2	17.3	73	0.6%	-19.20 [-25.51, -12.89]
Li 2015	27.2	21.3	32	52.5	25.9	32	0.5%	-25.30 [-36.92, -13.68]
Li 2016	12.8	8.2	25	45.7	21.7	25	0.5%	-32.90 [-41.99, -23.81]
Li 2018	15.1	12.1	25	39.8	19.9	25	0.5%	-24.70 [-33.83, -15.57]
Lim 2018	19.4	18.4	97	36.1	21.5	97	0.6%	-16.70 [-22.33, -11.07]
Lin 2010	23.4	24.7	43	51.5	25.4	43	0.5%	-28.10 [-38.69, -17.51]
Lin 2013	17.6	16.2	12	43.9	41.1	12	0.2%	-26.30 [-51.30, -1.30]
Lin 2014	26.5	23.5	35	50.6	16.6	35	0.5%	-24.10 [-33.63, -14.57]
Lin 2015	21.9	23.5	39	43.9	32.3	39	0.4%	-22.00 [-34.54, -9.46]
Liu 2013	29.5	22.5	51	65.6	17.2	51	0.5%	-36.10 [-43.87, -28.33]
MacKay 2013 (JLO)	14.5	9.8	17	36.3	11.9	17	0.5%	-21.80 [-29.13, -14.47]
Meraj 2017	21.5	19.3	101	44.9	21.6	101	0.6%	-23.40 [-29.05, -17.75]
Miyazaki 1998	2.9	4.3	10	14.4	19.9	10	0.4%	-11.50 [-24.12, 1.12]
Montevecchi 2018	13.5	10.3	111	33.4	19.5	111	0.6%	-19.90 [-24.00, -15.80]
Mora 2012	9	3.1	21	31.8	13	21	0.6%	-22.80 [-28.52, -17.08]
Moxness 2014	16.6	12.9	59	18.2	13.7	59	0.6%	-1.60 [-6.40, 3.20]
Mure 2019	11.6	9.6	11	29.5	21.7	11	0.4%	-17.90 [-31.92, -3.88]
Mutlu 2017	12.8	6.6	25	18.9	8.4	25	0.6%	-6.10 [-10.29, -1.91]
Nakata 2006	30.1	24	30	69	28.4	30	0.4%	-38.90 [-52.21, -25.59]
Neruntarat 2011	13.4	5.2	83	45.6	10.3	83	0.6%	-32.20 [-34.68, -29.72]
Pang 2016	12.6	5.8	73	26.3	17.7	73	0.6%	-13.70 [-17.97, -9.43]

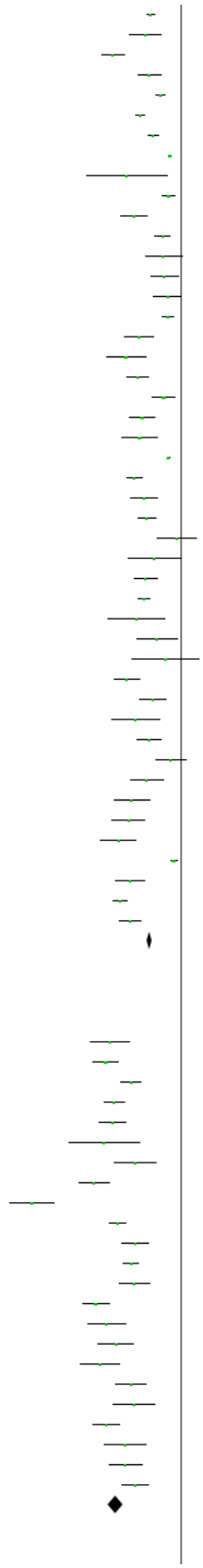
Pang 2020 (L p.551)	14.6	10.8	326	35	19.2	326	0.6%	-20.40 [-22.79, -18.01]
Park 2018	13.8	15.6	29	37.8	24.7	29	0.5%	-24.00 [-34.63, -13.37]
Peng 2016	18	15.2	52	64.2	25.5	52	0.5%	-46.20 [-54.27, -38.13]
Peng 2019	26	20.2	48	47	18.9	48	0.5%	-21.00 [-28.83, -13.17]
Plaza 2019	8.6	6.7	75	22.1	12.2	75	0.6%	-13.50 [-16.65, -10.35]
Rotenberg 2014	19.8	11.1	126	47.3	11.4	126	0.6%	-27.50 [-30.28, -24.72]
Roustan 2018	5.2	3.2	20	23.6	6.5	20	0.6%	-18.40 [-21.58, -15.22]
Salapatas 2016	12.7	7.6	601	19.8	5.9	601	0.6%	-7.10 [-7.87, -6.33]
Sanders 1990	29.5	38.8	15	66.3	38.5	15	0.2%	-36.80 [-64.46, -9.14]
Santos 2007	4.4	5.7	10	12.4	4.6	10	0.6%	-8.00 [-12.54, -3.46]
Sarber 2020	3.1	4.9	31	34.8	25.6	31	0.5%	-31.70 [-40.88, -22.52]
Sezen 2011	12.4	7.7	20	24.6	9.1	20	0.6%	-12.20 [-17.42, -6.98]
Shah 2018	28.8	25.4	20	40.3	12.4	20	0.4%	-11.50 [-23.89, 0.89]
Shin 2013	18.5	15.6	30	29.6	21.6	30	0.5%	-11.10 [-20.63, -1.57]
Shuaib 2015	16	16.1	26	24.7	18.8	26	0.5%	-8.70 [-18.21, 0.81]
Simsek 2015	9	8.4	37	17.6	9.1	37	0.6%	-8.60 [-12.59, -4.61]
Suh 2013	26.9	24.2	50	55.3	24.6	50	0.5%	-28.40 [-37.96, -18.84]
Sun 2008	28.6	29.1	31	65.9	23.8	31	0.4%	-37.30 [-50.53, -24.07]
Sundman 2021	23.6	20.2	60	52.9	20.5	65	0.5%	-29.30 [-36.44, -22.16]
Suslu 2017	19.6	14	28	31	15.2	28	0.5%	-11.40 [-19.05, -3.75]
Thaler 2016	31.4	28.6	75	57.5	23.9	75	0.5%	-26.10 [-34.54, -17.66]
Toh 2014	13.5	17.1	20	41.3	22.1	20	0.4%	-27.80 [-40.05, -15.55]
Tuncel 2012	6.2	1.7	35	14.6	2.4	35	0.6%	-8.40 [-9.37, -7.43]
Turhan 2015 (EAO p3411)	20.5	17.7	90	51.8	18.8	90	0.6%	-31.30 [-36.63, -25.97]
Turhan 2015 (EAO p995)	20	19.5	31	44.7	17	31	0.5%	-24.70 [-33.81, -15.59]
Turhan 2020	18.8	14.9	64	41.7	21.3	64	0.6%	-22.90 [-29.27, -16.53]
Verse 2002	28.9	24.7	26	31.6	25.6	26	0.4%	-2.70 [-16.37, 10.97]
Vicini 2010 (ORL)	20.6	17.3	10	38.3	23.5	10	0.3%	-17.70 [-35.79, 0.39]
Vicini 2014 (HN)	14.8	11.3	24	38.5	17	24	0.5%	-23.70 [-31.87, -15.53]
Vicini 2014 (ORL)	17.9	18.4	243	43	22.6	243	0.6%	-25.10 [-28.76, -21.44]
Vicini 2015	13.6	15.4	10	43.6	26.8	10	0.3%	-30.00 [-49.16, -10.84]
Vilaseca 2002	44.6	27	20	60.4	16.5	20	0.4%	-15.80 [-29.67, -1.93]
Walker 1989	60.1	27.6	11	70.3	26.7	11	0.3%	-10.20 [-32.89, 12.49]
Wang 2013	23.2	18.4	36	59.8	20.5	36	0.5%	-36.60 [-45.60, -27.60]
Wee 2015	18.7	14.8	27	37.7	18.6	27	0.5%	-19.00 [-27.97, -10.03]
Wetmore 1986	33.5	29.1	27	64	32.7	27	0.4%	-30.50 [-47.01, -13.99]
Woodson 2005	26.4	21.2	74	48.1	27.8	74	0.5%	-21.70 [-29.67, -13.73]
Xiao 2016	43.1	21.9	30	49.7	19.5	30	0.5%	-6.60 [-17.09, 3.89]
Yaremchuk 2011	13	17.4	29	35.9	24.9	29	0.5%	-22.90 [-33.96, -11.84]
Yi 2009	28.2	32.7	34	61.5	12.5	34	0.5%	-33.30 [-45.07, -21.53]
Yi 2011	30.1	23.1	26	65.6	17.6	26	0.5%	-35.50 [-46.66, -24.34]
Yin 2007	21.4	20.3	18	63.8	16.3	18	0.5%	-42.40 [-54.43, -30.37]
Yuksel 2018	11.6	5.6	63	16.1	7.1	63	0.6%	-4.50 [-6.73, -2.27]
Zhang 2013	20.4	17.5	36	54.7	25.2	36	0.5%	-34.30 [-44.32, -24.28]
Zhang 2014	20.8	18.4	119	62.1	18.2	119	0.6%	-41.30 [-45.95, -36.65]
Zhao 2017	20	11.9	40	54.2	20.5	40	0.5%	-34.20 [-41.55, -26.85]
Subtotal (95% CI)			**7326**			**7338**	**72.5%**	**-21.03 [-22.75, -19.32]**

Heterogeneity: Tau² = 88.24; Chi² = 2793.85, df = 141 (P < 0.00001); I² = 95%
Test for overall effect: Z = 24.00 (P < 0.00001)

1.1.2 Skeletal procedures

Bettega 2000	11.1	8.9	20	59.3	29	20	0.4%	-48.20 [-61.49, -34.91]
Blumen 2009	14.4	14.5	50	65.5	26.7	50	0.5%	-51.10 [-59.52, -42.68]
Boyd 2013	17.6	15.9	106	51.3	32.9	106	0.5%	-33.70 [-40.66, -26.74]
Brevi 2011 (JOMS)	10.5	7.4	33	55.6	19.3	33	0.5%	-45.10 [-52.15, -38.05]
Conradt 1997	5	5.8	15	51.4	16.9	15	0.5%	-46.40 [-55.44, -37.36]
Gerbino 2014	17.3	16.7	10	69.8	35.2	10	0.3%	-52.50 [-76.65, -28.35]
Giarda 2013	16.1	17.5	16	47.1	22.5	16	0.4%	-31.00 [-44.97, -17.03]
Goh 2003	11.4	7.4	11	70.7	15.9	11	0.5%	-59.30 [-69.66, -48.94]
Goodday 2016	16.1	26.2	13	117.9	9.2	13	0.4%	-101.80 [-116.89, -86.71]
Hochban 1997	2.5	3.9	38	45.2	17.1	38	0.6%	-42.70 [-48.28, -37.12]
Hsieh 2014	4.8	4.4	16	35.7	18	16	0.5%	-30.90 [-39.98, -21.82]
Islam 2015 (BJOMS v1)	7.6	6.3	50	41.3	17.8	50	0.6%	-33.70 [-38.93, -28.47]
Lin 2011	4.6	4.1	12	35.9	18	12	0.5%	-31.30 [-41.75, -20.85]
Liu 2012	9.4	7.2	11	67.4	13.3	11	0.5%	-58.00 [-66.94, -49.06]
Liu 2015	9.3	7.1	16	59.8	25.6	16	0.4%	-50.50 [-63.52, -37.48]
Liu 2016	9.5	7.4	20	53.6	26.6	20	0.5%	-44.10 [-56.20, -32.00]
Lye 2008	13.9	12.8	15	69.1	23.3	15	0.4%	-55.20 [-68.65, -41.75]
Makovey 2017	15.6	12.1	20	49.4	20	20	0.5%	-33.80 [-44.04, -23.56]
Manikandhan 2014	12.9	13.8	13	44.9	22.3	13	0.4%	-32.00 [-46.26, -17.74]
Ronchi 2013	8.1	7.8	15	58.7	16	15	0.5%	-50.60 [-59.61, -41.59]
Schendel 2014	5.2	8.3	10	42.9	21.2	10	0.4%	-37.70 [-51.81, -23.59]
Varghese 2012	7.8	10.5	24	45.4	26.4	24	0.5%	-37.60 [-48.97, -26.23]
Vigneron 2017	25.5	14.7	29	56.7	20.4	29	0.5%	-31.20 [-40.35, -22.05]
Subtotal (95% CI)			**563**			**563**	**10.8%**	**-44.55 [-49.61, -39.48]**

Heterogeneity: Tau² = 121.61; Chi² = 137.76, df = 22 (P < 0.00001); I² = 84%
Test for overall effect: Z = 17.23 (P < 0.00001)



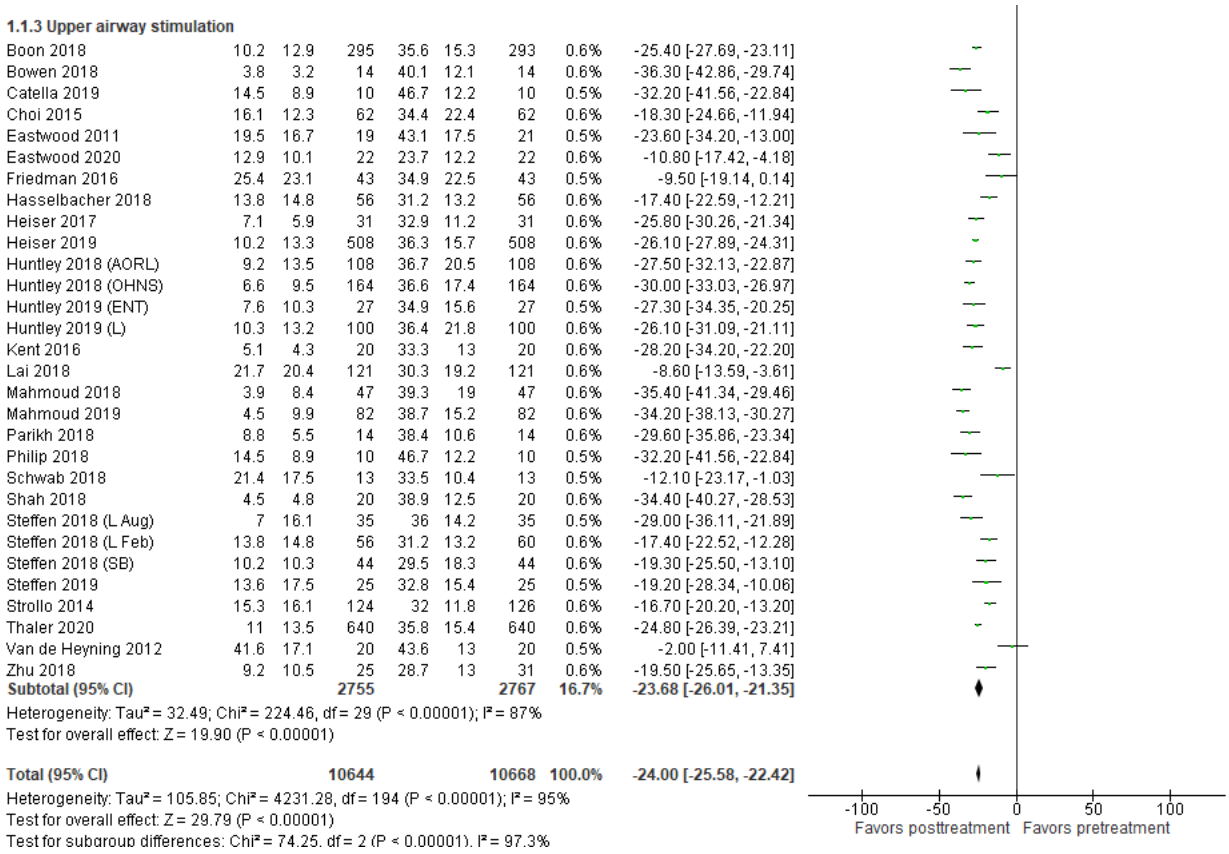


Figure S21. OR-based surgery vs. control (RDI) [RCT]

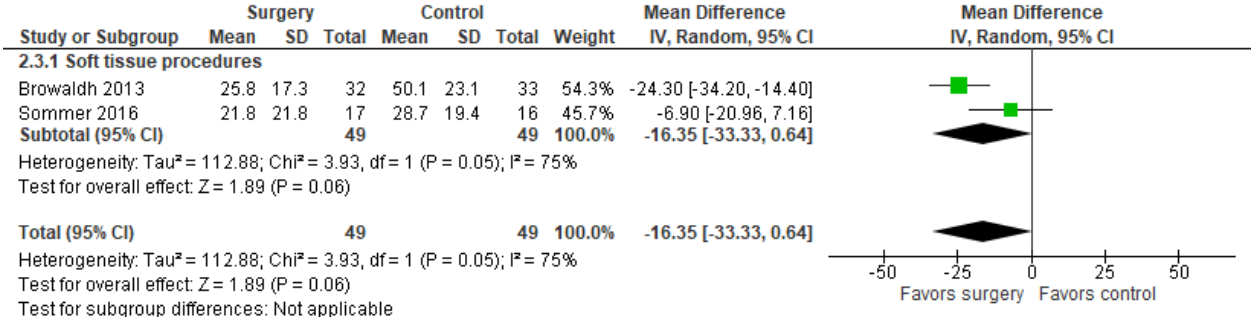


Figure S22. OR-based surgery pretreatment vs. posttreatment (RDI)

