National Institute for Health and Care Excellence

Final

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

Evidence review I: Oxygen therapy

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 All NICE guidance is subject to regular review and may be updated or withdrawn.

Copyright

ISBN: 978-1-4731-4229-9

Contents

1 Oxygen therapy

1.1 Review question: What is the clinical and cost effectiveness of oxygen therapy adjunctive to ventilatory support for people who do not fulfil long term oxygen therapy (LTOT) criteria for managing obstructive sleep apnoea/hypopnoea syndrome (OSAHS), obesity hypoventilation syndrome (OHS) and COPD-OSAHS overlap syndrome?

1.2 Review Question: What is the clinical and cost effectiveness of oxygen therapy (alone) for managing obstructive sleep apnoea/hypopnoea syndrome (OSAHS), obesity hypoventilation syndrome (OHS) and COPD-OSAHS overlap syndrome?

1.3 Introduction

Oxygen therapy is prescribed for patients with persistent hypoxaemia to increase alveolar oxygen tension and decrease the work of breathing. The concentration of oxygen used to treat hypoxaemia is adjusted to achieve normal or near normal oxygen saturations (94-96%), except in a group of patients with chronic hypercapnia in whom a lower target saturation of 88-92% is required. Patients with obstructive sleep apnoea/hypopnea syndrome, obesity hypoventilation and COPD-OSAHS overlap syndrome all experience varying levels of hypoxaemia. The extent of the hypoxaemia depends on the underlying pathology and patients may have obstructive sleep apnoea/hypopnoea syndrome in combination with chronic obstructive pulmonary disease, or obesity hypoventilation.

Although it is known that peple with OSAHS and OHS suffer hypoxaemia as a consequence of their disease, primary treatment is through the use of CPAP or NIV therapy. The use of ventilatory support reduces periods of apnoea and the work of breathing. However, some patients have persistent hypoxaemia despite the use of ventilatory support. Oxygen therapy is not commonly used in these patients unless they also meet the criteria for long term oxygen therapy (LTOT) as a result of pre-existing chronic respiratory failure with hypoxaemia during wakefulness. The aim of this review question was to determine the clinical and cost effectiveness of oxygen therapy on its own or as an adjunctive to ventilatory support for patients in whom LTOT is not otherwise indicated.

1.4 PICO table

For full details see the review protocol in appendix A.

Table 1: PICO characteristics of review question oxygen therapy adjunctive to ventilatory support

Population	People (16 and older) with OSAHS, OHS or COPD-OSAHS overlap syndrome and nocturnal hypoxaemia despite optimised ventilatory support who do not already fulfil criteria for long term oxygen therapy (LTOT)
Intervention	Oxygen therapy + CPAP or NIV (non-invasive ventilation)
Comparison	• CPAP or NIV (non-invasive ventilation) without oxygen therapy • any other OSAHS/OHS/COPD-OSAHS overlap syndrome treatment • no treatment/sham treatment

Outcomes	**Critical**
	• generic or disease specific quality of life measures (continuous)
	• mortality (dichotomous)
	
	Important
	• sleepiness scores (continuous, e.g. Epworth)
	• apnoea-Hypopnoea index or respiratory disturbance index (continuous)
	• oxygen desaturation index (continuous)
	• daytime pO2 (continuous)
	• daytime pCO2 (continuous)
	• daytime bicarbonate (continuous)
	• nocturnal transcutaneous CO_2 control (continuous)
	• nocturnal oximetry (continuous)
	• minor adverse effects of treatment (rates or dichotomous)
	• adherence (continuous)
	• driving outcomes (continuous)
	• neurocognitive outcomes (continuous)
	• pulmonary artery pressure by Transthoracic echocardiography (TTE) (continuous)
	• patient preference (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 only • minimum duration of follow-up 1 months • parallel or crossover to be included

Table 2: PICO characteristics of review question oxygen therapy alone

Population	People (16 and older) with OSAHS, OHS or COPD-OSAHS overlap syndrome
Intervention	Oxygen therapy
Comparison	• CPAP • non-invasive ventilation (NIV) • no treatment/sham treatment Comparison for people with OSAHS:

	• oxygen vs CPAP • oxygen vs no treatment Comparison for people with OHS: • oxygen vs NIV (non-invasive ventilation) • oxygen vs no treatment Comparison for people with COPD-OSAHS overlap syndrome: • oxygen vs NIV (non-invasive ventilation) • oxygen vs no treatment

Outcomes	**Critical**
	• generic or disease specific quality of life measures (continuous)
	• mortality (dichotomous)
	
	Important
	• sleepiness scores (continuous, e.g. Epworth)
	• apnoea-Hypopnoea index or respiratory disturbance index (continuous)
	• oxygen desaturation index (continuous)
	• daytime pO2 (continuous)
	• daytime pCO2 (continuous)
	• daytime bicarbonate (continuous)
	• nocturnal transcutaneous CO_2 control (continuous)
	• nocturnal oximetry (continuous)
	• minor adverse effects of treatment (rates or dichotomous)
	• adherence (continuous)
	• driving outcomes (continuous)
	• neurocognitive outcomes (continuous)
	• pulmonary artery pressure by Transthoracic echocardiography (TTE) (continuous)
	• patient preference (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 only • minimum duration of follow-up 1 months • parallel or crossover to be included

1.5 Clinical evidence

1.5.1 Included studies

Oxygen therapy (with CPAP/non-invasive ventilation)

No relevant clinical studies were identified.

Oxygen therapy (alone)

OSAHS

Three papers from two studies[10, 21, 42] were included in this review. These are summarised in Table 3 below.

One study (two papers) was an RCT parallel design and one study was an RCT crossover design. All papers included two comparisons; comparing oxygen therapy alone with CPAP and comparing oxygen therapy alone with a placebo or no treatment.

All studies included a moderate severity population of OSAHS (based on mean AHI). Studies were stratified based on the AHI/ODI severity of the population into mild, moderate or severe OSAHS. When a mixed severity population was included the severity of the majority of the population was used. This was calculated by taking the mean baseline AHI/ODI of the patients included and the study was downgraded for indirectness.

OHS

There was no evidence for people with OHS.

COPD-OSAHS overlap syndrome

There was no evidence for people with COPD-OSAHS overlap syndrome

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 H.

1.5.2 Excluded studies

See the excluded studies list in appendix I.