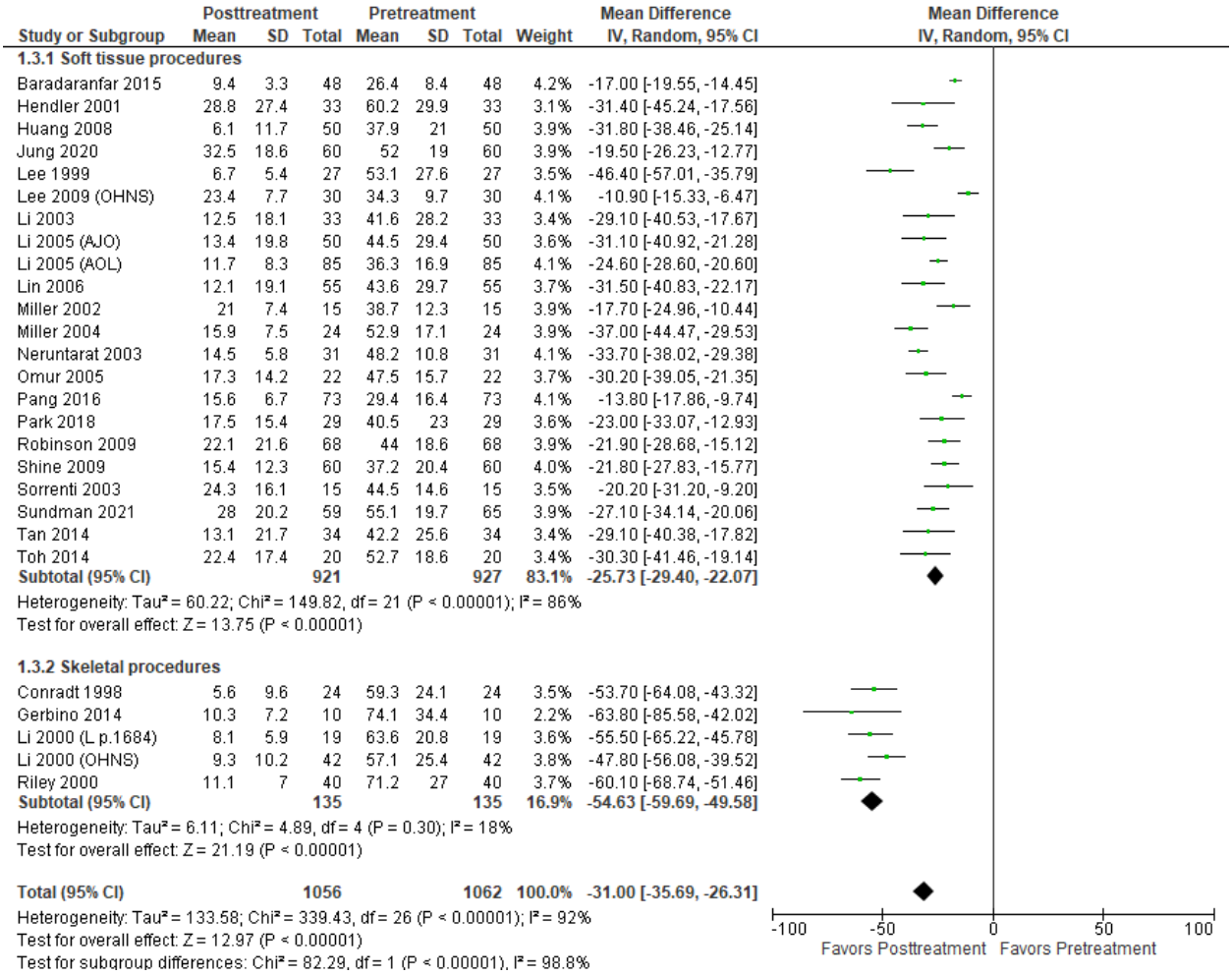


Figure S23. OR-based surgery pretreatment vs. posttreatment (persistent dysphagia, incidence)

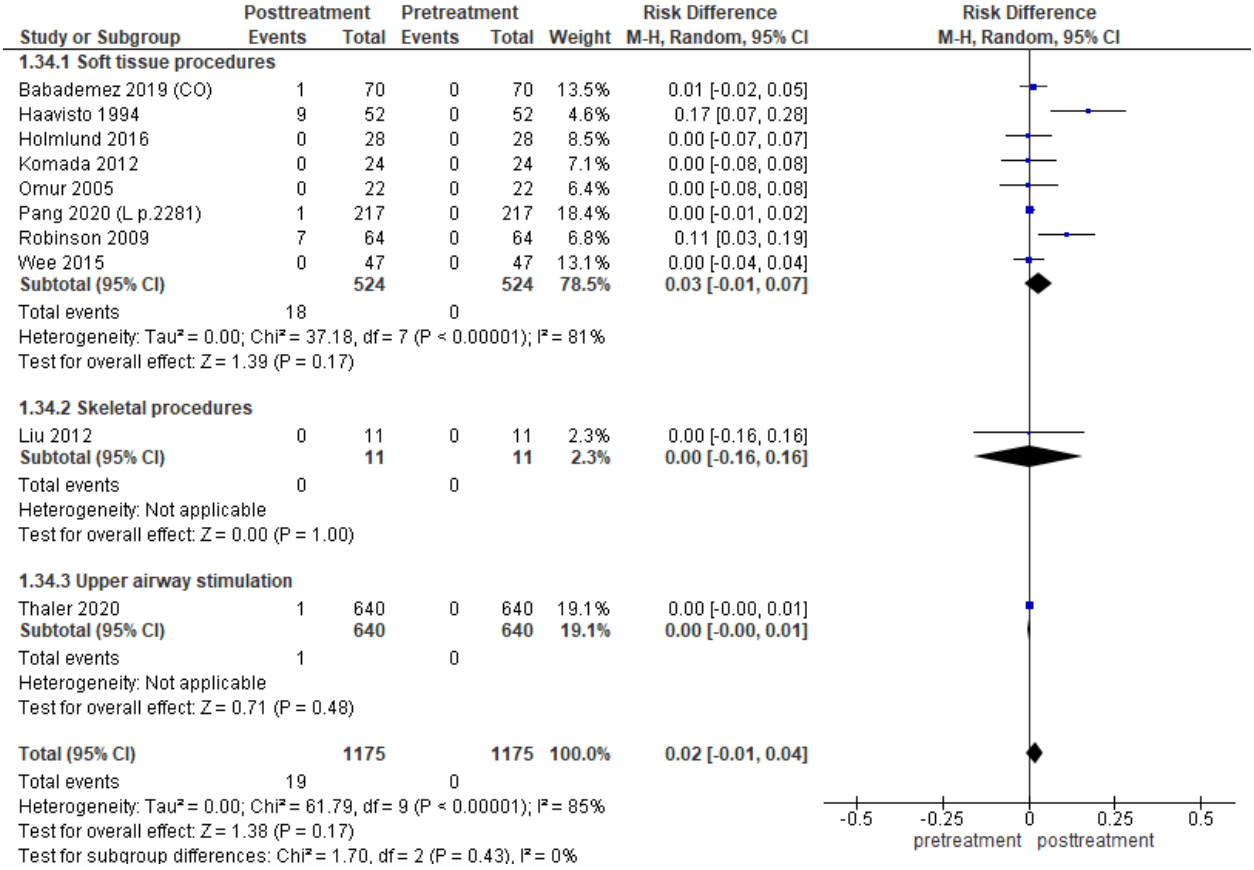


Figure S24. OR-based surgery pretreatment vs. posttreatment (MD Anderson dysphagia score)

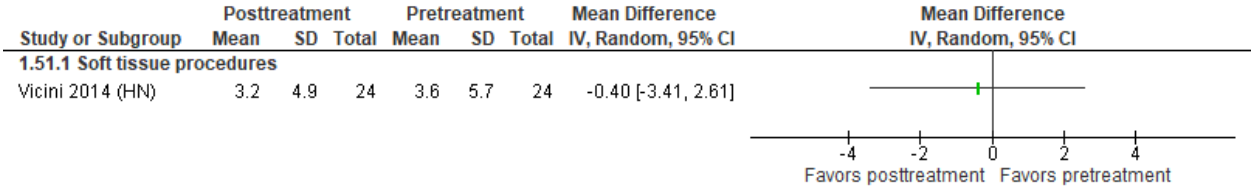


Figure S25. OR-based surgery vs. control (LSAT) [RCT]

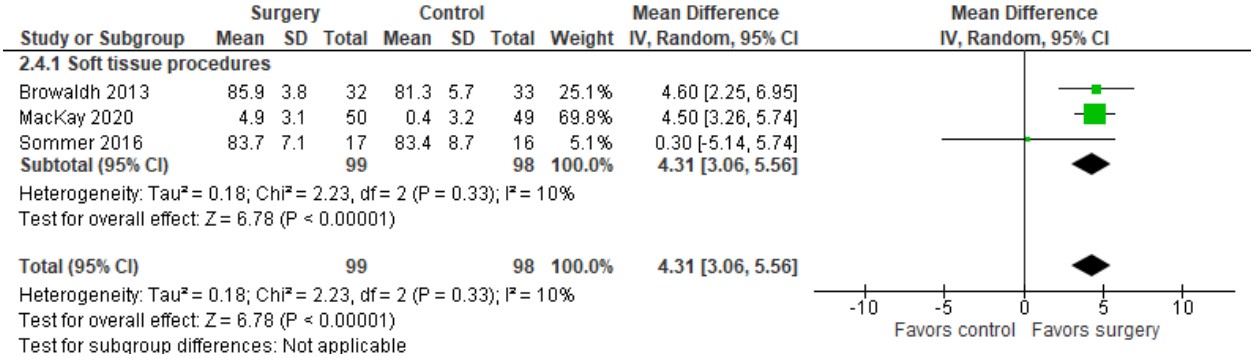
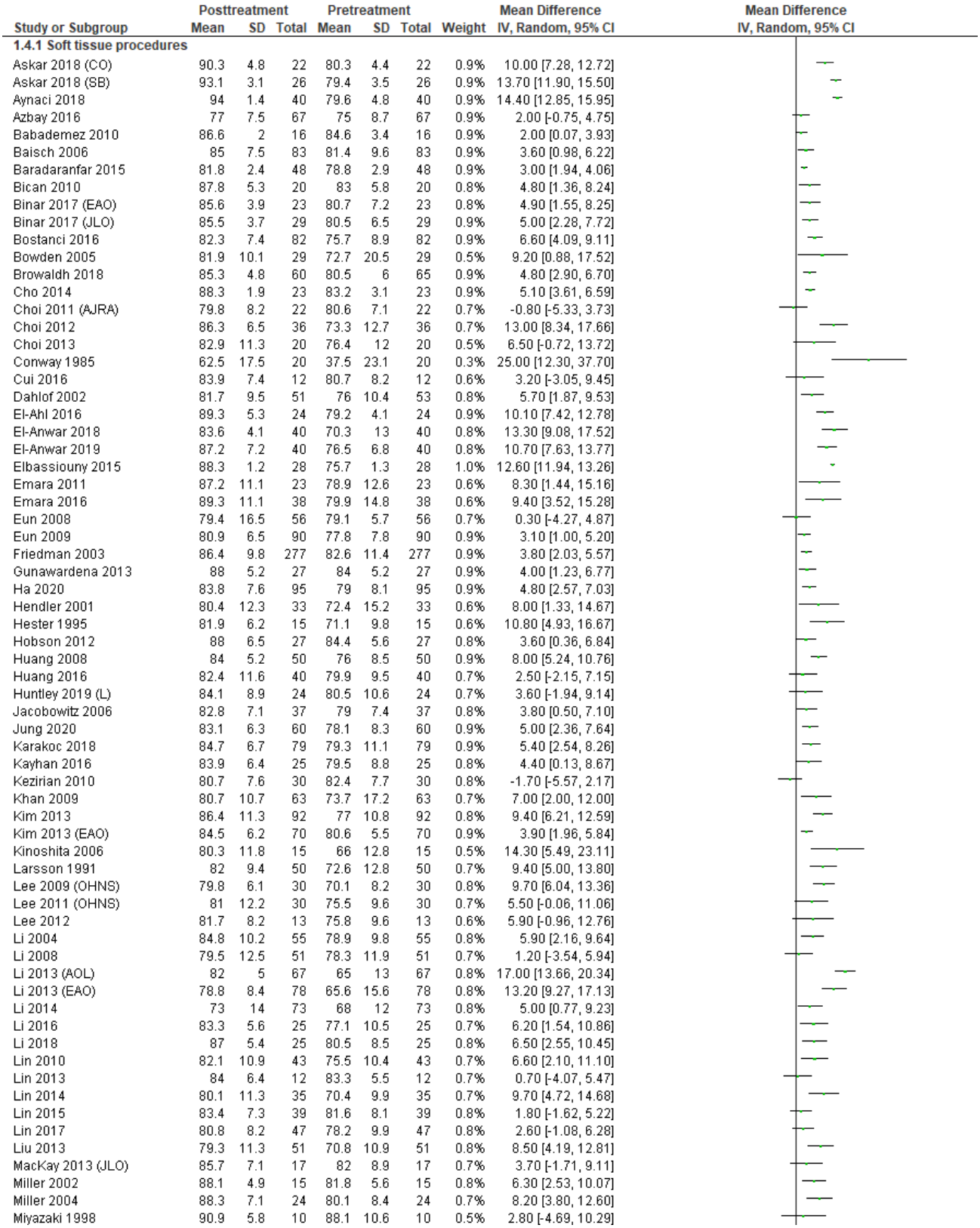


Figure S26. OR-based surgery pretreatment vs. posttreatment (LSAT)



Study or Subgroup	Posttreatment Mean	SD	Total	Pretreatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
1.4.1 Soft tissue procedures								
Askar 2018 (CO)	90.3	4.8	22	80.3	4.4	22	0.9%	10.00 [7.28, 12.72]
Askar 2018 (SB)	93.1	3.1	26	79.4	3.5	26	0.9%	13.70 [11.90, 15.50]
Aynaci 2018	94	1.4	40	79.6	4.8	40	0.9%	14.40 [12.85, 15.95]
Azbay 2016	77	7.5	67	75	8.7	67	0.9%	2.00 [-0.75, 4.75]
Babademez 2010	86.6	2	16	84.6	3.4	16	0.9%	2.00 [0.07, 3.93]
Baisch 2006	85	7.5	83	81.4	9.6	83	0.9%	3.60 [0.98, 6.22]
Baradaranfar 2015	81.8	2.4	48	78.8	2.9	48	0.9%	3.00 [1.94, 4.06]
Bican 2010	87.8	5.3	20	83	5.8	20	0.8%	4.80 [1.36, 8.24]
Binar 2017 (EAO)	85.6	3.9	23	80.7	7.2	23	0.8%	4.90 [1.55, 8.25]
Binar 2017 (JLO)	85.5	3.7	29	80.5	6.5	29	0.9%	5.00 [2.28, 7.72]
Bostanci 2016	82.3	7.4	82	75.7	8.9	82	0.9%	6.60 [4.09, 9.11]
Bowden 2005	81.9	10.1	29	72.7	20.5	29	0.5%	9.20 [0.88, 17.52]
Browaldh 2018	85.3	4.8	60	80.5	6	65	0.9%	4.80 [2.90, 6.70]
Cho 2014	88.3	1.9	23	83.2	3.1	23	0.9%	5.10 [3.61, 6.59]
Choi 2011 (AJRA)	79.8	8.2	22	80.6	7.1	22	0.7%	-0.80 [-5.33, 3.73]
Choi 2012	86.3	6.5	36	73.3	12.7	36	0.7%	13.00 [8.34, 17.66]
Choi 2013	82.9	11.3	20	76.4	12	20	0.5%	6.50 [-0.72, 13.72]
Conway 1985	62.5	17.5	20	37.5	23.1	20	0.3%	25.00 [12.30, 37.70]
Cui 2016	83.9	7.4	12	80.7	8.2	12	0.6%	3.20 [-3.05, 9.45]
Dahlof 2002	81.7	9.5	51	76	10.4	53	0.8%	5.70 [1.87, 9.53]
El-Ahl 2016	89.3	5.3	24	79.2	4.1	24	0.9%	10.10 [7.42, 12.78]
El-Anwar 2018	83.6	4.1	40	70.3	13	40	0.8%	13.30 [9.08, 17.52]
El-Anwar 2019	87.2	7.2	40	76.5	6.8	40	0.8%	10.70 [7.63, 13.77]
Elbassiouny 2015	88.3	1.2	28	75.7	1.3	28	1.0%	12.60 [11.94, 13.26]
Emara 2011	87.2	11.1	23	78.9	12.6	23	0.6%	8.30 [1.44, 15.16]
Emara 2016	89.3	11.1	38	79.9	14.8	38	0.6%	9.40 [3.52, 15.28]
Eun 2008	79.4	16.5	56	79.1	5.7	56	0.7%	0.30 [-4.27, 4.87]
Eun 2009	80.9	6.5	90	77.8	7.8	90	0.9%	3.10 [1.00, 5.20]
Friedman 2003	86.4	9.8	277	82.6	11.4	277	0.9%	3.80 [2.03, 5.57]
Gunawardena 2013	88	5.2	27	84	5.2	27	0.9%	4.00 [1.23, 6.77]
Ha 2020	83.8	7.6	95	79	8.1	95	0.9%	4.80 [2.57, 7.03]
Hendler 2001	80.4	12.3	33	72.4	15.2	33	0.6%	8.00 [1.33, 14.67]
Hester 1995	81.9	6.2	15	71.1	9.8	15	0.6%	10.80 [4.93, 16.67]
Hobson 2012	88	6.5	27	84.4	5.6	27	0.8%	3.60 [0.36, 6.84]
Huang 2008	84	5.2	50	76	8.5	50	0.9%	8.00 [5.24, 10.76]
Huang 2016	82.4	11.6	40	79.9	9.5	40	0.7%	2.50 [-2.15, 7.15]
Huntley 2019 (L)	84.1	8.9	24	80.5	10.6	24	0.7%	3.60 [-1.94, 9.14]
Jacobowitz 2006	82.8	7.1	37	79	7.4	37	0.8%	3.80 [0.50, 7.10]
Jung 2020	83.1	6.3	60	78.1	8.3	60	0.9%	5.00 [2.36, 7.64]
Karakoc 2018	84.7	6.7	79	79.3	11.1	79	0.9%	5.40 [2.54, 8.26]
Kayhan 2016	83.9	6.4	25	79.5	8.8	25	0.8%	4.40 [0.13, 8.67]
Kezirian 2010	80.7	7.6	30	82.4	7.7	30	0.8%	-1.70 [-5.57, 2.17]
Khan 2009	80.7	10.7	63	73.7	17.2	63	0.7%	7.00 [2.00, 12.00]
Kim 2013	86.4	11.3	92	77	10.8	92	0.8%	9.40 [6.21, 12.59]
Kim 2013 (EAO)	84.5	6.2	70	80.6	5.5	70	0.9%	3.90 [1.96, 5.84]
Kinoshita 2006	80.3	11.8	15	66	12.8	15	0.5%	14.30 [5.49, 23.11]
Larsson 1991	82	9.4	50	72.6	12.8	50	0.7%	9.40 [5.00, 13.80]
Lee 2009 (OHNS)	79.8	6.1	30	70.1	8.2	30	0.8%	9.70 [6.04, 13.36]
Lee 2011 (OHNS)	81	12.2	30	75.5	9.6	30	0.7%	5.50 [-0.06, 11.06]
Lee 2012	81.7	8.2	13	75.8	9.6	13	0.6%	5.90 [-0.96, 12.76]
Li 2004	84.8	10.2	55	78.9	9.8	55	0.8%	5.90 [2.16, 9.64]
Li 2008	79.5	12.5	51	78.3	11.9	51	0.7%	1.20 [-3.54, 5.94]
Li 2013 (AOL)	82	5	67	65	13	67	0.8%	17.00 [13.66, 20.34]
Li 2013 (EAO)	78.8	8.4	78	65.6	15.6	78	0.8%	13.20 [9.27, 17.13]
Li 2014	73	14	73	68	12	73	0.8%	5.00 [0.77, 9.23]
Li 2016	83.3	5.6	25	77.1	10.5	25	0.7%	6.20 [1.54, 10.86]
Li 2018	87	5.4	25	80.5	8.5	25	0.8%	6.50 [2.55, 10.45]
Lin 2010	82.1	10.9	43	75.5	10.4	43	0.7%	6.60 [2.10, 11.10]
Lin 2013	84	6.4	12	83.3	5.5	12	0.7%	0.70 [-4.07, 5.47]
Lin 2014	80.1	11.3	35	70.4	9.9	35	0.7%	9.70 [4.72, 14.68]
Lin 2015	83.4	7.3	39	81.6	8.1	39	0.8%	1.80 [-1.62, 5.22]
Lin 2017	80.8	8.2	47	78.2	9.9	47	0.8%	2.60 [-1.08, 6.28]
Liu 2013	79.3	11.3	51	70.8	10.9	51	0.8%	8.50 [4.19, 12.81]
MacKay 2013 (JLO)	85.7	7.1	17	82	8.9	17	0.7%	3.70 [-1.71, 9.11]
Miller 2002	88.1	4.9	15	81.8	5.6	15	0.8%	6.30 [2.53, 10.07]
Miller 2004	88.3	7.1	24	80.1	8.4	24	0.7%	8.20 [3.80, 12.60]
Miyazaki 1998	90.9	5.8	10	88.1	10.6	10	0.5%	2.80 [-4.69, 10.29]