1.5.3 Summary of clinical studies included in the evidence review

Table 3: Summary of studies included in the evidence review for oxygen therapy alone

Study	Intervention and comparison	Population	Outcomes	Comments
Lewis 2017[21] and Gottlieb 2014[10] RCT USA	N= 106 **Oxygen therapy** - in addition to health and lifestyle advice patients in the oxygen therapy group received a stationary oxygen concentrator (EverFlo, Philips Respironics) which was used to provide nightly treatment with oxygen at a rate of 2 litres per minute through a nasal cannula. **CPAP** – In addition to health and lifestyle advice patients were in the CPAP group received a CPAP device with automatic adjustment (REMstar Auto CPAP, Philips Respironics), set at a pressure range of 4 to 20 cm of water for 7 days and then reset to the best fixed pressure for each patient (defined as the 90th percentile of the pressure distribution generated through automatic adjustment during those 7 days). N= 106	Patients were recruited from four participating cardiology practices in the USA. Inclusion criteria were; Age 45-75, Berlin Questionnaire score1 of 2 or 3, established coronary heart disease (prior myocardial infarction or coronary artery revascularization or angiographically documented >70% stenosis of a major coronary artery). Or 3 or more of the following established cardiovascular risk factors: hypertension (SBP >140 or DBP >90 or use of antihypertensive medication), diabetes mellitus, obesity (BMI over 30 kg/m^2), dyslipidaemia (total cholesterol >240 mg/dl, LDL >160 mg/dl, HDL <45 mg/dl, or taking lipid-lowering medication). Baseline AHI: Nasal O2 group - 24.0 (8.1) CPAP group – 25.4 (8.7)	SF-36 Physical SF-36 Mental SF-36 Vitality PHQ – 9 (depression score) ODI change score Mortality Daytime mean systolic BP 24-hour mean systolic BP Cardiovascular events - myocardial infarction, unstable angina, stroke, percutaneous coronary intervention for worsening angina, tachyarrhythmia Atrial Fibrillation Motor vehicle accidents Follow up – 12 weeks	Moderate OSAHS All patients had cardiovascular disease or multiple cardiovascular risk factors

Study	Intervention and comparison	Population	Outcomes	Comments
	No treatment/health and lifestyle advice – The no treatment group received education in healthy sleep habits included suggestions for maintaining a regular sleep schedule, avoiding alcohol near bedtime, and maintaining sleep duration of 7-8 hours per night. Subjects were also provided with education on a heart-healthy lifestyle, including recommendations for weight loss (for overweight and obese subjects), healthy diet, regular exercise, smoking cessation, and medication adherence. N=106	Health and lifestyle advice group – 25.5 (8.8)		
Phillips 1990[42] Crossover trial USA	N= 8 **Oxygen therapy –** Patients in the oxygen therapy group received nocturnal nasal O2 at 4 LPM nightly for one month. N= 8 **CPAP –** subjects in the CPAP group received one month of nasal CPAP. Levels of nasal CPAP were established initially in the lab and then adjusted at home by a respiratory therapist based on behaviour during	Participants were recruited from the sleep apnoea laboratory at the university of Kentucky college of medicine. Either through routine clinic referral or as part of an ongoing study of sleep disordered breathing in the elderly. Inclusion criteria; AHI of equal to or over 5 and at least one of the following; daytime hypersomnolence with a mean sleep latency of ≤10 minutes on multiple	Stamford sleepiness score AHI Systolic BP Neuropsychological outcomes: Attention (2&7 test), digit symbol, selective reminding - long term storage, selective reminding - consistent retrieval, Benton visual Retention, Rey Figure – copy, Rey Figure - immediate recall, Rey Figure - delayed recall, finger	Cross over study with no wash out period. All patients received CPAP last to avoid potential carry over affects. Moderate OSAHS Small study 8 patients

Study	Intervention and comparison	Population	Outcomes	Comments
	napping. The CPAP levels ranged between 2.5 and 12.5cm H20. N= 8 **Placebo -** subjects in the placebo arm received nasal compressed air nightly for one month.	sleep latency testing, hypertension with a mean of at least 5 measurements of either systolic blood pressure >150mm hg and/or diastolic blood pressure ≥95 mm hg, significant cardiac arrhythmias, including marked sinus arrhythmia, sinus bradycardia, frequent (>6 per hour0 premature ventricular contractions, one or more sinus pauses > 2 seconds associated with apnoea or hypoxemia, or supraventricular tachycardia. Baseline AHI – 20.5 (4.8) SEM	tapping (dominant hand), finger tapping (non-dominant hand) Follow up – 1 month	

1.5.4. Quality assessment of clinical studies included in the evidence review oxygen therapy (alone) – oxygen therapy compared to other interventions/no interventions

Table 4: Clinical evidence summary: Oxygen therapy compared to CPAP (moderate OSAHS)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CPAP	Risk difference with Oxygen therapy versus CPAP (95% CI)
Mortality	200 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	RD 0.00 (-0.02 to 0.02)	No events	No events
Atrial Fibrillation	200 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	RR 1.96 (0.18 to 21.28)	Moderate	
				10 per 1000	10 more per 1000 with oxygen (from 8 fewer to 203 more)
Cardiovascular complications – (unstable angina, myocardial infarction, percutaneous coronary intervention for worsening angina, stroke)	200 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	RD 0.00 (-0.02 to 0.02)	Moderate	
				0 per 1000	No events
Number of motor vehicle accidents	200 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3,5] due to risk of bias, indirectness, imprecision	OR 7.24 (0.14 to 365.16)	Moderate	
				0 per 1000	10 more per 1000 with oxygen (from 20 fewer to 40 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CPAP	Risk difference with Oxygen therapy versus CPAP (95% CI)
24-hour mean systolic BP	184 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean 24-hour systolic BP in the control groups was 123.4 mmHg	The 24-hour mean systolic BP in the oxygen group was 3.5 higher (0.76 lower to 7.76 higher)
SF36 - physical Scale from: 0 to 100. Higher is better	199 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean SF36 - physical in the control groups was 44.6	The mean SF36 - physical in the oxygen group was 0.5 lower (3.38 lower to 2.38 higher)
SF36 - mental Scale from: 0 to 100. Higher is better	199 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean SF36 - mental in the control groups was 52.6	The mean SF36 - mental in the oxygen group was 0.7 lower (3.49 lower to 2.09 higher)
SF36 - vitality Scale from: 0 to 100. Higher is better	200 (1 study) 12	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, indirectness		The mean SF36 - vitality in the control groups was 51.8	The mean SF36 - vitality in the oxygen group was 2 lower (4.80 lower to 0.80 higher)
PHQ-9 (depression) Scale from: 0 to 27 Lower is better	200 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, indirectness		The mean PHQ-9 (depression) in the control groups was 3.5	The mean PHQ-9 (depression) in the oxygen group was 0.7 higher (0.42 lower to 1.82 higher)
ODI change score	200 (1 study) 12 weeks	⨁⨁⊖⊖ LOW[1,2]		The mean ODI change score in the control groups was 17.2	The mean ODI change score in the oxygen group was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CPAP	Risk difference with Oxygen therapy versus CPAP (95% CI)
		due to risk of bias, indirectness			2.4 higher (1.37 lower to 6.17 higher)
Daytime mean systolic BP	204 (2 studies) 8 weeks	⨁⊖⊖⊖ VERY LOW[1,2,6] due to risk of bias, indirectness, imprecision		The mean daytime mean systolic BP in the control groups was 133.8 mmHg	The daytime mean systolic BP in the oxygen group was 1.33 higher (1.86 lower to 4.53 higher)
AHI (events/hr) Lower is better	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean AHI in the control groups was 3	The mean AHI in the oxygen group was 13.8 higher (7.28 to 20.32 higher)
Stamford Sleepiness Score Scale from: 1 to 7. Lower is better	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean Stamford sleepiness score in the control groups was 2.5	The mean Stamford sleepiness score in the oxygen group was 0 higher (0.71 lower to 0.71 higher)
Attention (2&7 test)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean attention (2&7 test) in the control groups was 269.4	The mean attention (2&7 test) in the oxygen group was 11.5 lower (52.53 lower to 29.53 higher)
Neurocognitive (digit symbol)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias,		The mean neurocognitive (digit symbol) in the control groups was 52.4	The mean neurocognitive (digit symbol) in the oxygen group was 1.6 lower (13.52 lower to 10.32 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CPAP	Risk difference with Oxygen therapy versus CPAP (95% CI)
		indirectness, imprecision			
Neurocognitive (selective reminding - long term storage)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean neurocognitive (selective reminding - long term storage) in the control groups was 92.4	The mean neurocognitive (selective reminding - long term storage) in the oxygen group was 4.8 lower (46.38 lower to 36.78 higher)
Neurocognitive (selective reminding - consistent retrieval)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean neurocognitive (selective reminding - consistent retrieval) in the control groups was 61.4	The mean neurocognitive (selective reminding - consistent retrieval) in the oxygen group was 7.8 lower (53.99 lower to 38.39 higher)
Neurocognitive - Benton visual Retention	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean neurocognitive - Benton visual retention in the control groups was 6.1	The mean neurocognitive - Benton visual retention in the oxygen group was 0.8 lower (2.73 lower to 1.13 higher)
Neurocognitive - (Rey Figure - copy)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean neurocognitive - (rey figure - copy) in the control groups was 33.1	The mean neurocognitive - (rey figure - copy) in the oxygen group was 2.5 lower (5.56 lower to 0.56 higher)
Neurocognitive - (Rey Figure - immediate recall)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias,		The mean neurocognitive - (rey figure - immediate recall) in the control groups was 24.8	The mean neurocognitive - (rey figure - immediate recall) in the oxygen group was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with CPAP	Risk difference with Oxygen therapy versus CPAP (95% CI)
		indirectness, imprecision			1.5 lower (8.02 lower to 5.02 higher)
Neurocognitive - (Rey Figure - Delayed recall)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean neurocognitive - (rey figure - delayed recall) in the control groups was 22.6	The mean neurocognitive - (rey figure - delayed recall) in the oxygen group was 2.2 lower (9.28 lower to 4.88 higher)
Neurocognitive - finger tapping (dominant hand)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean neurocognitive - finger tapping (dominant hand) in the control groups was 46.9	The mean neurocognitive - finger tapping (dominant hand) in the oxygen group was 0.9 lower (7.3 lower to 5.5 higher)
Neurocognitive - finger tapping (non-dominant hand)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean neurocognitive - finger tapping (non-dominant hand) in the control groups was 41.3	The mean neurocognitive - finger tapping (non-dominant hand) in the oxygen group was 0.3 higher (6.91 lower to 7.51 higher)