Nakata 2006	93.4	6.4	30	81.4	9.1	30	0.8%	12.00 [8.02, 15.98]
Neruntarat 2003	88.8	2.9	31	81.8	3.8	31	0.9%	7.00 [5.32, 8.68]
Neruntarat 2011	89.2	4.8	83	82.6	5.4	83	0.9%	6.60 [5.05, 8.15]
Pang 2016	89	3.6	73	83.8	8.2	73	0.9%	5.20 [3.15, 7.25]
Pang 2020 (L p.551)	83.9	7.7	326	79.6	9.5	326	0.9%	4.30 [2.97, 5.63]
Park 2018	86.1	5.2	29	78.2	9.4	29	0.8%	7.90 [3.99, 11.81]
Peng 2016	75.6	14	52	55.5	17	52	0.6%	20.10 [14.11, 26.09]
Peng 2019	80.9	8.3	48	70.3	7.1	48	0.8%	10.60 [7.51, 13.69]
Philip-Joet 1991	67.4	19.5	14	51.2	19.6	14	0.2%	16.20 [1.72, 30.68]
Robinson 2009	86.5	6.8	77	81.4	8.1	77	0.9%	5.10 [2.74, 7.46]
Ryan 1990	74	15.5	60	64	18.6	60	0.6%	10.00 [3.87, 16.13]
Salapatas 2016	88.7	4.7	601	87.4	5.4	601	1.0%	1.30 [0.73, 1.87]
Sanders 1990	83.6	9.2	21	73.4	11.5	21	0.6%	10.20 [3.90, 16.50]
Santos 2007	88.4	3.1	10	82.9	9	10	0.6%	5.50 [-0.40, 11.40]
Sarber 2020	89	3.1	31	82.3	4.5	31	0.9%	6.70 [4.78, 8.62]
Shin 2013	86.5	8.2	30	82.7	10	30	0.7%	3.80 [-0.83, 8.43]
Shine 2009	87.4	4.3	60	83.9	5.4	60	0.9%	3.50 [1.75, 5.25]
Shuaib 2015	87.6	5.5	26	85.6	5.5	26	0.8%	2.00 [-0.99, 4.99]
Sorrenti 2003	84.9	5.2	15	79.1	5.6	15	0.8%	5.80 [1.93, 9.67]
Suh 2013	82.6	7.7	50	78.2	9.4	50	0.8%	4.40 [1.03, 7.77]
Sun 2008	75	12.4	31	72.6	11.9	31	0.6%	2.40 [-3.65, 8.45]
Sundman 2021	85.3	4.8	60	80.5	6	65	0.9%	4.80 [2.90, 6.70]
Tan 2014	83.9	8.6	34	73.3	13.7	34	0.7%	10.60 [5.16, 16.04]
Thaler 2016	83.1	7.3	75	78.8	9.6	75	0.9%	4.30 [1.57, 7.03]
Toh 2014	84.5	7.1	20	72.9	19.3	20	0.4%	11.60 [2.59, 20.61]
Turhan 2015 (EAO p3411)	82.4	6.6	90	75.6	9.3	90	0.9%	6.80 [4.44, 9.16]
Turhan 2020	85.1	5.8	64	80.4	9.4	64	0.9%	4.70 [1.99, 7.41]
Vicini 2010 (ORL)	80.2	8.7	10	72.7	10.9	10	0.5%	7.50 [-1.14, 16.14]
Vicini 2014 (ORL)	83.8	6.4	243	79.2	9.1	243	0.9%	4.60 [3.20, 6.00]
Vilaseca 2002	74.8	12.2	20	68.3	10.1	20	0.6%	6.50 [-0.44, 13.44]
Walker 1989	58.3	21.4	11	50	19.1	11	0.2%	8.30 [-8.65, 25.25]
Wang 2013	85.6	10	36	70.5	12.4	36	0.7%	15.10 [9.90, 20.30]
Wee 2015	83.1	6.4	27	78	9.1	27	0.8%	5.10 [0.90, 9.30]
Wetmore 1986	76.4	13	27	64.8	20.3	27	0.4%	11.60 [2.51, 20.69]
Xiao 2016	76	9.9	30	73.8	8.5	30	0.7%	2.20 [-2.47, 6.87]
Yi 2009	83.4	9.8	34	71.7	11.8	34	0.7%	11.70 [6.54, 16.86]
Yin 2007	81.3	13.3	18	72.4	7.1	18	0.6%	8.90 [1.94, 15.86]
Zonato 2006	80	11	17	82.2	9.8	17	0.6%	-2.20 [-9.20, 4.80]
Subtotal (95% CI)			**5431**			**5443**	**78.9%**	**6.53 [5.61, 7.44]**

Heterogeneity: Tau² = 17.83; Chi² = 1221.11, df = 104 (P < 0.00001); I² = 91%
Test for overall effect: Z = 14.02 (P < 0.00001)

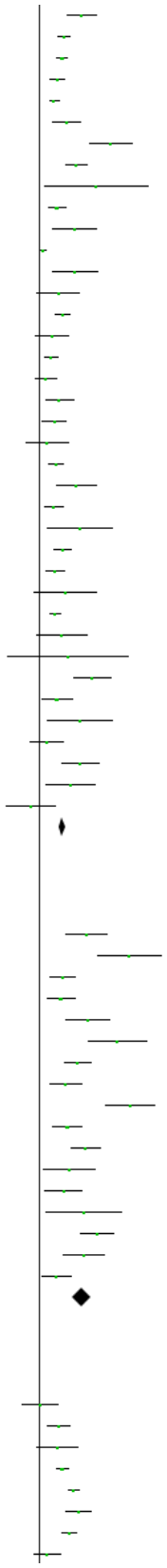
1.4.2 Skeletal procedures

Blumen 2009	84.2	7.4	50	70.7	19.2	50	0.7%	13.50 [7.80, 19.20]
Goh 2003	83.9	8.8	11	58.6	12.3	11	0.4%	25.30 [16.36, 34.24]
Islam 2015 (BJOMS v1)	82.3	7.4	50	75.6	10.7	50	0.8%	6.70 [3.09, 10.31]
Islam 2015 (BJOMS v3)	81.9	6.7	39	75.7	10.5	39	0.8%	6.20 [2.29, 10.11]
Li 2000 (L p.1684)	88.1	4	19	74.3	13.2	19	0.6%	13.80 [7.60, 20.00]
Li 2000 (L p.982)	86	7.9	21	63.9	17.7	21	0.5%	22.10 [13.81, 30.39]
Li 2000 (OHNS)	87.3	4.6	42	76.3	11.2	42	0.8%	11.00 [7.34, 14.66]
Lin 2011	90.6	3.6	12	83	7.2	12	0.7%	7.60 [3.05, 12.15]
Liu 2012	88.6	4.6	11	63	10.7	11	0.6%	25.60 [18.72, 32.48]
Liu 2015	88.9	3.4	16	80.8	7.6	16	0.8%	8.10 [4.02, 12.18]
Liu 2016	94.1	3.5	20	80.9	8.9	20	0.8%	13.20 [9.01, 17.39]
Lye 2008	85	8.2	15	76.5	11.4	15	0.6%	8.50 [1.39, 15.61]
Makovey 2017	85.1	6.2	19	78.1	9.8	19	0.7%	7.00 [1.79, 12.21]
Manikandhan 2014	81	12.9	13	68.4	14.7	13	0.4%	12.60 [1.97, 23.23]
Riley 2000	85.8	3.9	40	69.5	14.8	40	0.7%	16.30 [11.56, 21.04]
Ronchi 2013	86.5	5.7	15	73.9	10.1	15	0.6%	12.60 [6.73, 18.47]
Varghese 2012	86.2	5.6	21	81.2	8	21	0.8%	5.00 [0.82, 9.18]
Subtotal (95% CI)			**414**			**414**	**11.1%**	**11.95 [9.37, 14.53]**

Heterogeneity: Tau² = 21.12; Chi² = 67.24, df = 16 (P < 0.00001); I² = 76%
Test for overall effect: Z = 9.07 (P < 0.00001)

1.4.3 Upper airway stimulation

Bowen 2018	83.5	6.6	14	83.1	7	13	0.7%	0.40 [-4.74, 5.54]
Choi 2015	83.6	6.7	62	77.9	10.9	62	0.8%	5.70 [2.52, 8.88]
Heiser 2017	79.3	11.6	31	74.1	11.4	31	0.6%	5.20 [-0.53, 10.93]
Huntley 2018 (AORL)	87.5	3.9	108	80.9	7.5	108	0.9%	6.60 [5.01, 8.19]
Huntley 2018 (OHNS)	88.1	4	164	78.2	8.8	164	0.9%	9.90 [8.42, 11.38]
Huntley 2019 (ENT)	89.9	1.8	27	78.8	9	27	0.8%	11.10 [7.64, 14.56]
Huntley 2019 (L)	88.8	3.2	76	80.3	8.4	76	0.9%	8.50 [6.48, 10.52]
Kent 2016	82.2	5.2	20	79.8	6.8	20	0.8%	2.40 [-1.35, 6.15]



Lai 2018	82.4	10.9	121	80.1	10.8	121	0.9%	2.30 [-0.43, 5.03]
Mahmoud 2018	91	2.7	47	78	8.9	47	0.9%	13.00 [10.34, 15.66]
Mahmoud 2019	91	2.3	82	78	11.6	82	0.9%	13.00 [10.44, 15.56]
Steffen 2018 (L Feb)	80.9	6.4	56	71.4	11.4	60	0.8%	9.50 [6.16, 12.84]
Subtotal (95% CI)			**808**			**811**	**10.0%**	**7.60 [5.56, 9.64]**
Heterogeneity: Tau² = 10.40; Chi² = 77.62, df = 11 (P < 0.00001); I² = 86%								
Test for overall effect: Z = 7.29 (P < 0.00001)								
Total (95% CI)			**6653**			**6668**	**100.0%**	**7.22 [6.41, 8.03]**
Heterogeneity: Tau² = 17.98; Chi² = 1447.77, df = 133 (P < 0.00001); I² = 91%								
Test for overall effect: Z = 17.40 (P < 0.00001)								
Test for subgroup differences: Chi² = 15.23, df = 2 (P = 0.0005), I² = 86.9%								

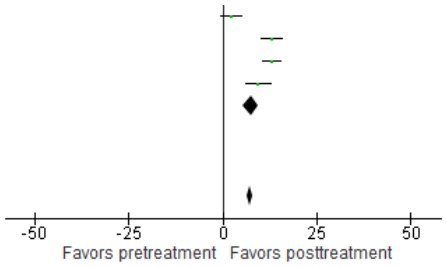


Figure S27. OR-based surgery vs. control (ODI) [RCT]

	Surgery			Control				Mean Difference
Study or Subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Random, 95% CI**
2.8.1 Soft tissue procedures								
Browaldh 2013	14	13.1	32	35.6	21.3	33	43.6%	-21.60 [-30.17, -13.03]
MacKay 2020	-17.7	16	50	-4.2	16.6	49	56.4%	-13.50 [-19.92, -7.08]
Subtotal (95% CI)			**82**			**82**	**100.0%**	**-17.03 [-24.91, -9.16]**
Heterogeneity: Tau² = 17.88; Chi² = 2.20, df = 1 (P = 0.14); I² = 54%								
Test for overall effect: Z = 4.24 (P < 0.0001)								
Total (95% CI)			**82**			**82**	**100.0%**	**-17.03 [-24.91, -9.16]**
Heterogeneity: Tau² = 17.88; Chi² = 2.20, df = 1 (P = 0.14); I² = 54%								
Test for overall effect: Z = 4.24 (P < 0.0001)								
Test for subgroup differences: Not applicable								

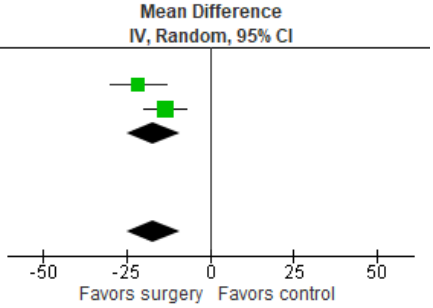
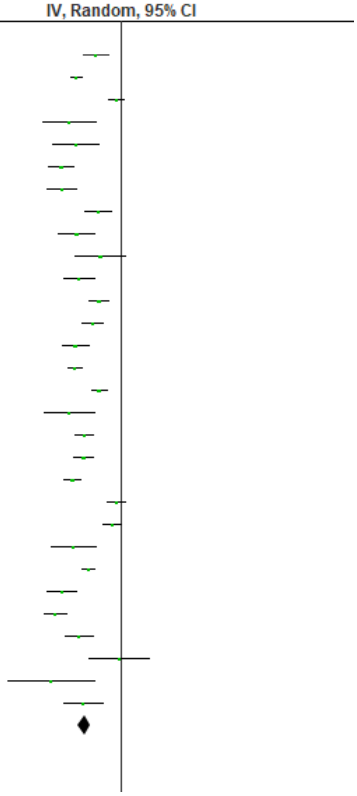


Figure S28. OR-based surgery pretreatment vs. posttreatment (ODI)

	Posttreatment			Pretreatment				Mean Difference
Study or Subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Random, 95% CI**
1.11.1 Soft tissue procedures								
Azbay 2016	37.8	17	67	49	16.7	67	2.3%	-11.20 [-16.91, -5.49]
Barbieri 2019	3.6	2.8	42	23.8	8.4	42	2.5%	-20.20 [-22.88, -17.52]
Bayir 2016	10.2	7.3	35	12	7.6	35	2.5%	-1.80 [-5.29, 1.69]
Binar 2017 (EAO)	9.4	11	23	32.8	28	23	1.6%	-23.40 [-35.69, -11.11]
Binar 2017 (JLO)	11.4	12.8	29	31.9	26.4	29	1.7%	-20.50 [-31.18, -9.82]
Bostanci 2016	17.7	15.9	82	44.8	21.4	82	2.3%	-27.10 [-32.87, -21.33]
Browaldh 2018	16	16.4	60	42.8	22.7	65	2.2%	-26.80 [-33.70, -19.90]
Dahlof 2002	12.6	15.1	51	22.8	15.3	53	2.3%	-10.20 [-16.04, -4.36]
Gunawardena 2013	6.1	5.2	27	26.1	21.8	27	2.0%	-20.00 [-28.45, -11.55]
Hasselbacher 2018 (SB)	16.1	13	15	25.3	18.6	15	1.7%	-9.20 [-20.68, 2.28]
Jung 2020	21	17.4	60	40	20.9	60	2.2%	-19.00 [-25.88, -12.12]
Karakoc 2018	14.5	13.6	79	24.2	15.5	79	2.4%	-9.70 [-14.25, -5.15]
Kim 2013 (EAO)	11.6	13	70	24.4	15.3	70	2.4%	-12.80 [-17.50, -8.10]
Larsson 1991	11.8	9.4	50	32.4	20.6	50	2.2%	-20.60 [-26.88, -14.32]
Larsson 1994	11	5	48	32	10	48	2.5%	-21.00 [-24.16, -17.84]
Lee 2009 (OHNS)	18.4	5.8	30	28.2	7.7	30	2.5%	-9.80 [-13.25, -6.35]
Lee 2011 (OHNS)	14.1	15.4	30	37.5	27.9	30	1.7%	-23.40 [-34.80, -12.00]
Lundkvist 2009	10.8	11.5	120	27.2	20.1	120	2.4%	-16.40 [-20.54, -12.26]
Montevecchi 2018	12.7	10.8	111	29.6	20.7	111	2.4%	-16.90 [-21.24, -12.56]
Mora 2012	2.1	1.8	21	24	8.7	21	2.5%	-21.90 [-25.70, -18.10]
Moxness 2014	12.4	10.7	59	14.2	12.2	59	2.4%	-1.80 [-5.94, 2.34]
Mutlu 2017	8.8	6.4	25	12.5	8.1	25	2.4%	-3.70 [-7.75, 0.35]
Park 2018	11.6	15.4	29	33.1	24.1	29	1.8%	-21.50 [-31.91, -11.09]
Roustan 2018	4.8	2.7	20	19.3	6.1	20	2.5%	-14.50 [-17.42, -11.58]
Sundman 2021	16	16.4	60	42.8	22.7	65	2.2%	-26.80 [-33.70, -19.90]
Turhan 2015 (EAO p3411)	18.2	15.5	90	48	19.5	90	2.3%	-29.80 [-34.95, -24.65]
Turhan 2020	14.5	13.9	64	33.4	22.7	64	2.2%	-18.90 [-25.42, -12.38]
Verse 2002	39.7	25.2	26	40.3	25.8	26	1.4%	-0.60 [-14.46, 13.26]
Vicini 2015	12.9	16.3	10	44.7	27.3	10	1.0%	-31.80 [-51.51, -12.09]
Wee 2015	14.7	13.1	27	31.7	20.7	27	1.9%	-17.00 [-26.24, -7.76]
Subtotal (95% CI)			**1460**			**1472**	**64.2%**	**-16.55 [-19.51, -13.58]**
Heterogeneity: Tau² = 55.86; Chi² = 277.61, df = 29 (P < 0.00001); I² = 90%								
Test for overall effect: Z = 10.92 (P < 0.00001)								



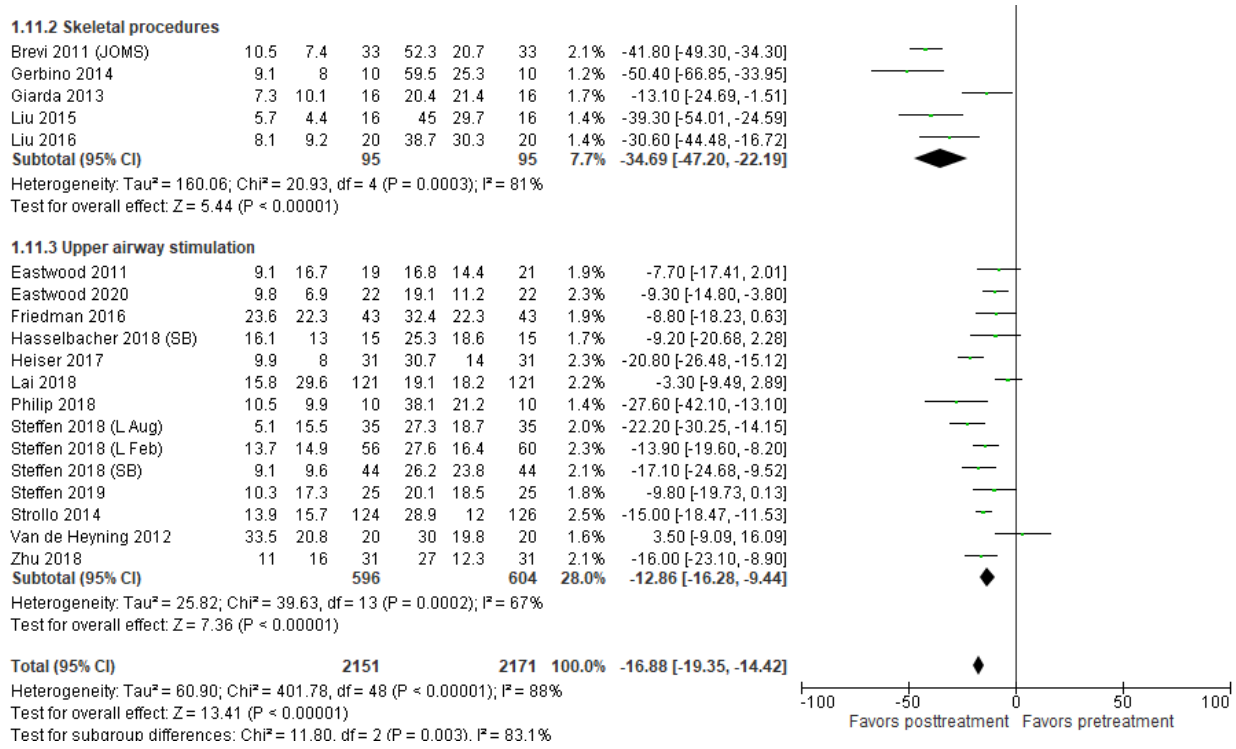


Figure S29. OR-based surgery pretreatment vs. posttreatment (persistent taste alteration, incidence)

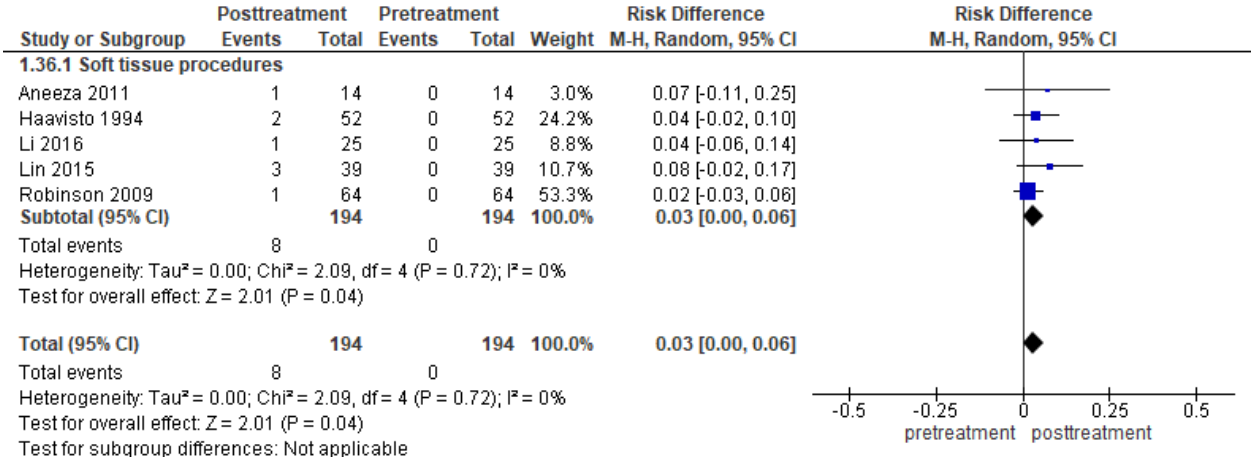


Figure S30. OR-based surgery pretreatment vs. posttreatment (persistent mandibular paresthesia, incidence)

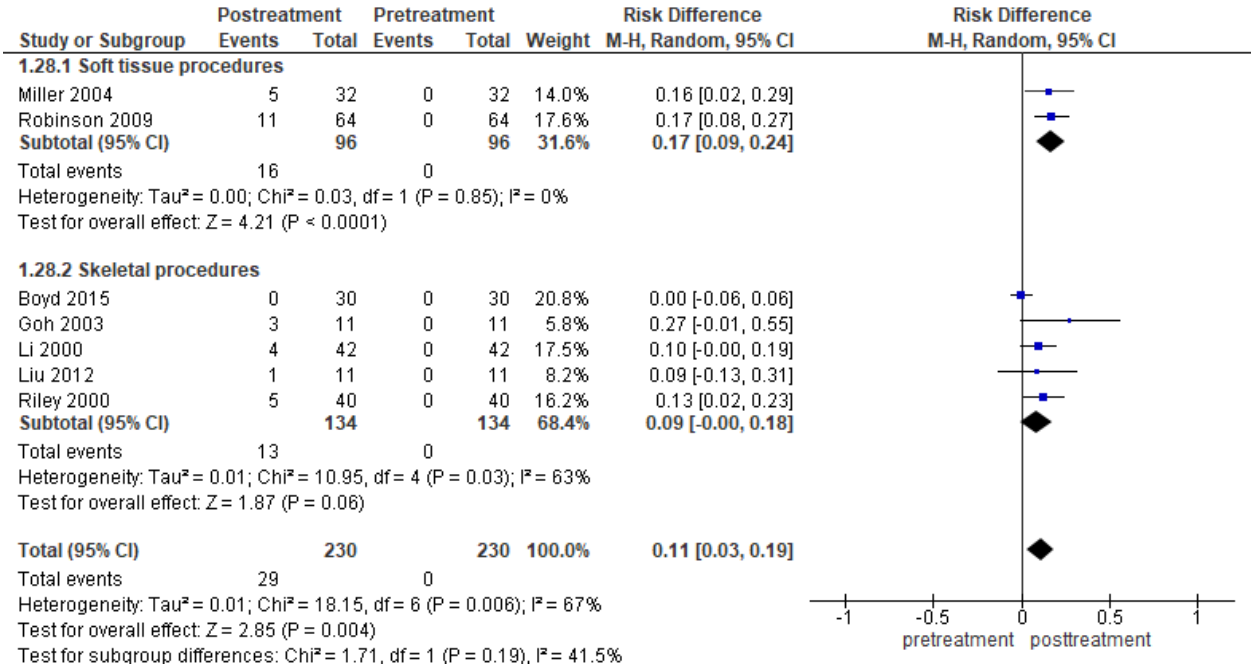


Figure S31. OR-based surgery pretreatment vs. posttreatment (persistent perceived worsening of facial appearance, incidence)

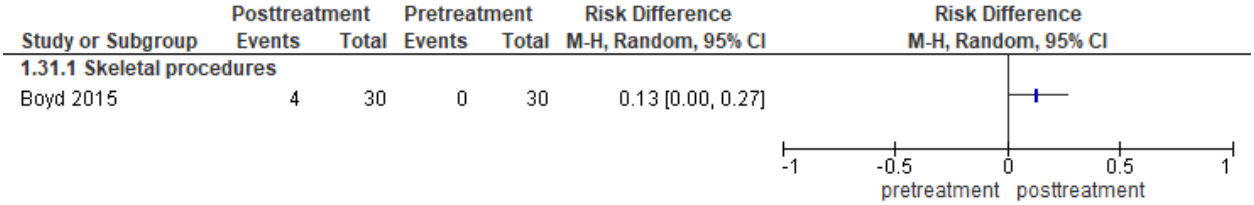


Figure S32. OR-based surgery pretreatment vs. posttreatment (persistent aspiration, incidence)

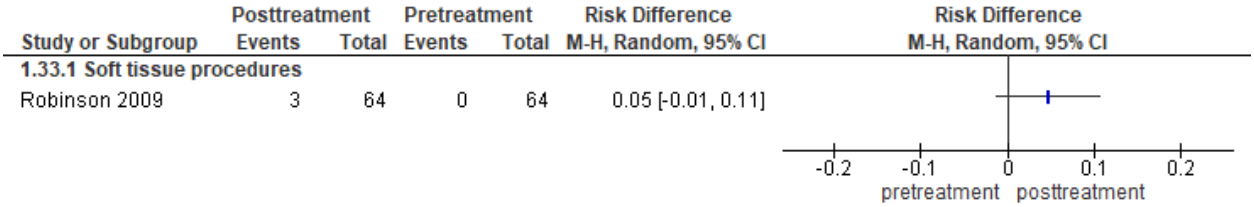


Figure S33. OR-based surgery pretreatment vs. posttreatment (persistent hemorrhage, incidence)

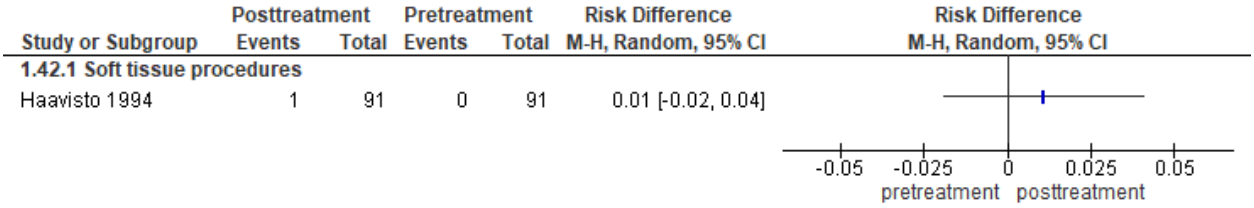


Figure S34. OR-based surgery pretreatment vs. posttreatment (persistent globus pharyngeus, incidence)

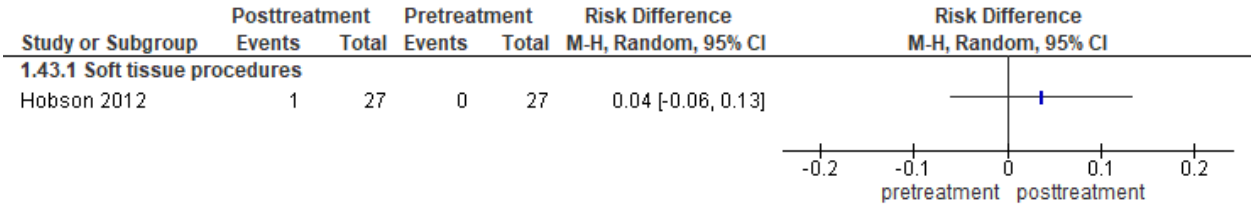


Table S1. Summary of findings table for the surgical treatment of obstructive sleep apnea in adults who were intolerant or unaccepting of CPAP

References: Adzreil 2017 (1); Aneeza 2011 (2); Arora 2016 (3); Askar 2018 CO (4); Askar 2018 SB (5); Aynaci 2018 (6); Azbay 2016 (7); Bababemez 2010 (8); Bababemez 2019 (9); Bababemez 2019 (CO) (10); Baisch 2006 (11); Baradaranfar 2015 (12); Barbieri 2019 (13); Barrera 2016 (14); Bayir 2016 (15); Benazzo 2008 (16); Bettega 2000 (17); Bican 2010 (18); Binar 2017 JLO (19); Binar 2017 EAO (20); Blumen 2009 (21); Blumen 2015 (22); Boon 2018 (23); Bostanci 2016 (24); Bowden 2005 (25); Bowen 2018 (26); Boyd 2013 (27); Boyd 2015 (28); Brevi 2011 (29); Browaldh 2013 (30); Browaldh 2016 (31); Browaldh 2018 (32); Cammaroto 2017 (33); Catella 2019 (34); Chen 2014 (35); Cho 2014 (36); Choi 2011a (37); Choi 2011b (38); Choi 2012 (39); Choi 2013 (40); Choi 2015 (41); Choi 2017 (42); Cillo 2013 (43); Conradt 1997 (44); Conradt 1998 (45); Conway 1985 (46); Cui 2016 (47); Dahlof 2002 (48); de Paula Soares 2014 (49); den Herder 2005 (50); Eastwood 2011 (51); Eastwood 2020 (52); El-Ahl 2016 (53); El-Anwar 2018 (54); El-Anwar 2019 (55); Elbassiouny 2015 (56), Emara 2011 (57); Emara 2016 (58); Eun 2008 (59); Eun 2009 (60); Evans 2019 (61); Fibbi 2009 (62); Fiorita 2018 (63); Freidman 2003 (64); Friedman 2004 (65); Friedman 2007 (66); Friedman 2016 (67); Gerbino 2014 (68); Giarda 2013 (69); Goh 2003 (70); Goodday 2016 (71); Gunawardena 2013 (72); Gunbey 2015 (73); Ha 2020 (74); Haavisto 1994 (75); Hamans 2008 (76); Hasselbacher 2018 EAORL (77) Hasselbacher 2018 SB (78); Heiser 2017 (79); Heiser 2019 (80); Hendler 2001 (81); Hester 1995 (82); Hobson 2012 (83); Hochban 1997 (84); Holmlund 2016 (85); Hsieh 2014 (86); Huang 2008 (87); Huang 2016 (88); Huntley 2018 AORL (89); Huntley 2018 OHNS (90); Huntley 2019 ENT (91); Huntley 2019 L (92); Islam 2015a (93); Islam 2015b (94); Jacobowitz 2006 (95); Jung 2020 (96); Karakoc 2018 (97); Karatayli-Ozgursoy 2012 (98); Katsantonis 1990 (99); Kayhan 2016 (100); Kent 2016 (101); Kezirian 2010 (102); Khan 2009 (103); Kim 2013 CO (104); Kim 2013 EAO (105); Kinoshita 2006 (106); Komada 2012 (107); Lai 2018 (108); Larsson 1991 (109); Larsson 1994 (110); Lee 1999 (111); Lee 2009 OHNS (112); Lee 2010 (1113); Lee 2011 (114); Lee 2012 (115); Lee 2018 (116); Li 2000 (117); Li 2000 (L) (118); Li 2000 (L p.982) (119); Li 2003 (120); Li 2004 (121); Li 2005a (122); Li 2005b (123); Li 2006 (124); Li 2008 (125); Li 2009 (126); Li 2013a (127); Li 2013b (128); Li 2013c (129); Li 2014 (130); Li 2015 (131); Li 2016 (132); Li 2018 (133); Lim 2018 (134); Lin 2006 (135); Lin 2010 (136); Lin 2011 (137); Lin 2013 (138); Lin 2014 (139); Lin 2015 (140); Lin 2017 (141); Liu 2012 (142); Liu 2013 (143); Liu 2015 (144); Liu 2016 (145); Lundkvist 2009 (146); Lye 2008 (147); Mackay 2013a (148); MacKay 2013b (149); MacKay 2020 (150); Mahmoud 2018 (151); Mahmoud 2019 (152); Makovey 2017 (153); Manikandhan 2014 (154); Meraj 2017 (155); Miller 2002 (156); Miller 2004 (157); Miyazaki 1998 (158); Montevecchi 2018 (159); Mora 2012 (160); Moxness 2014 (161); Mure 2019 (162); Mutlu 2017 (163); Nakata 2006 (164); Neruntarat 2003 (165); Neruntarat 2011 (166); Omur 2005 (167); Pang 2016 (168); Pang 2017 (169); Pang 2020 (L p.551) (170); Pang 2020 (L p.2281) (171); Parikh 2018 (172); Park 2018 (173); Peng 2016 (174); Peng 2019 (175); Philip 2018 (176); Philip-Joet 1991 (177); Plaza 2019 (178); Riley 2000 (179); Robinson 2009 (180); Ronchi 2013 (181); Rotenberg 2014 (182); Roustan 2018 (183); Ryan 1990 (184); Salapatas 2016 (185); Sanders 1990 (186); Santos 2007 (187); Sarber 2020 (188); Schendel 2014 (189); Schwab 2018 (190); Sezen 2011 (191); Shah 2018 (192); Shin 2013 (193); Shine 2009 (194); Shuaib 2015 (195); Simsek 2015 (196); Sommer 2016 (197); Sorrenti 2003 (198); Steffen 2018 (L Aug) (199); Steffen 2018 L Feb (200); Steffen 2018 (SB) (201); Steffen 2019 (202); Strollo 2014 (203); Suh 2013 (204); Sun 2008 (205); Sundman 2021 (206); Suslu 2017 (207); Tan 2014 (208); Thaler 2016 (209); Thaler 2020 (210); Toh 2014 (211); Tsui 2020 (212); Tuncel 2012 (213); Turhan 2015a (214); Turhan 2015b (215); Turhan 2020 (216); Van de Heyning 2012 (217); Varghese 2012 (218); Verse 2002 (219); Vicini 2010 (220); Vicini 2014 HN (221); Vicini 2014 ORL (222); Vicini 2015 (223); Vigneron 2017 (224); Vilaseca 2002 (225); Walia 2019 (226); Walker 1989 (227); Wang 2013 (228); Weaver 2011 (229); Wee 2015 (230); Wetmore 1986 (231); Woodson 2005 (232); Xiao 2016 (233); Yaremchuk 2011 (234); Yi 2009 (235); Yi 2011 (236); Yin 2007 (237); Yuksel 2018 (238); Zhang 2013 (239); Zhang 2014 (240); Zhao 2017 (241); Zhu 2018 (242); Zonato 2006 (243)