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. The majority of the evidence included an indirect population (downgrade by one increment) or a very indirect population (downgrade by two increments). Downgraded due to a mixed severity OSHAS population was included based on mean AHI.

3. Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs. MID for Systolic BP 5mmhg; Established MIDs for SF-36 physical/mental- 2/3.. GRADE default MID (0.5XSD) for all other continuous outcomes.

4. Outcomes 1-11 are from a 3-arm trial, results from the participants receiving the oxygen therapy are used in both the CPAP and placebo/no treatment comparisons

5. Peto odds ratio analysis used as there were zero events in one treatment arm.

6. GIV analysis used as cross over and parallel design RCTS combined

Table 5: Clinical evidence summary: Oxygen therapy compared to Placebo (moderate OSAHS

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with no treatment/placebo	Risk difference with Oxygen therapy versus placebo/no treatment (95% CI)
Mortality	202 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	RD 0.00 (-0.02 to 0.02)	No events	No events
Atrial fibrillation[6]	202 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision	RR 2.00 (0.18 to 21.71)	Moderate	
				10 per 1000	10 more per 1000 with oxygen (from 8 fewer to 207 more)
Cardiovascular complications – (unstable angina, myocardial infarction, percutaneous coronary intervention for worsening angina, stroke)	202 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW[1,2,3,7] due to risk of bias, indirectness, imprecision	OR 0.13 OR 0.13 (0.02 to 0.95)	Moderate	
				40 per 1000	40 fewer per 1000 with oxygen (from 80 fewer to 0 more)
Number of motor vehicle accidents	202 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW[1,2,3,7] due to risk of bias, indirectness, imprecision	OR 7.39 (0.15 to 372.38)	Moderate	
				0 per 1000	10 more per 1000 with oxygen (from 20 fewer to 40 more)
24-hour mean systolic BP	191 (1 study) 12 weeks	⊕⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The 24-hour mean systolic BP in the control groups was 124.7	The 24-hour mean systolic BP in the oxygen group was 2.2 higher (2.47 lower to 6.87 higher)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with no treatment/placebo	Risk difference with Oxygen therapy versus placebo/no treatment (95% CI)
SF36 physical Scale from: 0 to 100 Higher is better	199 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean SF36 physical in the control groups was 42.9	The mean SF36 physical in the oxygen group was 1.2 higher (1.56 lower to 3.96 higher)
SF36 - mental Scale from: 0 to 100. Higher is better	199 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean SF36 - mental in the control groups was 49.7	The mean SF36 - mental in the oxygen group was 2.2 higher (0.73 lower to 5.13 higher)
SF36 - vitality Scale from: 0 to 100. Higher is better	201 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW[1,2] due to risk of bias, indirectness		The mean SF36 - vitality in the control groups was 49.5	The mean SF36 - vitality in the oxygen group was 0.3 higher (2.24 lower to 2.84 higher)
PHQ-9 (depression) Scale from: 0 to 27. Lower is better	201 (1 study) 12 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean PHQ-9 (depression) in the control groups was 5.1	The mean PHQ-9 (depression) in the oxygen group was 0.9 lower (2.01 lower to 0.21 higher)
ODI change score	202 (1 study) 12 weeks	⨁⨁⊖⊖ LOW[1,2] due to risk of bias, indirectness		The mean ODI change score in the control groups was 1.5	The mean ODI change score in the oxygen group was 18.1 higher (14.32 to 21.88 higher)
Daytime mean systolic BP	208 (2 studies) 8 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3,6] due to risk of bias, indirectness, imprecision		The mean daytime mean systolic BP in the control groups was 136.3 mm hg	The mean daytime mean systolic BP in the oxygen group was 0.85 lower (4.40 lower to 2.70 higher)
AHI (events/hr) Lower is better	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias,		The mean AHI in the control groups was 22.1	The mean AHI in the oxygen group was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with no treatment/placebo	Risk difference with Oxygen therapy versus placebo/no treatment (95% CI)
		indirectness, imprecision			5.3 lower (18.11 lower to 7.51 higher)
Stamford Sleepiness Score Scale from: 1 to 7 Lower is better	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean Stamford sleepiness score in the control groups was 2.9	The mean Stamford sleepiness score in the oxygen group was 0.4 lower (1.11 lower to 0.31 higher)
Attention (2&7 test)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean attention (2&7 test) in the control groups was 254.1	The mean attention (2&7 test) in the oxygen group was 3.8 higher (36.96 lower to 44.56 higher)
Neurocognitive (digit symbol)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean neurocognitive (digit symbol) in the control groups was 49.8	The mean neurocognitive (digit symbol) in the oxygen group was 1 higher (10.79 lower to 12.79 higher)
Neurocognitive (selective reminding - long term storage)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean neurocognitive (selective reminding - long term storage) in the control groups was 81.4	The mean neurocognitive (selective reminding - long term storage) in the oxygen group was 6.2 higher (35.66 lower to 48.06 higher)
Neurocognitive (selective reminding - consistent retrieval)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean neurocognitive (selective reminding - consistent retrieval) in the control groups was 46.6	The mean neurocognitive (selective reminding - consistent retrieval) in the oxygen group was 7 higher (35.89 lower to 49.89 higher)
Neurocognitive - Benton visual Retention	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias,		The mean neurocognitive - benton visual retention in the control groups was 5.6	The mean neurocognitive - benton visual retention in the oxygen group was