Outcomes	Quality of the evidence (GRADE)	Anticipated absolute effects* (95% CI) MD between pre-surgery and post-surgery or surgery and control	№ of participants (studies)
ESS* (Surgery vs Control)	⊕⊕⊕⊕ HIGH	The mean difference in ESS score of the surgery group was 5.6 points [A] lower (7.3 lower to 4.0 lower) than the control group.	199 (3 RCTs) [31,150,197]
ESS* (Pre- vs. Post-surgery)	⊕⊕◯◯ LOW	The mean difference in ESS score after surgery was 5.8 points [A] lower (6.3 lower to 5.4 lower) than before surgery.	7875 (145 observational studies) [1,2,4-7,9-14,16,18-20,23,25,29,32-34,36-38,40,41,43,50-55,57-60,62,63,66,67,69,72,74,77,79,80,83,85,87-91,93-95,100-101,107,108,112,114-116,121,125-129,131-133,135,136,138-146,148-151,153,159-162,164-169,172-176,178,182,183,185,188,191-193,195,197-203,205,206,210-212,216-223,225,228,229,234-238,242]
Sleep-related QOL* (FOSQ) (Surgery vs Control)	⊕⊕⊕◯ MODERATE [1]	The mean difference in FOSQ score of the surgery group was 3.5 points[A] higher (2.6 higher to 4.4 higher) than the control group.	99 (1 RCT) [150]

Sleep-related QOL* (FOSQ) (Pre- vs. Post-surgery)	⊕⊕◯◯ LOW	The mean difference in FOSQ score after surgery was 3.5 points [A] higher (2.9 higher to 4.0 higher) than before surgery.	432 (11 observational studies) [28,51,52,77,102,147,200,203,212,217,229]
Sleep apnea related QOL* (SAQLI) (Pre- vs Post-surgery)	⊕◯◯◯ VERY LOW [1,2]	The mean difference in SAQLI score after surgery was 1 point [B] higher (0.03 higher to 1.7 higher) than before surgery.	80 (3 observational studies) [51,67,212]
General QOL* (SF-36 Physical Component Summary) (Surgery vs Control)	⊕⊕⊕◯ MODERATE [1]	The mean difference in SF-36 Physical Summary Score in the surgery group was 2.9 points [B] higher (1.6 lower to 7.4 higher) than the control group.	62 (1 RCT) [31]
General QOL* (SF-36 Physical Component Summary) (Pre- vs Post-surgery)	⊕◯◯◯ VERY LOW [1,2]	The mean difference in SF-36 Physical Summary Score after surgery was 6.3 points [A] higher (0.4 lower to 13.0 higher) than before surgery.	73 (2 observational studies) [121,212]
General QOL* (SF-36 Mental Component Summary) (Surgery vs Control)	⊕⊕⊕◯ MODERATE [1,2]	The mean difference in SF-36 Mental Summary Score in the surgery group was 5.4 points [A] higher (0.1 higher to 10.7 higher) than the control group.	62 (1 RCT) [31]
General QOL* (SF-36 Mental Component Summary) (Pre- vs Post-surgery)	⊕◯◯◯ VERY LOW [1]	The mean difference in SF-36 Mental Summary Score after surgery was 9.5 points [A] higher (2.0 lower to 20.9 higher) than before surgery.	73 (2 observational studies) [121,212]
General QOL* (SF-36 Vitality Score) (Surgery vs Control)	⊕⊕⊕◯ MODERATE [1,2]	The mean difference in SF-36 Vitality Score in the surgery group was 21.1 points [A] lower (32.7 lower to 9.5 lower) than the control group.	55 (1 RCT) [31]
General QOL* (SF-36 Vitality score) (Pre- vs Post-surgery)	⊕◯◯◯ VERY LOW [1,2]	The mean difference in SF-36 Vitality Score after surgery was 14.4 points [A] higher (9.0 higher to 19.8 higher) than before surgery.	73 (2 observational studies) [121,212]
Sleep quality* (PSQI) (Pre- vs. Post-surgery)	⊕⊕◯◯ LOW	The mean difference in PSQI score after surgery was 1.4 points [B] lower (2.0 lower to 0.7 lower) than before surgery.	148 (4 observational studies) [51,105,173,233]
Snoring*, VAS (Surgery vs Control)	⊕⊕⊕◯ MODERATE [1,2]	The mean difference in snoring score in the surgery group was 3.7 points [A] lower (5.3 lower to 2.1 lower) than the control group.	29 (1 RCT) [197]
Snoring*, VAS (Pre- vs Post-surgery)	⊕⊕◯◯ LOW	The mean difference in snoring score after surgery was 5.2 points [A] lower (5.9 lower to 4.6 lower) than before surgery.	2098 (36 observational studies) [1,4,5,9,19,36,38,53,54,56,64-66,72,83,87,104,116,131,133,136,138,139,141,142,166-169,185,197,211,228,229,235,236,238]
Systolic BP* (Surgery vs Control)	⊕⊕◯◯ LOW [2]	The mean difference in systolic BP in the surgery group was 0.2 points [B] lower (4.9 lower to 4.5 higher) than the control group.	99 (1 RCT) [150]
Systolic BP* (Pre- vs Post-surgery)	⊕◯◯◯ VERY LOW [2]	The mean difference in systolic BP after surgery was 6.3 mmHg [A] lower (11.6 lower to 0.9 lower) than before surgery.	540 (10 observational studies) [49,88,106,112,114,153,169,174,182,226]
Diastolic BP* (Surgery vs Control)	⊕⊕◯◯ LOW [2]	The mean difference in diastolic BP in the surgery group was 0.4 points [B] lower (3.5 lower to 2.7 higher) than the control group.	99 (1 RCT) [150]
Diastolic BP* (Pre- vs Post-surgery)	⊕⊕◯◯ LOW	The mean difference in diastolic BP after surgery was 2.7 mmHg [A] lower (7.9 lower to 2.5 higher) than before surgery.	518 (9 observational studies) [49,88,106,112,114,153,169,174,226]
AHI* (Surgery vs Control)	⊕⊕⊕⊕ HIGH	The mean difference in AHI in the surgery group was 18.4 events/hr [A] lower (26.4 lower to -10.5 lower) than in the control group.	198 (3 RCTs) [30,150,197]

AHI* (Pre- vs Post-surgery)	⊕⊕⊕◯ MODERATE [3]	The mean difference in AHI after surgery was 24.1 events/hr [A] lower (25.7 lower to 22.5 lower) than before surgery. The mean AHI before surgery was 39.0 (18.2) events/hr and after surgery was 16.6 (14.6) events/hr for a 58% reduction.	10652 (194 observational studies) [1,3-14,16-27,29,30,32-44,47,49-62,64,65,67-74,76-80,82-86,88-93,95-104,106-108,113-116,124-134,136-140,142-145,147,149-155,158-164,166,168,170,172-176,178,181-183,185-193,195-197,199-205,207,209-211,213-225,227,228,230-242]
RDI* (Surgery vs Control)	⊕⊕⊕◯ MODERATE [1,2]	The mean difference in RDI in the surgery group was 16.4 events/hr [A] lower (33.3 lower to 0.6 higher) than in the control group.	98 (2 RCTs) [30, 197]
RDI* (Pre- vs Post-surgery)	⊕⊕⊕◯ MODERATE [3]	The mean difference in RDI after surgery was 31.0 events/hr [A] lower (35.7 lower to 26.3 lower) than before surgery. The mean RDI before surgery was 45.3 (20.6) events/hr and after surgery was 16.1 (13.3) events/hr for a 64% reduction.	965 (27 observational studies) [11,44,64,76,82,104,105,110-113,115,116,128,145,146,154,156,157,159,166,167,180,184,190,192]
Permanent dysphagia, Risk* (Pre- vs Post-surgery)	⊕◯◯◯ VERY LOW [2]	The mean difference in risk of permanent dysphagia between pre- and post-surgery was 0.02 [B] higher (-0.01 lower to 0.04 higher).	1175 (10 observational studies) [10,75,85,107,142,167,171,180,210,230]
MD Anderson dysphagia score* (Pre- vs Post-surgery)	⊕◯◯◯ VERY LOW [1,2]	The mean difference in MD Anderson permanent dysphagia score was 0.4 points [B] lower (3.4 lower to 2.6 higher) than before surgery.	24 (1 observational study) [221]
LSAT (Surgery vs Control)	⊕⊕⊕◯ MODERATE [1,2]	The mean difference in LSAT in the surgery group was 4.3% [B] higher (3.1 higher to 5.6 higher) than in the control group.	197 (3 RCTs) [30,150,197]
LSAT (Pre- vs Post-surgery)	⊕⊕◯◯ LOW	The mean difference in LSAT after surgery was 7.2% [A] higher (6.4 higher to 8.0 higher) than before surgery. The mean LSAT before surgery was 78.0 (9.4)% and after surgery was 84.4 (7.2)% for a 8.2% increase.	6653 (133 observational studies) [4-8,11,12,18-21,24-26,30,32,36,37,39-41,46-48,53-60,64,70,72,74,79,81-83,87-97,100-106,108-109,112,114,115,117-119,121,125,127-130,132,133,136-145,147,149-154,156-158,164-166,168,170,173-175,177,179-181,184-188,193-195,197,198,200,204-206,208,209,211,215,216,218,220,222,225,227,228,230,231,233,235,237,243]
ODI (Surgery vs Control)	⊕⊕⊕⊕ HIGH	The mean difference in ODI in the surgery group was 17.0 points [A] lower (24.9 lower to 9.2 lower) than in the control group.	164 (2 RCTs) [30,150]
ODI (Pre- vs Post-surgery)	⊕⊕⊕◯ MODERATE [3]	The mean difference in ODI after surgery was 16.9 points [A] lower (19.4 lower to 14.4 lower) than before surgery. The mean ODI before surgery was 30.9 (18.2) and after surgery was 14.3 (13.8) for a reduction of 54%.	2151 (48 observational studies) [7,13,15,19,20,24,29,32,48,51,52,67-69,72,77-79,96,97,105,108-110,112,114,144-146,159-161,163,173,176,183,199-203,206,214,216,217,219,223,230,242]

*Critical Outcomes
[1]Quality of evidence was downgraded due to imprecision associated with small sample size (i.e., <100)
[2]Quality of evidence was downgraded due to imprecision (i.e., 95% CI of mean difference crosses clinical significance threshold)
[3]Quality of evidence was upgraded due to large effect
[A]Mean difference meets or exceeds clinical significance threshold
[B]Mean difference does not meet clinical significance threshold

Bariatric surgery for the treatment of OSA in obese adults (PICO 2)

Figure S35. Bariatric surgery pretreatment vs. posttreatment (ESS)

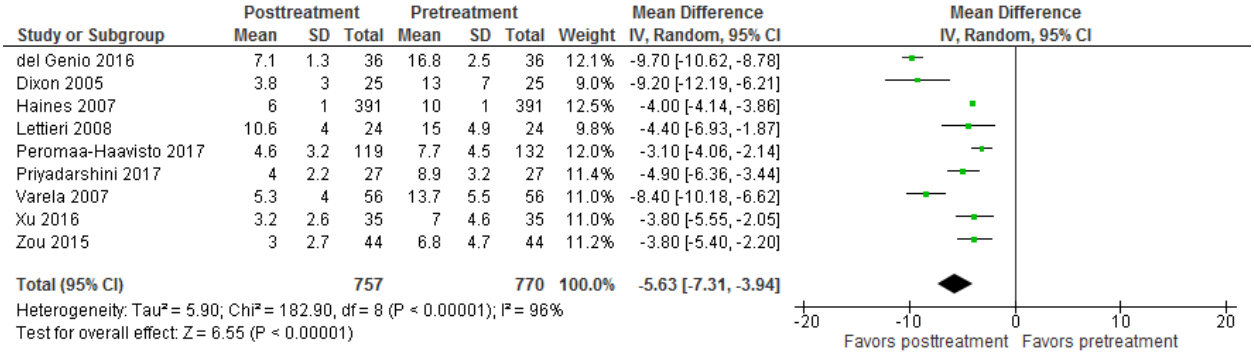


Figure S36. Bariatric surgery pretreatment vs. posttreatment (Daytime sleepiness frequency)

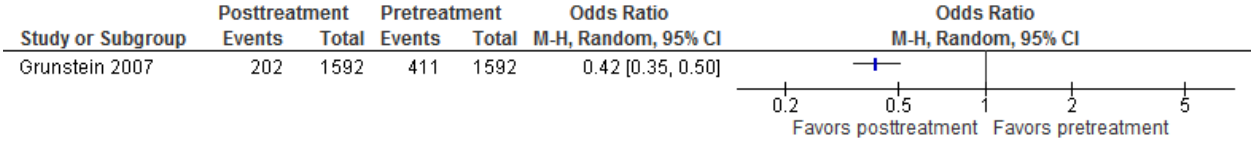


Figure S37. Bariatric surgery pretreatment vs. posttreatment (SBP, mmHg)

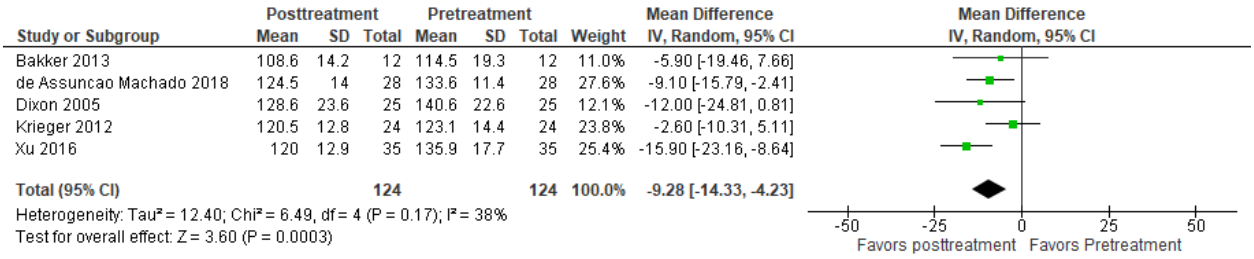


Figure S38. Bariatric surgery pretreatment vs. posttreatment (DBP, mmHg)

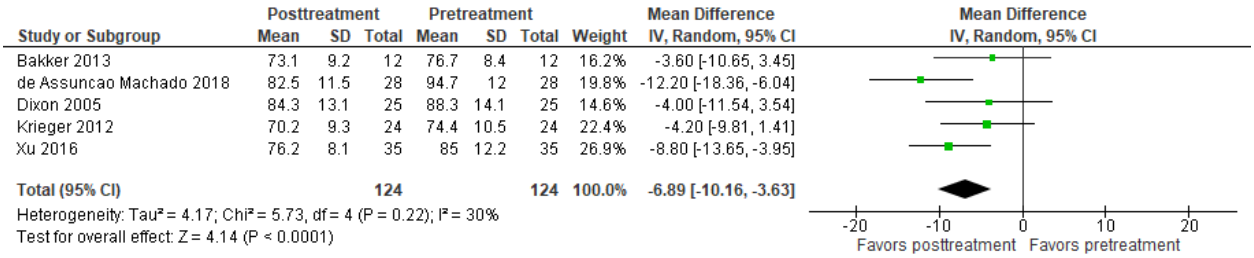


Figure S39. Bariatric surgery vs. nutritional care (AHI, events/hr) [RCT]

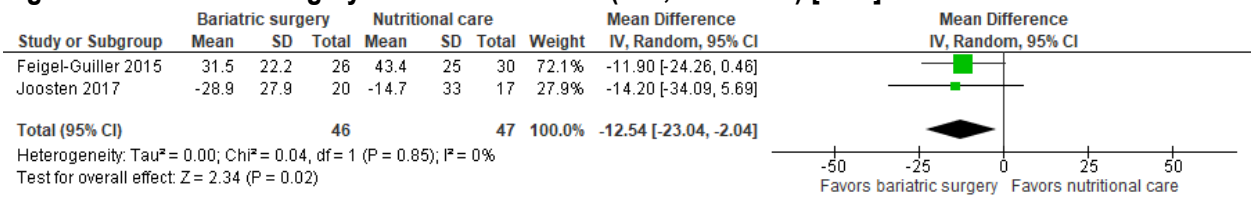
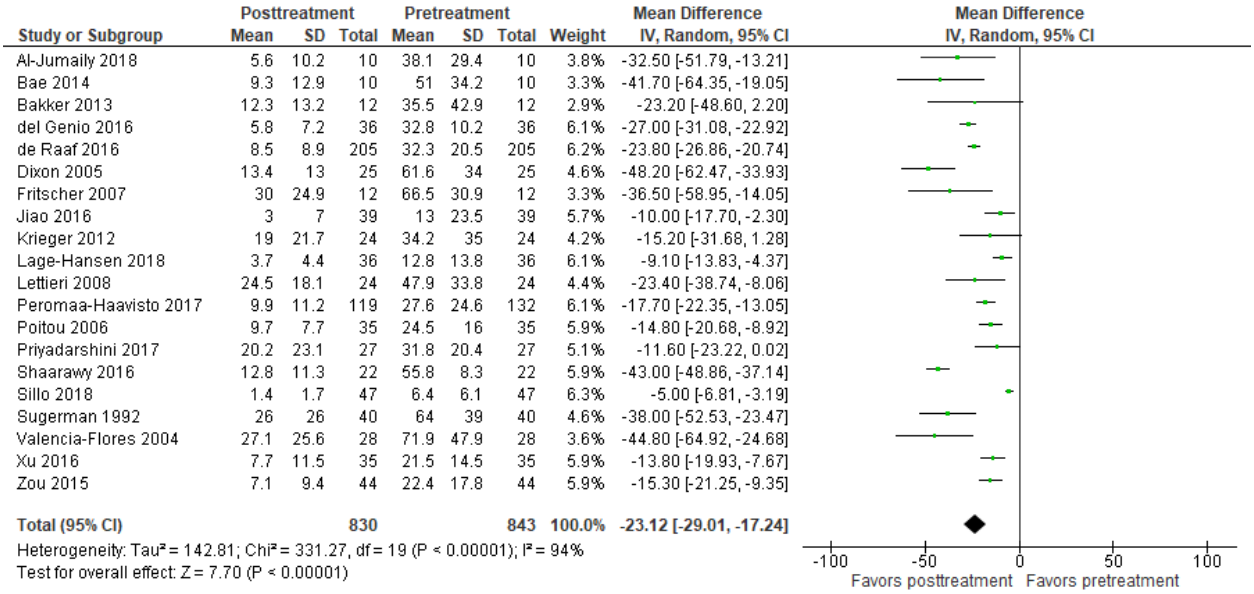
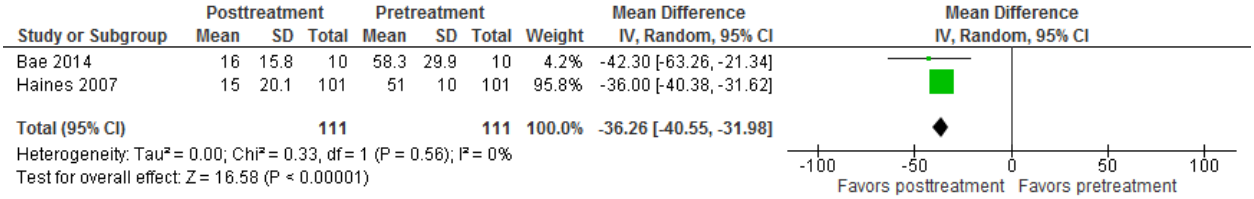


Figure S40. Bariatric surgery pretreatment vs. posttreatment (AHI, events/hr)



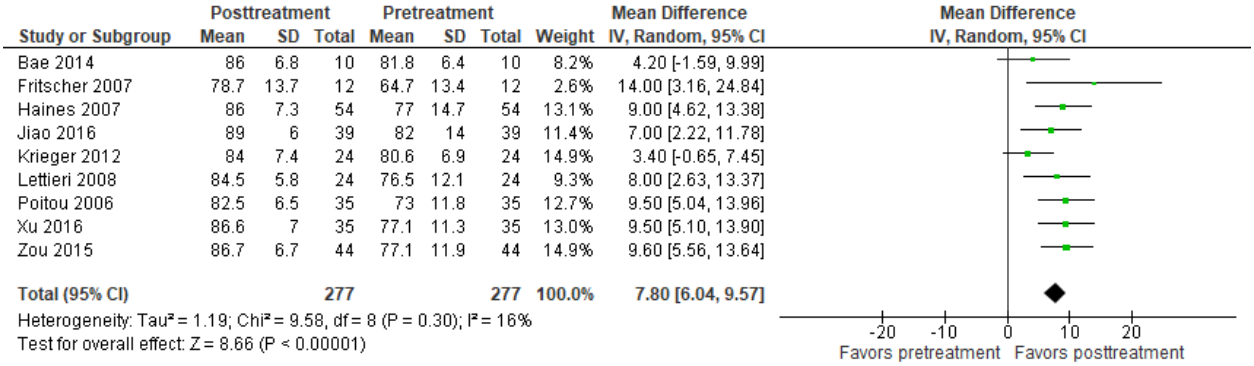
Heterogeneity: Tau² = 142.81; Chi² = 331.27, df = 19 (P < 0.00001); I² = 94%
Test for overall effect: Z = 7.70 (P < 0.00001)

Figure S41. Bariatric surgery pretreatment vs. posttreatment (RDI)



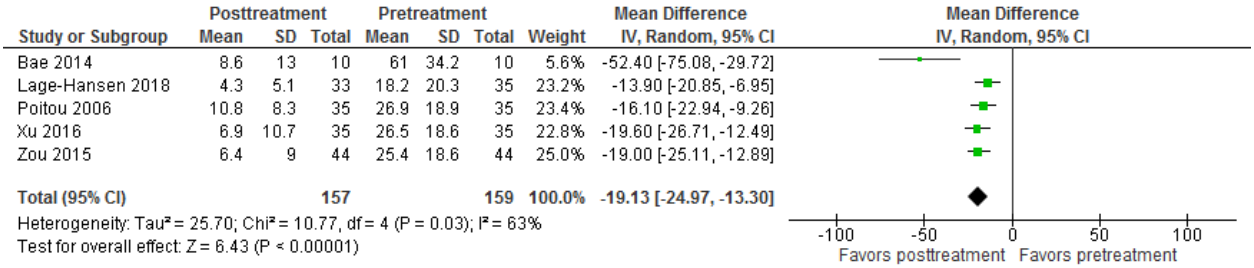
Heterogeneity: Tau² = 0.00; Chi² = 0.33, df = 1 (P = 0.56); I² = 0%
Test for overall effect: Z = 16.58 (P < 0.00001)

Figure S42. Bariatric surgery pretreatment vs. posttreatment (LSAT, %)



Heterogeneity: Tau² = 1.19; Chi² = 9.58, df = 8 (P = 0.30); I² = 16%
Test for overall effect: Z = 8.66 (P < 0.00001)

Figure S43. Bariatric surgery pretreatment vs. posttreatment (ODI)



Heterogeneity: Tau² = 25.70; Chi² = 10.77, df = 4 (P = 0.03); I² = 63%
Test for overall effect: Z = 6.43 (P < 0.00001)

Figure S44. Bariatric surgery pretreatment vs. posttreatment (Snoring, %)

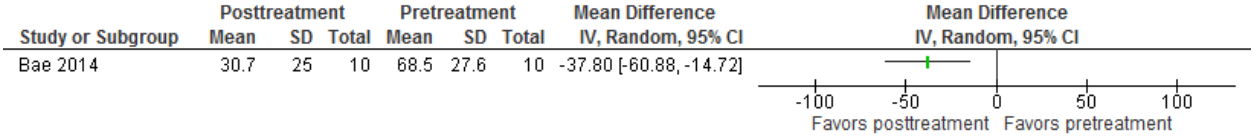


Figure S45. Bariatric surgery pretreatment vs. posttreatment (Snoring frequency)

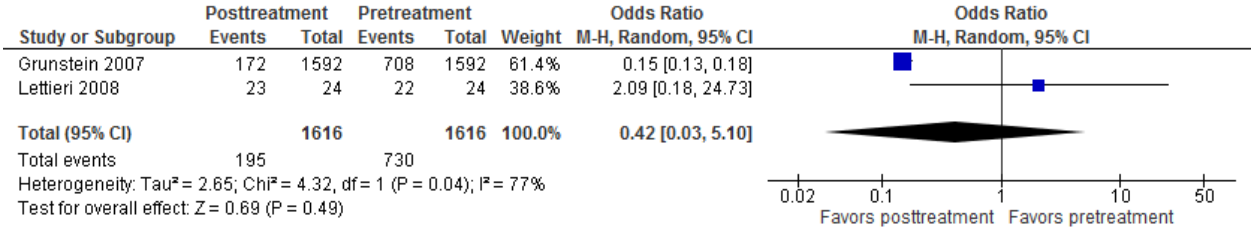


Figure S46. Bariatric surgery vs. nutritional care (BMI, kg/m²) [RCT]

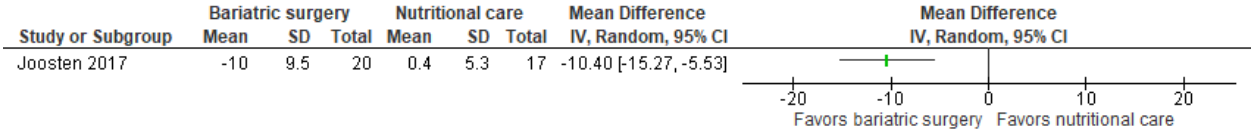


Figure S47. Bariatric surgery pretreatment vs. posttreatment (BMI, kg/m²)

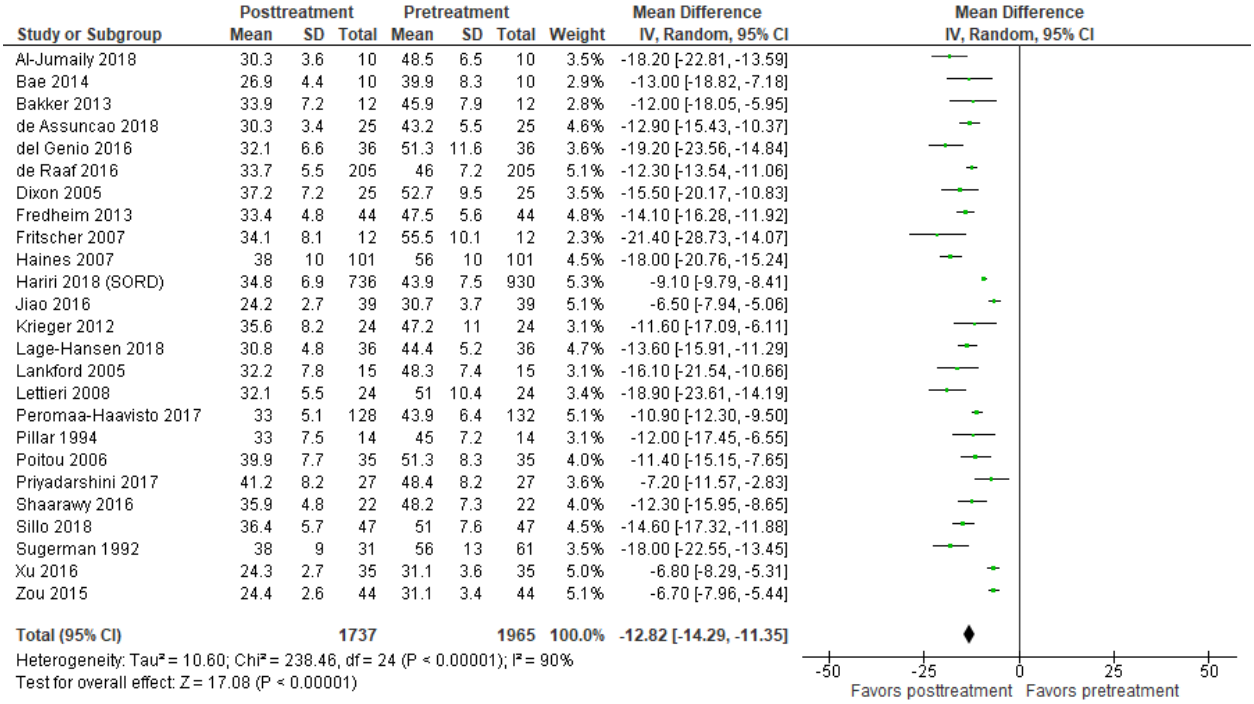


Figure S48. Bariatric surgery pretreatment vs. posttreatment (PAP pressure, cm H_2O)

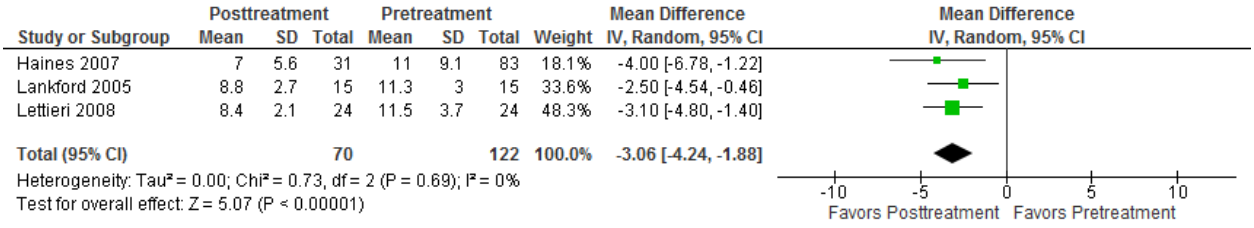


Figure S49. Bariatric surgery pretreatment vs. posttreatment (iron malabsorption, incidence)

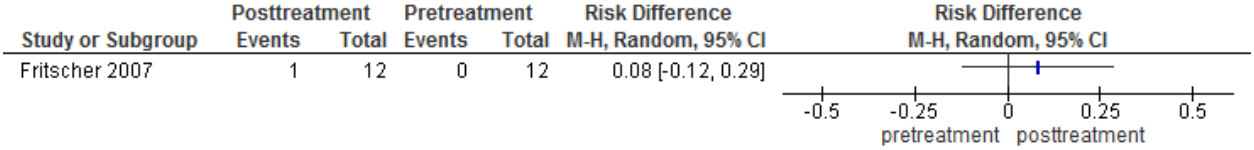


Figure S50. Bariatric surgery vs. nutritional care (gastric ulcer, incidence) [RCT]

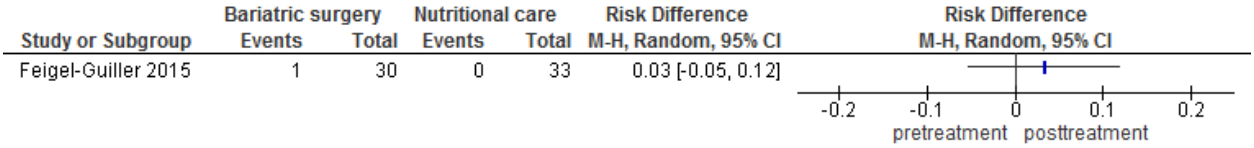


Table S2. Summary of findings table for the bariatric surgical treatment of obstructive sleep apnea in adults with obesity

References: Al-Jumaily 2018 (1); Bae 2014 (2); Bakker 2013 (3); de Assuncao Machado2018 (4); de Raaf 2016 (5); Del Genio 2016 (6); Dixon 2005 (7); Feigel-Guiller 2015 (8); Fredheim 2013 (9); Fritscher 2007 (10); Grunstein 2007 (11); Haines 2007 (12); Hariri 2018 (13); Jiao 2016 (14); Joosten 2017 (15); Krieger 2012 (16); Lage-Hansen 2018 (17); Lankford 2005 (18); Lettieri 2008 (19); Peroma-Haavisto 2017 (20); Pillar 1994 (21); Poitou 2006 (22); Priyadarshini 2017 (23); Shaarawy 2016 (24); Sillo 2018 (25); Sugerman 1992 (26); Valencia-Flores 2004 (27); Varela 2007 (28); Xu 2016 (29); Zou 2015 (30)

Outcomes	Quality of the evidence (GRADE)	Anticipated absolute effects* (95% CI) MD between pre-surgery and post-surgery or surgery and control	№ of participants (studies)
ESS* (Pre- vs Post-surgery)	⊕⊕⊕◯ MODERATE [3]	The mean ESS after surgery was 5.6 points [A] lower (7.3 lower to 3.9 lower) than before surgery.	757 (9 observational studies) [6,7,12,19,20,23,28-30]
Daytime sleepiness* (Pre- vs Post-surgery)	⊕⊕◯◯ LOW	The odds ratio for daytime sleepiness was 0.4 [A] (0.35 to 0.5) after surgery vs before surgery.	1592 (1 observational study) [12]
SBP* (Pre- vs Post-surgery)	⊕⊕⊕◯ MODERATE [3]	The mean SBP after surgery was 9.3 mmHg [A] lower (14.3 lower to 4.2 lower) than before surgery.	124 (5 observational studies) [3,4,7,16,29]
DBP* (Pre- vs Post-surgery)	⊕⊕⊕◯ MODERATE [3]	The mean DBP after surgery was 6.9 mmHg [A] lower (10.2 lower to 3.6 lower) than before surgery.	124 (5 observational studies) [3,4,7,16,29]
AHI* (Surgery vs Control)	⊕⊕⊕◯ MODERATE [2]	The mean AHI in the surgery group was 12.5 events/hr [A] lower (23.0 lower to 2.0 lower) than in the control group.	46 (2 RCTs) [8,15]
AHI* (Pre- vs Post-surgery)	⊕⊕⊕◯ MODERATE [3]	The mean AHI after surgery was 23.1 events/hr [A] lower (29.0 lower to 17.2 lower) than before surgery. The mean AHI before surgery was 34.8 events/hr and was 11.7 events/hr after surgery for a reduction of 66%.	830 (20 observational studies) [1-3,5-7,10,14,16,17,19,20,22-27,29,30]
RDI* (Pre- vs Post-surgery)	⊕⊕⊕◯ MODERATE [1]	The mean RDI after surgery was 36.3 events/hr [A] lower (40.6 lower to 32.0 lower) than before surgery. The mean RDI before surgery was 51.3events/hr and was 15.0 events/hr after surgery for a reduction of 71%.	111 (2 observational studies) [2,12]

LSAT (Pre-vs Post-surgery)	⊕⊕◯◯ LOW	The mean LSAT after surgery was 7.8% [A] higher (6.0 higher to 9.6 higher) than before surgery. The mean LSAT was 77.7% before surgery and was 85.5% after surgery for an increase of 10%.	277 (9 observational studies) [2,10,12,14,16,19,22,29,30]
ODI (Pre- vs. Post-surgery)	⊕⊕⊕◯ MODERATE [3]	The mean ODI after surgery was 19.1 points [A] lower (25.0 lower to 13.3 lower) than before surgery. The mean ODI was 26.3 before surgery and was 7.2 after surgery for a reduction of 73%.	157 (5 observational studies) [2,17,22,29,30]
% of Patients Snoring (Pre- vs Post-surgery)	⊕◯◯◯ VERY LOW [1]	The mean % of patients snoring after surgery was 37.8% [A] lower (60.9 lower to 14.7 lower) than before surgery.	10 (1 observational study) [2]
Snoring Frequency (Pre- vs Post-surgery)	⊕◯◯◯ VERY LOW [2]	The odds ratio of snoring after surgery was 0.4 lower (0.03 lower to 5.10 higher) than before surgery.	1616 (2 observational studies) [11,19]
BMI (Surgery vs Control)	⊕⊕⊕◯ MODERATE [1]	The mean BMI in the surgery group was 10.4 kg/m^2 [A] lower (15.3 lower to 5.5 lower) than in the control group.	37 (1 RCT) [15]
BMI (Pre- vs Post-surgery)	⊕⊕⊕◯ MODERATE [3]	The mean BMI after surgery was 12.8 kg/m^2 [A] lower (14.3 lower to 11.4 lower) than before surgery.	1737 (25 observational studies) [1-7,9,10,12-14,16-26,29,30]
Optimal PAP Pressure (Pre- vs Post-surgery)	⊕◯◯◯ VERY LOW [1]	The mean optimal PAP pressure after surgery was 3.1 cm [A] H_2O lower (4.2 lower to 1.9 lower) than before surgery.	70 (3 observational studies) [12,18,19]

*Critical Outcomes
[1]Quality of evidence was downgraded due to imprecision associated with small sample size (<100 participants)
[2]Quality of evidence was downgraded due to imprecision (i.e., 95% CI of mean difference crosses clinical significance threshold)
[3]Quality of evidence was upgraded due to large effect
[A]Mean difference meets or exceeds clinical significance threshold
[B]Mean difference does not meet clinical significance threshold

Surgical treatment of OSA in adults to facilitate use of CPAP (PICO 3)

Figure S51. OR-based adjunctive surgery pretreatment vs. posttreatment (ESS)