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with no treatment/placebo	Risk difference with Oxygen therapy versus placebo/no treatment (95% CI)
		indirectness, imprecision			0.3 lower (2.23 lower to 1.63 higher)
neurocognitive - (Rey Figure - copy)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean neurocognitive - (rey figure - copy) in the control groups was 32.3	The mean neurocognitive - (rey figure - copy) in the oxygen group was 1.7 lower (5.31 lower to 1.91 higher)
Neurocognitive - (Rey Figure - Immediate recall)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean neurocognitive - (rey figure - immediate recall) in the control groups was 20.1	The mean neurocognitive - (rey figure - immediate recall) in the oxygen group was 3.2 higher (4.48 lower to 10.88 higher)
Neurocognitive - (Rey Figure - Delayed recall)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean neurocognitive - (rey figure - delayed recall) in the control groups was 19.4	The mean neurocognitive - (rey figure - delayed recall) in the oxygen group was 1 higher (6.35 lower to 8.35 higher)
Neurocognitive - finger tapping (dominant hand)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean neurocognitive - finger tapping (dominant hand) in the control groups was 45.6	The mean neurocognitive - finger tapping (dominant hand) in the oxygen group was 0.4 higher (6.53 lower to 7.33 higher)
Neurocognitive - finger tapping (non-dominant hand)	8 (1 study) 4 weeks	⨁⊖⊖⊖ VERY LOW[1,2,3] due to risk of bias, indirectness, imprecision		The mean neurocognitive - finger tapping (non-dominant hand) in the control groups was 38.7	The mean neurocognitive - finger tapping (non-dominant hand) in the oxygen group was 2.9 higher (4.45 lower to 10.25 higher)

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

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with no treatment/placebo	Risk difference with Oxygen therapy versus placebo/no treatment (95% CI)

2 Downgraded by 1 or 2 increments because: The majority of the evidence included an indirect population (mixed severity population) (downgrade by one increment) or a very indirect population (downgrade by two increments. Downgraded due to a mixed severity OSHAS population was included based on mean AHI.

3 Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs.MID for systolic BP 5mmhg. Established MIDs for SF-36 physical/mental- 2/3; GRADE default MID (0.5XSD) for all other continuous outcomes.

4 No events were reported in the control group while one event occurred in the O2 group. However, the absolute effect was not estimable on GRADE.

5. Outcomes 1-11 are from a 3-arm trial, results from the participants receiving the oxygen therapy are used in both the CPAP and placebo/no treatment comparisons

6. Outcome includes atrial fibrillation and 1 episode of unspecified tachyarrhythmia requiring hospitalisation in the placebo group.

7. Peto odds ratio analysis used as there were zero events in one treatment arm.

8. GIV analysis used as cross over and parallel design RCTS combined

1.6 Economic evidence

1.6.1 Included studies

No relevant health economic studies were identified.

1.6.2 Excluded studies

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

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

1.6.3 Health economic modelling

Original modelling was not conducted for this question.

1.6.4 Health economic evidence statements

No relevant economic evaluations were identified.

1.7 The committee's discussion of the evidence

1.7.1 Interpreting the evidence

1.7.1.1 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), Apnoea hypopnoea index (AHI), Oxygen desaturation index (ODI), minor adverse effects of treatment, daytime PO2, daytime PCO2, daytime bicarbonate, nocturnal transcutaneous CO_2 control, nocturnal oximetry, driving outcomes, neurocognitive outcomes, pulmonary artery pressure by transthoracic echocardiography (TTE), and patient preference. The committee was 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 daytime PCO2, daytime bicarbonate, nocturnal transcutaneous CO_2 control, adherence, pulmonary artery pressure by TTE, patient preference and impact on HbA1c for diabetes.

OSAHS

Oxygen therapy + CPAP or non-invasive ventilation vs CPAP or non-invasive ventilation without oxygen therapy

No relevant clinical studies were identified for this comparison.

Oxygen therapy + CPAP or non-invasive ventilation vs any other OSAHS treatment

No relevant clinical studies were identified for this comparison.

Oxygen therapy + CPAP or non-invasive ventilation vs no treatment/sham treatment

No relevant clinical studies were identified for this comparison.

Oxygen therapy vs non-invasive ventilation

No relevant clinical studies were identified for this comparison.

Oxygen therapy (alone) vs CPAP or no treatment

There was evidence from two studies (three papers); both studies had three treatment arms and compared oxygen therapy with CPAP and oxygen therapy with a placebo or no treatment.

One study was an RCT parallel design and one study was a cross over trial with no wash out period which was downgraded for risk of bias. In all studies nocturnal oxygen therapy was administered nightly via nasal cannulae at a rate of 2-4 LPM for between 4-12 weeks duration.

The populations recruited to the studies were predominately male with a diagnosis of OSAHS. One study recruited patients from cardiology practices and only included people with cardiovascular disease or multiple cardiovascular risk factors. The committee considered the applicability of this cohort to the general OSAHS population when making recommendations based on this evidence. At baseline the majority of the study populations had high BMIs and ESS scores under the arbitrary upper limit of normal of 9.

All studies included a mixed OSAHS severity population based on AHI scores. When a mixed severity population was included (i.e. mild and moderate severity OSAHS), the severity of the majority of the population was determined from the mean baseline value and the study was downgraded for indirectness. Based on mean AHI values, both studies were in moderate severity population.

The quality of the evidence varied from low to very low quality; the majority of the evidence was downgraded due to due to risk of bias, indirectness and imprecision. Risk of bias in all studies was due to selection and blinding bias. Indirectness was present in all studies due to the inclusion of mixed severity OSAHS populations, combining people with mild, moderate or severe OSAHS. Imprecision was also present for a number of the outcomes with confidence intervals crossing the MID thresholds. The low quality of evidence and uncertainty around the effect estimate was taken into consideration by the committee when assessing the evidence base in this review.

One study employed a crossover design with no washout period and therefore all outcomes were downgraded for risk of bias and ultimately all outcomes from this study were graded very low quality. The committee however agreed that this study should be included as oxygen therapy should not have any carry over effects and CPAP was administered last to all patients.

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.

OHS

There was no evidence available for people with OHS.