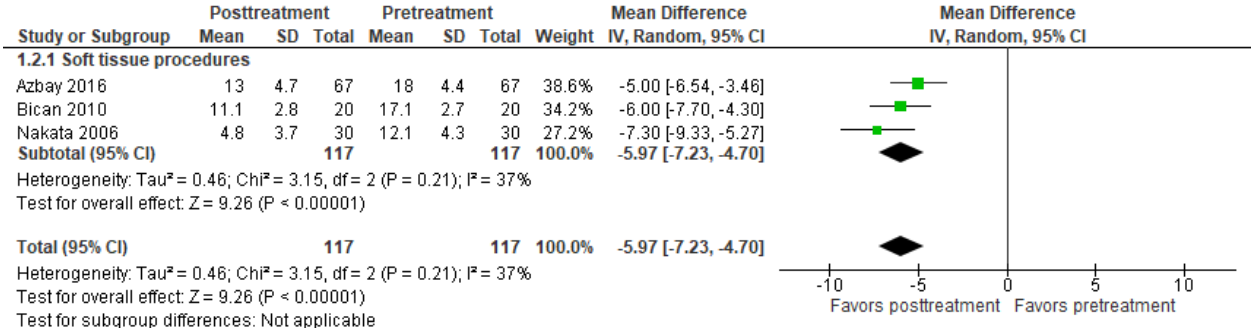


Figure S52. OR-based adjunctive surgery pretreatment vs. posttreatment (optimal CPAP level, cm H2O)

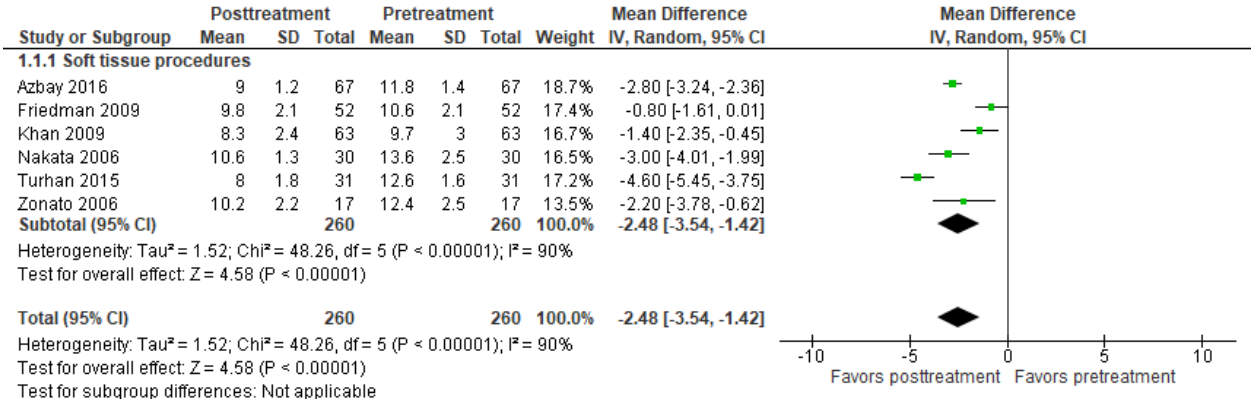


Figure S53. OR-based adjunctive surgery pretreatment vs. posttreatment (Adherence, hrs/night)

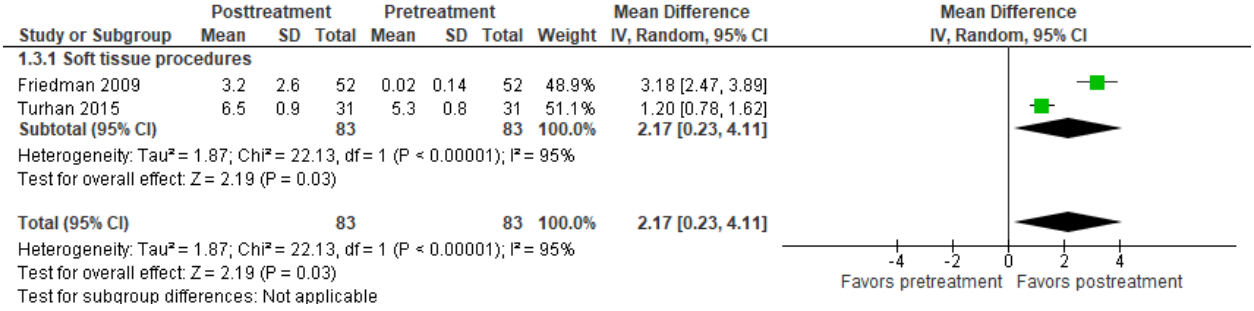


Figure S54. OR-based adjunctive surgery pretreatment vs. posttreatment (AHI)

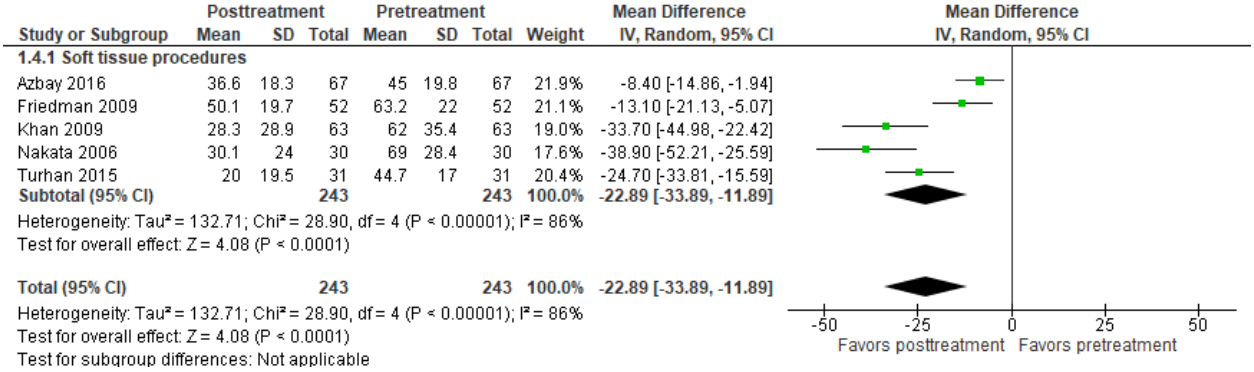


Figure S55. OR-based adjunctive surgery pretreatment vs. posttreatment (LSAT)

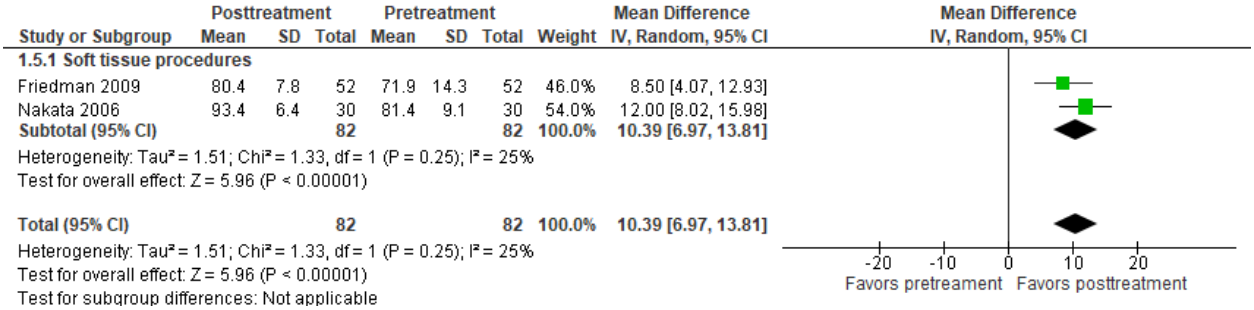


Table S3. Summary of findings table for the surgical treatment of obstructive sleep apnea in adults to facilitate CPAP therapy

References: Azbay 2016 (1); Bican 2010 (2); Friedman 2009 (3); Khan 2009 (4); Nakata 2006 (5); Turhan 2015 (6); Zonato 2006 (7)

Outcomes	Quality of the evidence (GRADE)	Anticipated absolute effects* (95% CI) MD between pre-surgery and post-surgery or surgery and control	№ of participants (studies)
ESS* (Pre- vs. Post-surgery)	⊕⊕◯◯ LOW	The mean difference in ESS score after surgery was 6.0 points [A] lower (7.2 lower to 4.7 lower) than before surgery.	117 (3 observational studies) [1,2,5]
Optimal CPAP Pressure* (Pre- vs Post-surgery)	⊕⊕◯◯ LOW	The mean difference in optimal CPAP pressure after surgery was 2.5 cm [A] H_2O lower (3.5 lower to 1.4 lower) than before surgery.	260 (6 observational studies) [1,3-7]
CPAP Adherence* (Pre- vs. Post-surgery)	⊕◯◯◯ VERY LOW [1,2]	The mean difference in CPAP adherence after surgery was 2.2 hrs/night [A] higher (0.2 higher to 4.11 higher) than before surgery.	83 (2 observational studies) [3,6]
AHI (Pre- vs Post-surgery)	⊕⊕◯◯ LOW	The mean difference in AHI after surgery was 22.9 events/hr [A] lower (33.9 lower to 11.9 lower) than before surgery. The mean AHI before surgery was 56.2 events/hr and after surgery was 33.3events/hr for a reduction of 41%.	243 (5 observational studies) [1,3-6]
LSAT (Pre- vs Post-surgery)	⊕◯◯◯ VERY LOW [1,2]	The mean difference in LSAT after surgery was 10.4% [A] higher (7.0 higher to 13.8 higher) than before surgery. The mean LSAT before surgery was 77.0% and after surgery was 87.4% for an increase of 13.5%	82 (2 observational studies) [3,5]

*Critical Outcomes
[1]Quality of evidence was downgraded due to imprecision associated with small sample size (<100 participants)
[2]Quality of evidence was downgraded due to imprecision (i.e., 95% CI of mean difference crosses clinical significance threshold)
[A]Mean difference meets or exceeds clinical significance threshold

Surgical treatment of adult OSA as an initial therapy (PICO 4)

Figure S56. OR-based surgery vs. control (ESS) [RCT]

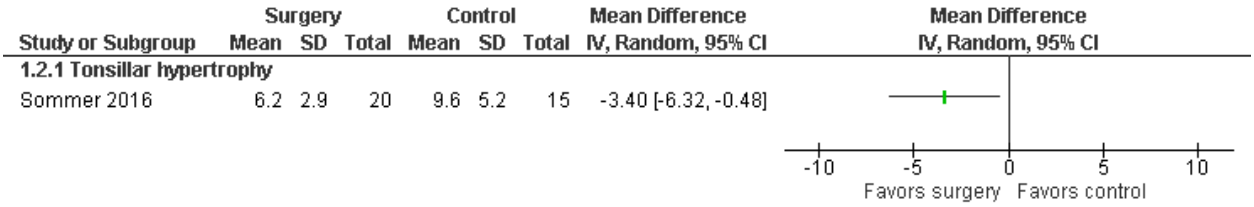


Figure S57. OR-based surgery pretreatment vs. posttreatment (ESS)

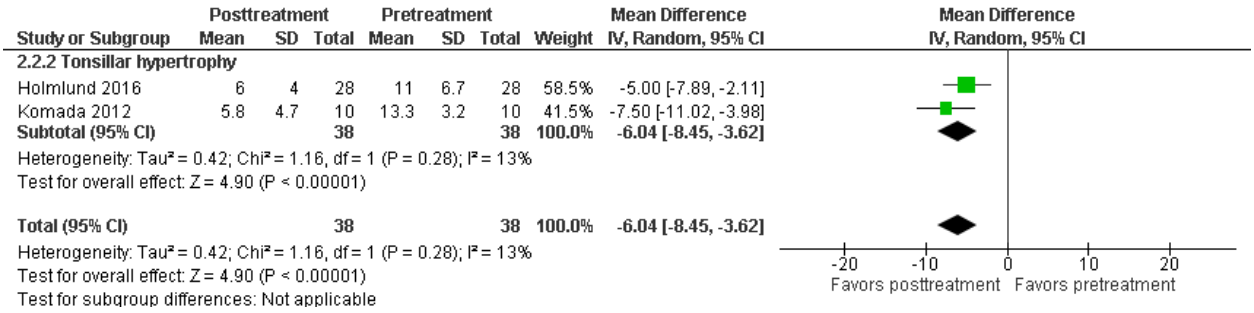


Figure S58. OR-based surgery vs. control (Snoring, VAS) [RCT]

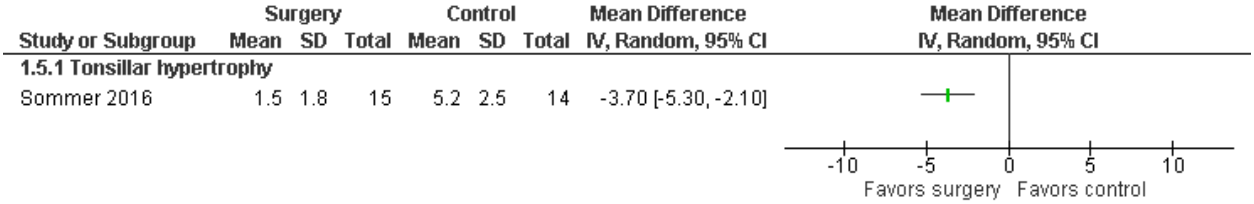


Figure S59. OR-based surgery pretreatment vs. posttreatment (Snoring, VAS)

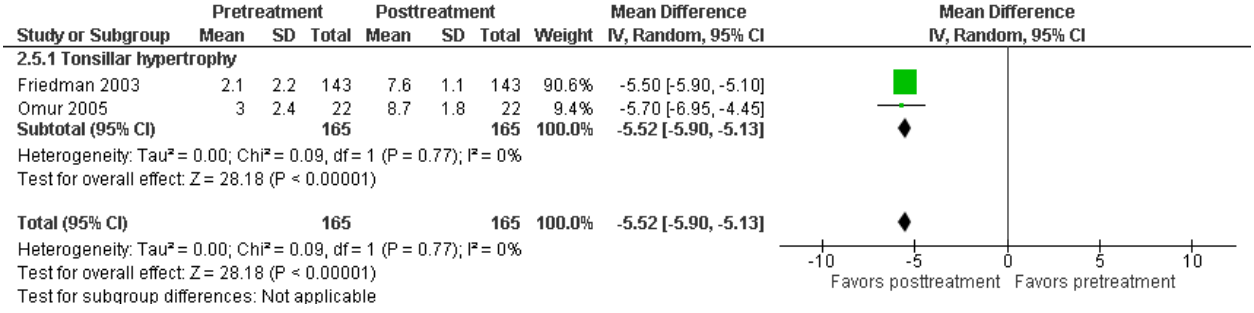


Figure S60. OR-based surgery vs. control (AHI) [RCT]

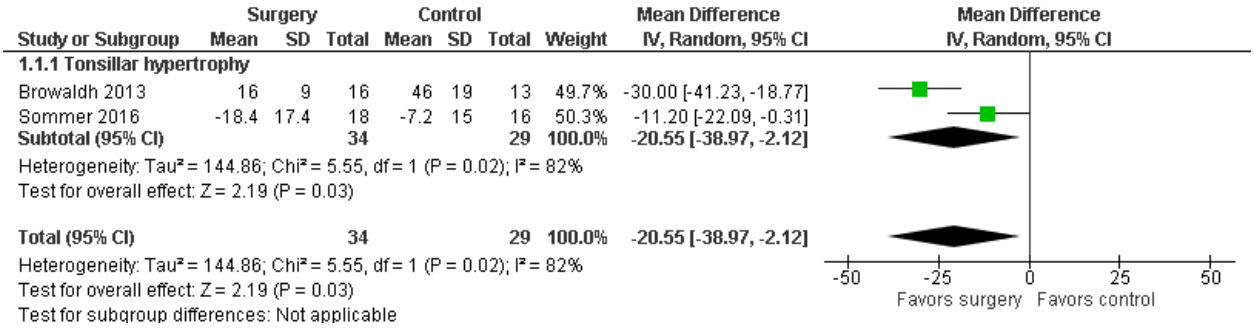


Figure S61. OR-based surgery pretreatment vs. posttreatment (AHI)

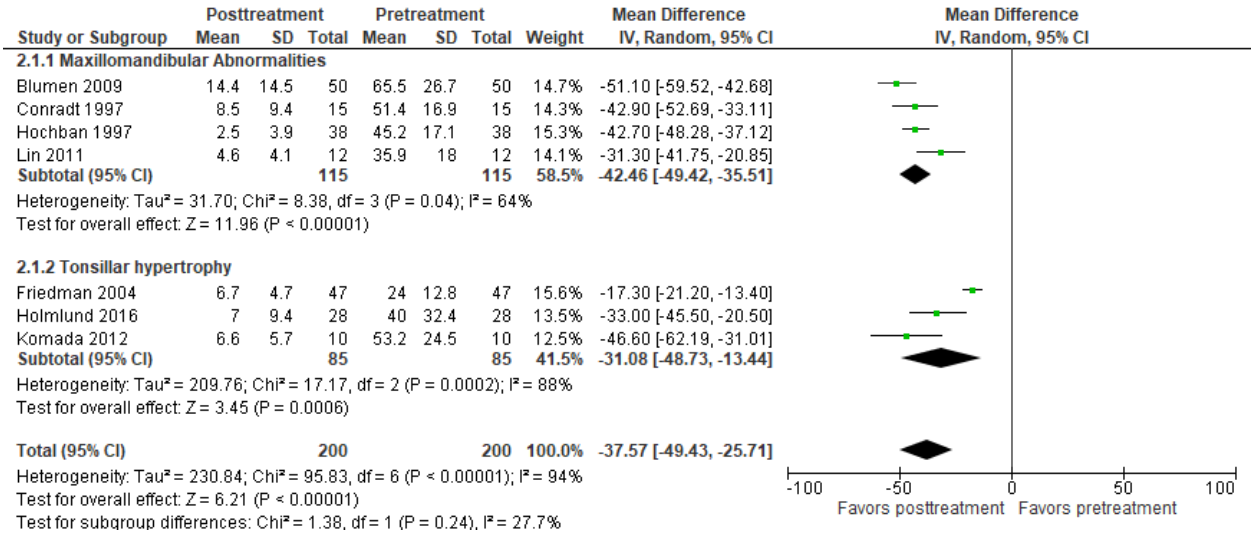


Figure S62. OR-based surgery vs. control (RDI) [RCT]

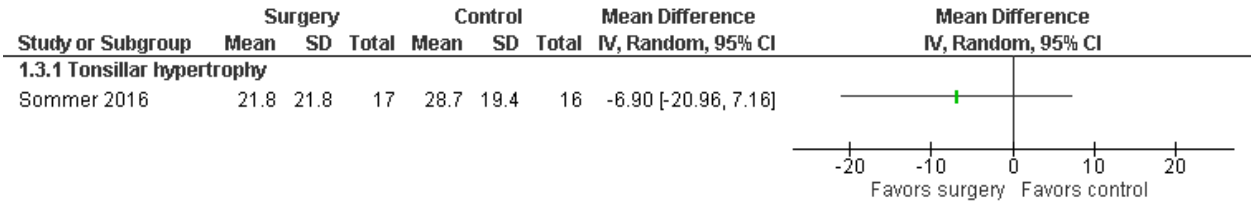


Figure S63. OR-based surgery pretreatment vs. posttreatment (RDI)

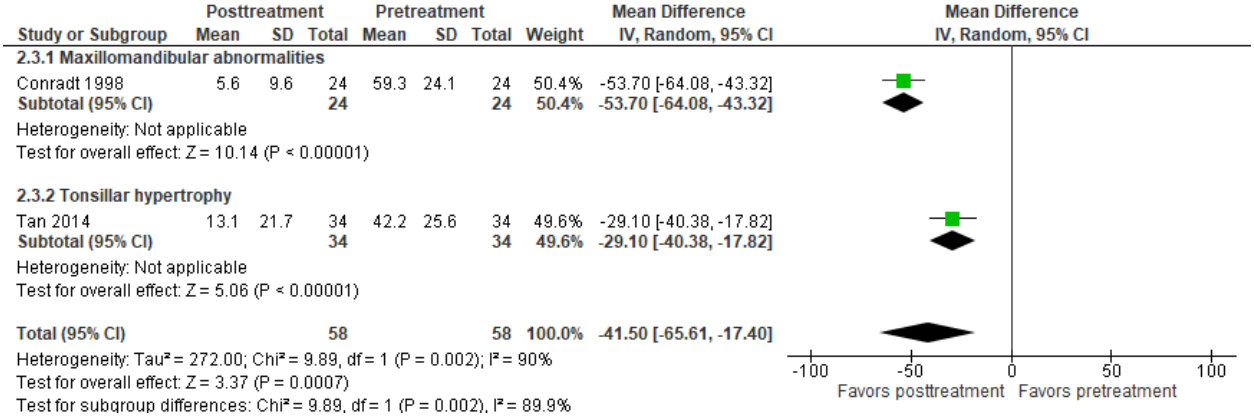


Figure S64. OR-based surgery vs. control (LSAT) [RCT]

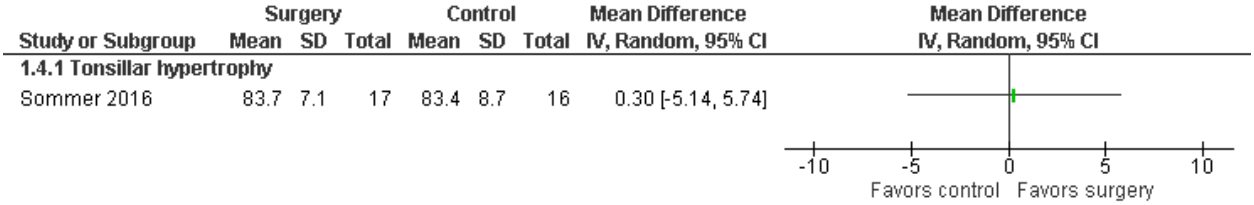


Figure S65. OR-based surgery pretreatment vs. posttreatment (LSAT)

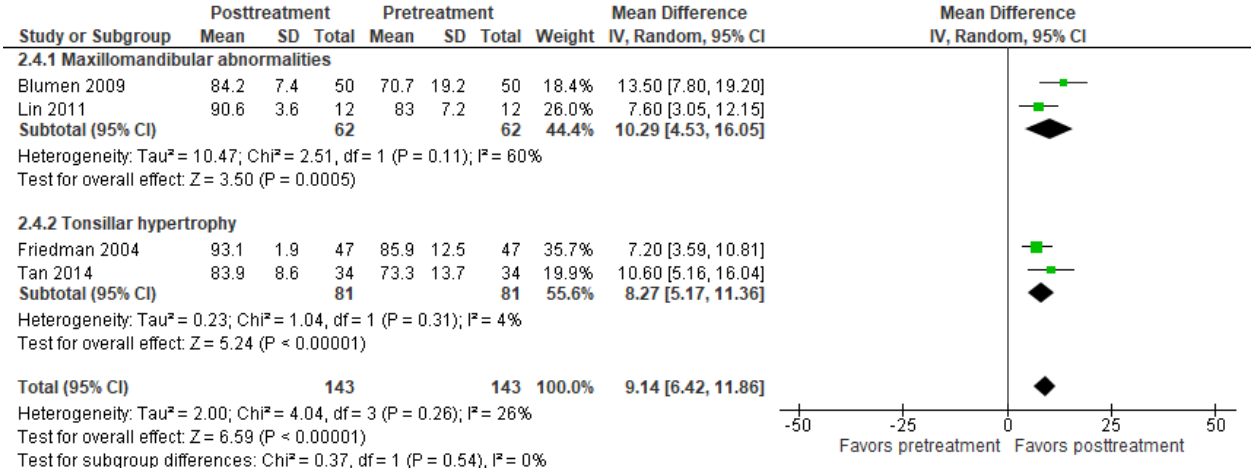


Figure S66. OR-based surgery pretreatment vs. posttreatment (ODI)

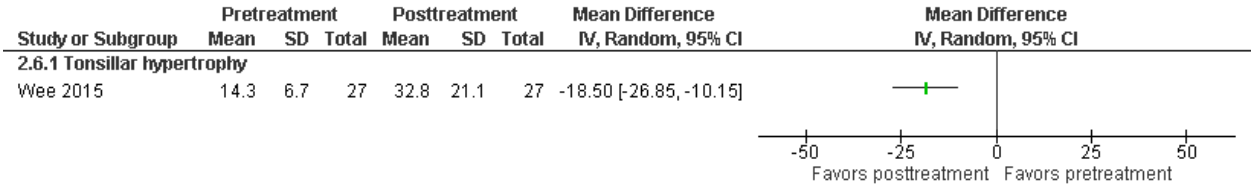


Figure S67. OR-based surgery pretreatment vs. posttreatment (persistent dysphagia, incidence)

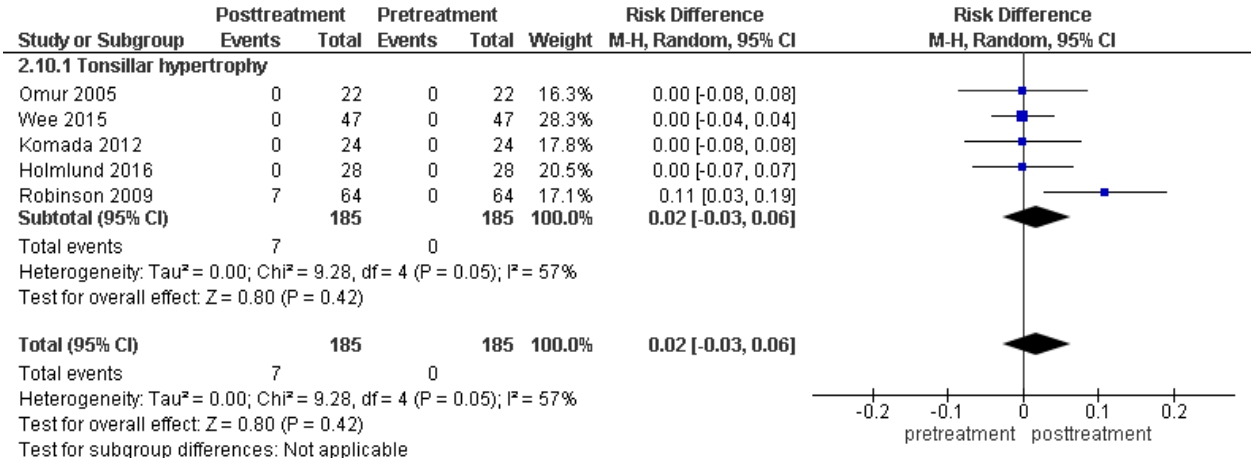


Figure S68. OR-based surgery pretreatment vs. posttreatment (mean SBP, mm Hg)

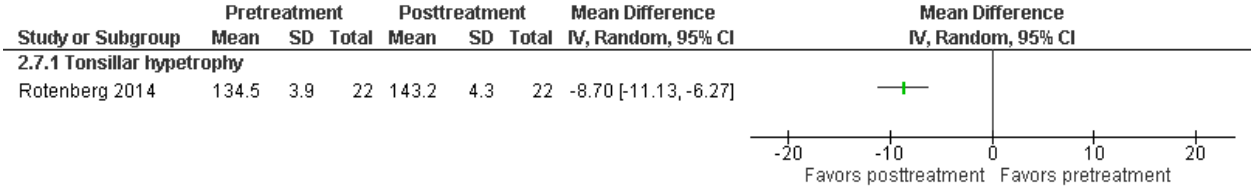


Figure S69. OR-based surgery pretreatment vs. posttreatment (persistent taste alteration, incidence)

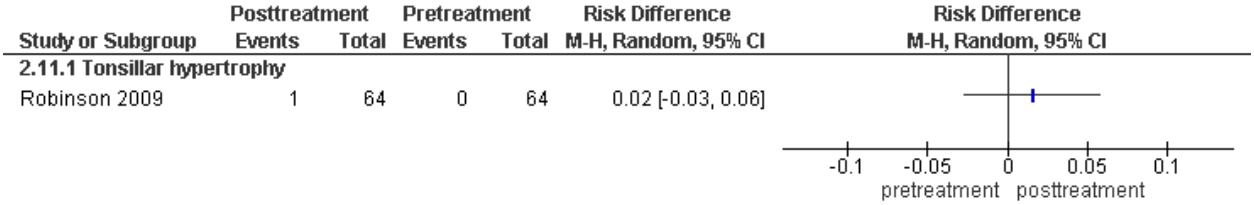


Figure S70. OR-based surgery pretreatment vs. posttreatment (persistent mandibular paresthesia, incidence)

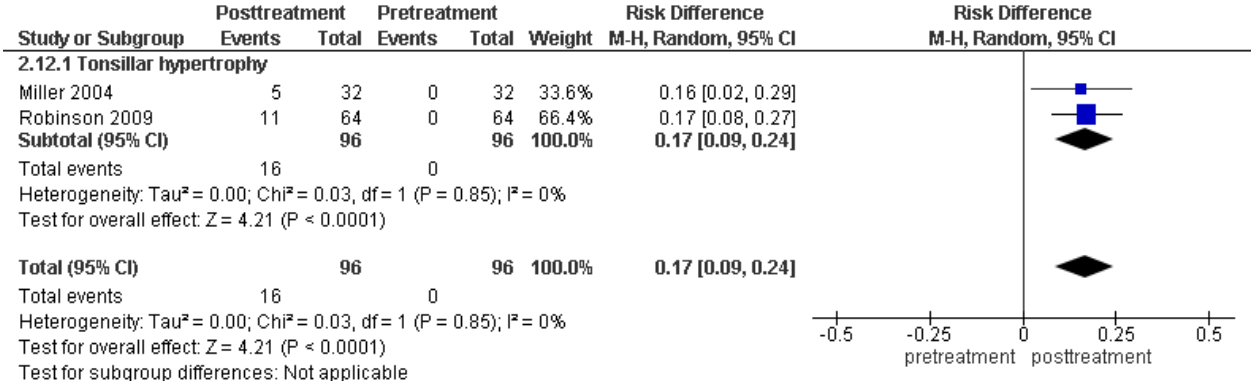


Figure S71. OR-based surgery pretreatment vs. posttreatment (persistent aspiration, incidence)

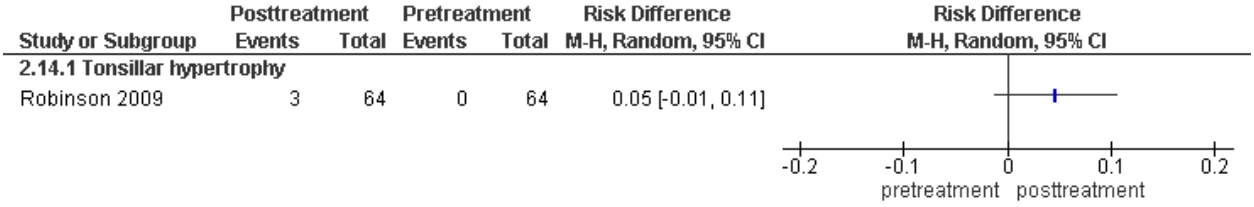


Table S4. Summary of findings table for the surgical treatment of obstructive sleep apnea in adults as an initial therapy

References: Blumen 2009 (1); Browaldh 2013 (2); Conradt 1997 (3); Conradt 1998 (4); Friedman 2003 (5); Friedman 2004 (6); Hochban 1997 (7); Holmlund 2016 (8); Komada 2012 (9); Lin 2011 (10); Miller 2004 (11); Omur 2005 (12); Robinson 2009 (13); Rotenberg 2014 (14); Sommer 2016 (15); Tan 2014 (16); Wee 2015 (17)

Outcomes	Quality of the evidence (GRADE)	Anticipated absolute effects* (95% CI) MD between pre-surgery and post-surgery or surgery and control	№ of participants (studies)
ESS* (Surgery vs Control)	⊕⊕⊕◯ MODERATE [1,2]	The mean ESS in the surgery group was 3.4 points [A] lower (6.3 lower to 0.5 lower) than in the control group.	35 (1 RCT) [15]
ESS* (Pre- vs Post-surgery)	⊕◯◯◯ VERY LOW [1]	The mean ESS after tonsillectomy was 6.0 points [A] lower (8.4 lower to 3.6 lower) than before surgery.	38 (2 observational studies) [8,9]
Snoring VAS* (Surgery vs Control)	⊕⊕⊕◯ MODERATE [1,2]	The mean snoring VAS in the surgery group was 3.7 points [A] lower (5.3 lower to 2.1 lower) than in the control group.	29 (1 RCT) [15]
Snoring VAS* (Pre- vs Post-surgery)	⊕⊕◯◯ LOW	The mean snoring VAS after surgery was 5.5 points [A] lower (5.9 lower to 5.1 lower) than before surgery.	165 (2 observational studies) [5,14]
AHI* (Surgery vs Control)	⊕⊕⊕◯ MODERATE [1,2]	The mean AHI in the tonsillectomy group was 20.6 events/hr [A] lower (39.0 lower to 2.1 lower) than in the control group.	63 (2 RCTs) [2,15]
AHI* (Pre- vs Post-surgery)	⊕⊕⊕◯ MODERATE [3]	The mean AHI after surgery was 37.6 events/hr [A] lower (49.4 lower to 25.7 lower) than before surgery. The mean AHI before surgery was 44.8 events/hr and after surgery was 7.2 events/hr for a reduction of 84%. The mean AHI after surgery for patients with maxillomandibular abnormalities was 42.5 events/hr lower (49.4 lower to 35.5 lower) than before surgery. The mean AHI was reduced by 85% with craniofacial procedures. The mean AHI after surgery for patients with tonsillar hypertrophy was 31.1 events/hr lower (48.7 lower to 13.4 lower) than before surgery. The mean AHI was reduced by 82% with tonsillectomy.	200 (7 observational studies) [1,3,5-9]
RDI* (Surgery vs Control)	⊕⊕⊕◯ MODERATE [1]	The mean RDI in the surgery group was 6.9 events/hr [A] lower (21.0 lower to 7.2 lower) than in the control group.	33 (1 RCT) [15]
RDI* (Pre- vs Post-surgery)	⊕◯◯◯ VERY LOW [1]	The mean RDI after surgery was 41.5 events/hr [A] lower (65.6 lower to 17.4 lower) than before surgery. The mean RDI before surgery was 50.8 events/hr and was 9.3 events/hr after surgery for a reduction of 82%. The mean RDI after surgery for patients with maxillomandibular abnormalities was 53.7 events/hr lower (64.1 lower to 43.3 lower) than before surgery for a reduction of 91%. The mean RDI after surgery for patients with tonsillar hypertrophy was 29.1 events/hr lower (40.4 lower to 17.8 lower) than before surgery for a reduction of 69%.	58 (2 observational studies) [4,16]
LSAT* (Surgery vs Control)	⊕◯◯◯ VERY LOW [1,4]	The mean LSAT in the surgery group was 0.3 % [B] higher (5.1 lower to 5.7 higher) than in the control group.	33 (1 RCT) [15]
LSAT* (Pre- vs Post-surgery)	⊕⊕◯◯ LOW	The mean LSAT after surgery was 9.1% [A] higher 6.4 higher to 11.9 higher) than before surgery for an increase of 16%. The mean LSAT after surgery for patients with maxillomandibular abnormalities was 10.3% higher (4.5 higher to 16.0 higher) than before surgery. The mean LSAT after surgery for patients with tonsillar hypertrophy was 8.3% higher (5.2 higher to 11.4 higher) than before surgery.	143 (4 observational studies) [1,6,10,16]
ODI* (Pre- vs. Post-surgery)	⊕◯◯◯ VERY LOW [1]	The mean ODI after surgery in patients with tonsillar hypertrophy was 18.5 points [A] lower (26.8 lower to 10.2 lower) than before surgery. The mean ODI before surgery was 32.8 and was 14.3 after surgery for a reduction of 56%.	27 (1 observational study) [17]

Permanent dysphagia* (Pre- vs Post-surgery)	⊕◯◯◯ VERY LOW [2]	The risk difference in permanent dysphagia after surgery was 0.02 [B] higher (0.03 lower to 0.06 higher)	185 (5 observational studies) [8,9,12,13,17]
SBP (Pre- vs Post-surgery)	⊕◯◯◯ VERY LOW [1]	The mean SBP after surgery for patients with tonsillar hypertrophy was 8.7 mmHg [A] lower (11.1 lower to 6.3 lower) than before surgery.	22 (1 observational study) [14]

*Critical Outcomes
[1]Quality of evidence was downgraded due to imprecision associated with small sample size (<100 participants)
[2]Quality of evidence was downgraded due to imprecision (i.e., 95% CI of mean difference crosses clinical significance threshold)
[3]Quality of evidence was upgraded due to large effect
[4]Quality of evidence was downgraded due to imprecision (i.e., 95% CI of mean difference crosses clinical significance threshold in both directions)
[A]Mean difference meets or exceeds clinical significance threshold
[B]Mean difference does not meet clinical significance threshold