COPD-OSAHS overlap syndrome

There was no evidence available for people with COPD-OSAHS overlap syndrome.

1.7.1.2 Benefits and harms

OSAHS

Oxygen therapy vs CPAP – moderate OSAHS

Across most of the outcomes reported by the 2 studies in this comparison there was no evidence of any clinically important difference between oxygen therapy and CPAP.

Evidence from one study reported a small benefit of CPAP on one aspect of patients QOL. This was based on improvements in the vitality component of their SF-36 score which may be explained by improved sleep quality and cerebral oxygenation, but this did not reach clinical significance.

Evidence from one small study showed a clinically important benefit of CPAP when compared to oxygen therapy for AHI score, although there was large uncertainty around the effect estimate with the confidence interval crossing both MIDs. The committee had expected this result as they did not anticipate any improvements in AHI with oxygen therapy alone, which has been found to lengthen apnoea duration.

The evidence also reported a benefit of CPAP for the neurocognitive outcome (Rey figure – copy). The committee found this unconvincing as neurocognitive outcomes would be more affected by the patients' sleepiness and vigilance levels and the results would therefore only be valid if the analysis adjusted for baseline sleepiness scores between patients which the study failed to do. The committee also acknowledged the uncertainty around the effect estimate.

The evidence showed a small improvement in ODI change score with oxygen therapy but this change did not meet the threshold of clinically important difference.

The evidence from both studies in this comparison suggested that there were no clinically important differences between oxygen therapy and CPAP for the following outcomes; mortality, adverse events, number of motor vehicle accidents, systolic BP, ODI change scores, Stamford sleepiness scores, SF-36 mental and physical components and the PHQ-9 depression score. There were also no clinically important differences for the remaining nine neurocognitive outcomes reported by one study.

The committee acknowledged the low quality of the evidence, small number of studies and small patient populations. The committee therefore agreed that CPAP is the treatment of choice in patients with moderate OSAHS and noted there was no evidence to suggest oxygen therapy should be recommended as an alternative to CPAP in this population.

Oxygen therapy versus placebo/no treatment - moderate OSAHS

Across most of the outcomes reported by the 2 studies included in this comparison there was no evidence of any clinically important difference between oxygen therapy and a placebo/no treatment.

Evidence from one small study reported a clinically important benefit of oxygen in the ODI change score. The committee however agreed that this finding does not demonstrate any therapeutic benefits of oxygen therapy and merely proves that patients were being given oxygen at that time.

One study reported a reduction in the number of cardiovascular events (including: unstable angina, myocardial infarction, percutaneous coronary intervention for worsening angina and stroke) with oxygen therapy when compared to no treatment. However, although there were no established MIDs for this outcome the committee agreed that the difference was not enough to be of clinical significance, and they also acknowledged the uncertainty around the effect estimate. The committee also noted that this evidence came from a specific cohort of cardiovascular patients so its applicability to a general OSAHS population is limited and should be interpreted with caution.

The evidence showed that there were no clinically important differences between oxygen therapy and placebo/no treatment for the following outcomes: mortality, AHI, number of

motor vehicle accidents, systolic BP, Stamford sleepiness scores, QOL outcomes and neurocognitive outcomes.

Overall: oxygen therapy for OSAHS

There was no evidence for oxygen therapy as an adjunct to ventilatory support for people with OSAHS. There was a lack of convincing evidence in favour of oxygen therapy alone for people with moderate OSAHS and no evidence for people with mild and severe OSAHS.There was no overriding consensus within the committee to make a recommendation for oxygen therapy in people with OSAHS. They agreed that a research recommendation, specifically looking at the clinical effectiveness of oxygen therapy compared to a placebo in a CPAP intolerant population, would help to inform future guidance. The committee therefore used consensus opinion to make a research recommendation specifically for this CPAP intolerant OSAHS population as they agreed from their experience that this was a difficult group to manage and further trials on this population could identify if oxygen therapy would be effective.

OHS

There was no evidence for oxygen therapy in OHS.

Based on their experience and current practice, the committee agreed that, whilst optimal CPAP or NIV will usually be sufficient to correct ventilatory failure, some people with OHS may remain hypoxaemic while asleep despite control of AHI and nocturnal hypercapnia.This would be shown on oximetry measures or on arterial blood gas during sleep. Addition of supplemental oxygen therapy to the CPAP or non-invasive ventilation during sleep may be needed to correct this hypoxia and any additional underlying causes of hypoxaemia should be addressed where possible. Usually only a low flow rate such as 1-2L/minute would be required. Repeating oximetry or arterial blood gas would allow the response to this oxygen therapy to be evaluated and any further adjustments to oxygen prescription to be made.

The committee agreed that recommendations on oxygen therapy reflect current practice in most NHS centres, so there is likely to be little impact on practice.

COPD-OSAHS overlap syndrome

The committee noted the lack of evidence for oxygen therapy in COPD-OSAHS overlap syndrome and decided to make consensus recommendations based on experience and current practice.

Some patients will be established users of long-term oxygen therapy because of their COPD, in which case supplemental oxygen can be can be given via CPAP or non-invasive ventilation whilst sleeping, with oxygen flow rate and non-invasive ventilation or CPAP settings titrated during respiratory polygraphy according to individual need. People with COPD-OSAHS are subject to greater falls in oxygen saturation while sleeping than those with COPD alone, and the committee therefore agreed that in people with COPD-OSAHS overlap syndrome who do not fulfil criteria for long term oxygen therapy, supplemental oxygen therapy may be required in those who remain hypoxaemic when asleep despite control of AHI and nocturnal hypercapnia with CPAP or non-invasive ventilation. This would be shown on overnight oximetry measures or on arterial blood gas measurement. Therefore, addition of supplemental oxygen therapy to the CPAP or non-invasive ventilation overnight may be needed to correct this hypoxia and any additional underlying causes of hypoxaemia should be addressed where possible. Usually only low oxygen concentrations such as 1-2L/minute are required. Repeating oximetry or arterial blood gas would allow the response to this oxygen therapy to be evaluated and any further adjustments to oxygen prescription to be made.

The committee noted that the recommendations reflect current practice in most NHS centres, so there is likely to be little impact on practice.

1.7.3 Cost effectiveness and resource use

There were no economic evaluations available for this question.

There was a consensus within the committee that further research was required to establish both the clinical and cost-effectiveness of oxygen therapy for people with OSAHS.

The committee agreed that in people with OHS and COPD-OSAHS overlap syndrome who do not fulfil requirements for long term oxygen therapy, supplemental oxygen therapy would be cost effective in those who remain hypoxaemic at night despite control of AHI with CPAP or NIV.This use of oxygen reflects current practice in most centres.

1.7.4 Other factors the committee took into account

Oxygen therapy may be entrained into a CPAP or NIV device if used in the treatment of patients requiring long term oxygen therapy for COPD as per the NICE guideline NG115 (https://www.nice.org.uk/guidance/ng115) (2018) and NICE quality standard QS10 (https://www.nice.org.uk/guidance/QS10) (2011, updated 2016).

References

1. Bardwell WA, Norman D, Ancoli-Israel S, Loredo JS, Lowery A, Lim W et al. Effects of 2-week nocturnal oxygen supplementation and continuous positive airway pressure treatment on psychological symptoms in patients with obstructive sleep apnea: A randomized placebo-controlled study. Behavioral Sleep Medicine. 2007; 5(1):21-38

2. Bordier P, Lataste A, Hofmann P, Robert F, Bourenane G. Nocturnal oxygen therapy in patients with chronic heart failure and sleep apnea: A systematic review. Sleep Medicine. 2016; 17:149-157

3. Bordier P, Orazio S. Short- and long-term effects of nocturnal oxygen therapy on sleep apnea in chronic heart failure. Archives of Cardiovascular Diseases Supplements. 2013; 5(1):20

4. Bordier P, Orazio S, Hofmann P, Robert F, Bourenane G. Short- and long-term effects of nocturnal oxygen therapy on sleep apnea in chronic heart failure. Sleep & Breathing. 2015; 19(1):159-168

5. Frohnhofen H, Orth G, Hagen O, Holtmann BJ. Nocturnal oxygen therapy and cognitive function in conscious elderly patients with stroke and obstructive sleep apnea (OSA): A placebo controlled study. Cerebrovascular Diseases. 1998; 8(Suppl 4):41

6. Frohnhofen H, Orth G, Hagen O, Rang P, Holtmann B. Significance of nocturnal oxygen insufflation therapy in geriatric clinical patients with obstructive sleep apnoea. Pneumologie (Stuttgart, Germany). 1998; 52(11):645

7. George J, Wang L, Nawabit R, Kaffashi F, Walia HK, Punjabi NM et al. Impact of cpap versus supplemental oxygen on cardiac electophysiological indices in obstructive sleep apnea: The Heartbeat study. Sleep. 2018; 41(Suppl 1):A168

8. Gold AR, Schwartz AR, Bleecker ER, Smith PL. The effect of chronic nocturnal oxygen administration upon sleep apnea. American Review of Respiratory Disease. 1986; 134(5):925-929

9. Gottlieb DJ, Babineau D, Bhatt DL, Blumenthal R, Lewis EF, Mehra R et al. Effect of continuous positive airway pressure and nocturnal supplemental oxygen on blood pressure in patients with obstructive sleep apnea: The heartbeat randomized clinical trial. Sleep. 2013; 36(Suppl):A119

10. Gottlieb DJ, Punjabi NM, Mehra R, Patel SR, Quan SF, Babineau DC et al. CPAP versus oxygen in obstructive sleep apnea. New England Journal of Medicine. 2014; 370(24):2276-2285

11. Hagenah G, Andreas S, Clemens C, Figulla HR, Kreuzer H. Nocturnal oxygen administration and cardiac arrhythmias during Cheyne-Stokes respiration in patients with heart failure. Zeitschrift für Kardiologie. 1996; 85(6):435-441

12. Hollier C, Harmer A, Maxwell L, Menadue C, Unger G, Willson G et al. Effects of supplemental O2 on PCO2 and ventilation in people with obesity hypoventilation syndrome. European Respiratory Journal. 2012; 40(Suppl 56):P3186

13. Hollier CA, Harmer AR, Maxwell LJ, Menadue C, Piper AJ, Willson GN. Breathing low or moderate concentration O2 worsens hypercapnia in obesity hypoventilation syndrome. Respirology. 2012; 17(Suppl 1):80

14. Hollier CA, Harmer AR, Maxwell LJ, Menadue C, Willson GN, Unger G et al. Moderate concentrations of supplemental oxygen worsen hypercapnia in obesity

hypoventilation syndrome: A randomised crossover study. Thorax. 2014; 69(4):346-353

15. Hubatsch M. Prediction of therapeutic effectiveness of nocturnal oxygen supplementation in Cheyne-Stokes breathing. European Respiratory Journal. 2014; 44(Suppl 58):P2284

16. Janssens JP, Borel JC, Pepin JL. Non-PAP treatment modalities in obesity-hypoventilation syndrome: Role of exercise, nonsurgical and surgical weight reduction, tracheostomy, respiratory stimulants, and oxygen. Sleep Medicine Clinics. 2014; 9(3):357-364

17. Javaheri S, Ahmed M, Parker TJ, Brown CR. Effects of nasal O2 on sleep-related disordered breathing in ambulatory patients with stable heart failure. Sleep. 1999; 22(8):1101-1106

18. Kasai T, Usui Y, Yoshioka T, Yanagisawa N, Takata Y, Narui K et al. Effect of flow-triggered adaptive servo-ventilation compared with continuous positive airway pressure in patients with chronic heart failure with coexisting obstructive sleep apnea and Cheyne-Stokes respiration. Circulation: Heart Failure. 2010; 3(1):140-148

19. Kempf P, Mössinger B, Müller B, Kirchheiner T, Rühle KH. Comparative studies on the effect of nasal CPAP, theophylline and oxygen in patients with sleep apnea syndrome. Pneumologie (Stuttgart, Germany). 1991; 45(Suppl 1):279-282

20. Landry SA, Joosten SA, Sands SA, White DP, Malhotra A, Wellman A et al. Response to a combination of oxygen and a hypnotic as treatment for obstructive sleep apnoea is predicted by a patient's therapeutic CPAP requirement. Respirology. 2017; 22(6):1219-1224

21. Lewis EF, Wang R, Punjabi N, Gottlieb DJ, Quan SF, Bhatt DL et al. Impact of continuous positive airway pressure and oxygen on health status in patients with coronary heart disease, cardiovascular risk factors, and obstructive sleep apnea: A Heart Biomarker Evaluation in Apnea Treatment (HEARTBEAT) analysis. American Heart Journal. 2017; 189:59-67

22. Lewis EF, Wang R, Quan S, Gottlieb D, Bhatt D, Blumenthal R et al. Impact of treatment modalities on health status in patients with obstructive sleep apnea. Journal of the American College of Cardiology. 2014; 63(12 Suppl 1):A1536

23. Loredo JS, Ancoli-Israel S, Kim EJ, Lim WJ, Dimsdale JE. Effect of continuous positive airway pressure versus supplemental oxygen on sleep quality in obstructive sleep apnea: A placebo-CPAP-controlled study. Sleep. 2006; 29(4):564-571

24. Marrone O, Bellia V, Pieri D, Salvaggio A, Bonsignore G. Acute effects of oxygen administration on transmural pulmonary artery pressure in obstructive sleep apnea. Chest. 1992; 101(4):1023-1027

25. Masa JF, Corral J, Romero A, Caballero C, Teran-Santos J, Alonso-Alvarez ML et al. The effect of supplemental oxygen in obesity hypoventilation syndrome. Journal of Clinical Sleep Medicine. 2016; 12(10):1379-1388

26. Mayos M, Hernández Plaza L, Farré A, Mota S, Sanchis J. The effect of nocturnal oxygen therapy in patients with sleep apnea syndrome and chronic airflow limitation. Archivos de Bronconeumologia. 2001; 37(2):65-68

27. Mehta V, Vasu TS, Phillips B, Chung F. Obstructive sleep apnea and oxygen therapy: A systematic review of the literature and meta-analysis. Journal of Clinical Sleep Medicine. 2013; 9(3):271-279

28. Mehta V, Wong J, Phillips B, Chung F. Oxygen therapy for OSA patients: A systematic review. Canadian Journal of Anaesthesia. 2012; 59(Suppl 1):1

29. Mills PJ, Kennedy BP, Loredo JS, Dimsdale JE, Ziegler MG. Effects of nasal continuous positive airway pressure and oxygen supplementation on norepinephrine kinetics and cardiovascular responses in obstructive sleep apnea. Journal of Applied Physiology. 2006; 100(1):343-348

30. Mostafavi A, Aliabadi L, Sadeghniyat K, Hussein Tabatabaei SA. Comparison of the efficacy of continuous positive airway pressure and oxygen therapy in increasing heart rate variability in patients with obstructive sleep apnea. Iranian Heart Journal. 2017; 18(4):34-41

31. Murase K, Ono K, Akao M, Miki S, Nohara R, Mishima M et al. The clinical effects of adaptive servo ventilation versus nocturnal oxygen therapy for sleep disordered breathing in heart failure patients: A ramdomized trial. American Journal of Respiratory and Critical Care Medicine. 2015; 191:A6109

32. Murase K, Ono K, Yoneda T, Iguchi M, Yokomatsu T, Mizoguchi T et al. Adaptive servoventilation versus oxygen therapy for sleep disordered breathing in patients with heart failure: A randomised trial. Open Heart. 2016; 3:e000366

33. 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

34. NCT. The effect of adaptive servo ventilation and oxygen therapy in central sleep apnea patients. 2010. Available from: Https://clinicaltrials.gov/show/nct01187823 Last accessed: 25/10/2019.

35. NCT. Effects of nocturnal nasal oxygen on biomarkers in sleep apnea patients with heart failure. 2011. Available from: Https://clinicaltrials.gov/show/nct01431157 Last accessed: 25/10/2019.

36. NCT. Heart biomarker evaluation in apnea treatment. 2010. Available from: https://clinicaltrials.gov/show/NCT01086800 Last accessed: 23/07/2020.

37. NCT. Oxygen versus PAP for sleep apnea in heart failure. 2013. Available from: Https://clinicaltrials.gov/show/nct01807897 Last accessed: 25/10/2019.

38. Norman D, Loredo JS, Nelesen RA, Ancoli-Israel S, Mills PJ, Ziegler MG et al. Effects of continuous positive airway pressure versus supplemental oxygen on 24-hour ambulatory blood pressure. Hypertension. 2006; 47(5):840-845

39. Pagel JF, Parnes B. Treatment of complex apnea with supplemental oxygen and pap. Sleep. 2011; 34(Suppl):A139-A140

40. Patel SR, Weng J, Babineau DC, Bhatt DL, Blumenthal R, Gottlieb DJ et al. Effect of continuous positive airway pressure (CPAP) and nocturnal supplemental oxygen (NSO) on inflammatory markers in obstructive sleep apnea: The heartbeat trial. Sleep. 2013; 36(Suppl):A121

41. Petousi N, Turnbull CD, Sen D, Kohler M, Stradling JR. Effects of overnight supplemental oxygen on morning blood pressure in a randomized controlled CPAP withdrawal trial in patients with osa. American Journal of Respiratory and Critical Care Medicine. 2018; 197:A19

42. Phillips BA, Schmitt FA, Berry DT, Lamb DG, Amin M, Cook YR. Treatment of obstructive sleep apnea. A preliminary report comparing nasal CPAP to nasal oxygen in patients with mild OSA. Chest. 1990; 98(2):325-330

43. Pokorski M, Jernajczyk U. Nocturnal oxygen enrichment in sleep apnoea. Journal of International Medical Research. 2000; 28(1):1-8

44. Prezerakos G, Vagiakis E, Karatza MH, Vasilleiou R, Tzouvelekis A, Rigopoulou E et al. OSAS patients can exhibit nocturnal IL-6 upregulation triggered by apnea related strenuous diaphragmatic contractions and oxygen desaturation. A case controlled prospective study. European Respiratory Journal. 2014; 44(Suppl 58):4449

45. Quiroga MAS, Masa JF, Corral J, Romero A, Caballero C, Teran-Santos J et al. The effect of supplemental oxygen in obesity hypoventilation syndrome. European Respiratory Journal. 2016; 48(Suppl 60):PA2389

46. Roffe C, Frohnhofen H. The use of oxygen in the therapy of obstructive sleep apnoea in older and frail patients. Somnologie. 2014; 18(Suppl 1):40-41

47. Sasayama S, Izumi T, Seino Y, Ueshima K, Asanoi H, Chf-Hot Study Group. Effects of nocturnal oxygen therapy on outcome measures in patients with chronic heart failure and cheyne-stokes respiration. Circulation Journal. 2006; 70(1):1-7

48. Scherz N. CPAP (continuous positive pressure ventilation) reduces blood pressure in comparison with nocturnal oxygen therapy. Praxis. 2014; 103(20):1211-1212

49. Selim B, Gali B, Kashyap R, Morgenthaler T, Gay P. Nasal high flow therapy versus oxygen supplementation in postoperative care of obstructive sleep apnea patients with CPAP non-compliance: A pilot randomized study. Chest. 2018; 154(4 Suppl):1133A

50. Shafazand S, Badr MS. Commentary on CPAP vs. Oxygen for treatment of OSA. Journal of Clinical Sleep Medicine. 2014; 10(11):1257-1259

51. Shigemitsu M, Nishio K, Kusuyama T, Itoh S, Konno N, Katagiri T. Nocturnal oxygen therapy prevents progress of congestive heart failure with central sleep apnea. International Journal of Cardiology. 2007; 115(3):354-360

52. Staniforth AD, Kinnear WJ, Starling R, Hetmanski DJ, Cowley AJ. Effect of oxygen on sleep quality, cognitive function and sympathetic activity in patients with chronic heart failure and Cheyne-Stokes respiration. European Heart Journal. 1998; 19(6):922-928

53. Sugimura K, Shinozaki T, Fukui S, Ogawa H, Shimokawa H. End-Tidal CO2 tension is predictive of effective nocturnal oxygen therapy in patients with chronic heart failure and central sleep apnea. Tohoku Journal of Experimental Medicine. 2016; 239(1):39-45

54. Tan L, Li T, Zhou J, Zhang Y, Ren R, Luo L et al. Effect of nocturnal oxygen treatment on obstructive sleep apnea/hypopnea syndrome in highlanders: randomized, placebo-controlled, double-blinded trial. Sleep Medicine. 2019; 64(Suppl 1):S373-S374

55. Tan M, Joosten S, Landry S, Mansfield D, Hamilton G, Edwards B. Oxygen therapy for treating patients with residual sleep apnoea following upper airway surgery. Journal of Sleep Research. 2017; 26(Suppl 1):41

56. Teramoto S, Kume H, Matsuse T, Ishii T, Miyashita A, Akishita M et al. Oxygen administration improves the serum level of nitric oxide metabolites in patients with obstructive sleep apnea syndrome. Sleep Medicine. 2003; 4(5):403-407

57. Turnbull CD, Petousi N, Sen D, Stradling JR, Kohler M. The effects of supplemental oxygen on blood pressure in obstructive sleep apnoea during CPAP withdrawal. Thorax. 2017; 72(Suppl 3):A202

58. Turnbull CD, Sen D, Kohler M, Petousi N, Stradling JR. Effect of supplemental oxygen on blood pressure in obstructive sleep apnea (SOX): A randomised, CPAP withdrawal trial. American Journal of Respiratory and Critical Care Medicine. 2019; 199(2):211-219

59. Turnbull CD, Sen D, Petousi N, Stradling JR. Effect of supplemental oxygen on the return of sleepiness during cpap withdrawal: Which derivative of the osler maintenance of wakefulness test should be used? American Journal of Respiratory and Critical Care Medicine. 2018; 197:A4397

60. Ulrich S, Keusch S, Hildenbrand F, Huber LC, Speich R, Tanner F et al. Nocturnal oxygen therapy improves exercise performance in patients with pulmonary hypertension and sleep disturbed breathing compared to acetazolamide and placebo. Randomized, double-blind, cross-over trial. Respiration. 2013; 85(6):527

61. Walsh JT, Andrews R, Starling R, Cowley AJ, Johnston ID, Kinnear WJ. Effects of captopril and oxygen on sleep apnoea in patients with mild to moderate congestive cardiac failure. British Heart Journal. 1995; 73(3):237-241

62. Wijesinghe M, Williams M, Perrin K, Weatherall M, Beasley R. The effect of supplemental oxygen on hypercapnia in subjects with obesity-associated hypoventilation: A randomized, crossover, clinical study. Chest. 2011; 139(5):1018-1024

63. Yamamoto H, Teramoto S, Yamaguchi Y, Hanaoka Y, Ishii M, Hibi S et al. Long-term oxygen administration reduces plasma adrenomedullin levels in patients with obstructive sleep apnea syndrome. Sleep Medicine. 2007; 9(1):80-87

Appendices

Appendix A: Review protocols

Table 6: Review protocol: oxygen therapy (with CPAP/NIV)

Field	Content
PROSPERO registration number	Not registered
Review title	oxygen therapy
Review question	What is the clinical and cost effectiveness of oxygen therapy adjunctive to ventilatory support for people who do not fulfil LTOT criteria for managing obstructive sleep apnoea/hypopnoea syndrome, obesity hypoventilation syndrome and COPD-OSAHS overlap syndrome?
Objective	To determine the clinical and cost effectiveness of oxygen therapy adjunctive to ventilatory support for patients who do not fulfil LTOT criteria for managing obstructive sleep apnoea/hypopnoea syndrome, obesity hypoventilation syndrome and COPD-OSAHS overlap syndrome
Searches	The following databases (from inception) 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.
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).
Population	Inclusion: People (16 and older) with OSAHS, OHS or COPD-OSAHS overlap syndrome (only if formal diagnosis methods) and nocturnal hypoxaemia despite optimised ventilatory support who do not already fulfil criteria for LTOT

	Population will be stratified by: • OSAHS vs OHS vs COPD-OSAHS overlap syndrome • Mild vs moderate vs severe (based on AHI/ODI) 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.
Intervention/Exposure/Test	Oxygen therapy + CPAP or non-invasive ventilation (NIV)
Comparator/Reference standard/Confounding factors	CPAP or non-invasive ventilation (NIV) without oxygen therapy • Any other OSAHS/OHS/ COPD-OSAHS overlap syndrome treatment • No treatment/sham treatment
Types of study to be included	RCTs only Minimum duration of follow-up 1 months Parallel or crossover to be included
Other exclusion criteria	None
Context	NA
Primary outcomes (critical outcomes)	• Generic or disease specific quality of life measures (continuous) • Mortality (dichotomous) Outcomes will be separated into short term (latest follow-up to 6 months) and long term (latest follow-up beyond 6 months)
Secondary outcomes (important outcomes)	• Sleepiness scores (continuous, e.g. Epworth) • Apnoea-Hypopnoea index or respiratory disturbance index (continuous) • Oxygen desaturation index (continuous) • Daytime pO2 (continuous) • Daytime pCO2 (continuous) • Daytime bicarbonate (continuous) • Nocturnal transcutaneous CO_2 control (continuous) • Nocturnal oximetry (continuous) • Minor adverse effects of treatment (rates or dichotomous) • Adherence (continuous) • Driving outcomes (continuous) • Neurocognitive outcomes (continuous) • Pulmonary artery pressure by TTE (continuous) • Patient preference (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)
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.
Risk of bias (quality) assessment	Risk of bias will be assessed using the appropriate checklist as described in Developing NICE guidelines: the manual. For Intervention reviews • 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.
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.	
Analysis of sub-groups	Subgroups that will be investigated if heterogeneity is present: • 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	
Type and method of review	☒	Intervention
	☐	Diagnostic
	☐	Prognostic
	☐	Qualitative
	☐	Epidemiologic
	☐	Service Delivery
	☐	Other (please specify)
Language	English	
Country	England	
Anticipated or actual start date	NA – not registered on PROSPERO	
Anticipated completion date	NA – not registered on PROSPERO	
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	
Review team members	From the National Guideline Centre: Carlos Sharpin, Guideline lead Sharangini Rajesh, Senior systematic reviewer Audrius Stonkus, Systematic reviewer Madelaine Zucker, 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	

Funding sources/sponsor	This systematic review is being completed by the National Guideline Centre which receives funding from NICE.
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.
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
Other registration details	NA – not registered
Reference/URL for published protocol	NA – not registered
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.
Keywords	-
Details of existing review of same topic by same authors	NA
Additional information	-
Details of final publication	www.nice.org.uk

Table 7 Review protocol: oxygen therapy (alone)

Field	Content
PROSPERO registration number	Not registered
Review title	oxygen therapy

Review question	What is the clinical and cost effectiveness of oxygen therapy (alone) for managing obstructive sleep apnoea/hypopnoea syndrome, obesity hypoventilation syndrome and COPD-OSAHS overlap syndrome?
Objective	To determine the clinical and cost effectiveness of oxygen therapy for managing obstructive sleep apnoea/hypopnoea syndrome, obesity hypoventilation syndrome and COPD-OSAHS overlap syndrome
Searches	The following databases (from inception) will be searched: • Cochrane Central Register of Controlled Trials (CENTRAL) • Cochrane Database of Systematic Reviews (CDSR) • Embase • MEDLINE 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.
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).
Population	• Inclusion: People (16 and older) with OSAHS, OHS or COPD-OSAHS overlap syndrome Population will be stratified by: • OSAHS vs OHS vs COPD-OSAHS overlap syndrome • Mild vs moderate vs severe (based on AHI/ODI) 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.
Intervention/Exposure/Test	Oxygen therapy
Comparator/Reference standard/Confounding factors	• CPAP • non-invasive ventilation (NIV) • No treatment/sham treatment

	Comparison for people with OSA: • Oxygen vs CPAP • oxygen vs no treatment Comparison for people with OHS: • Oxygen vs NIV (non-invasive ventilation) • oxygen vs no treatment Comparison for people with COPD-OSAHS overlap syndrome: • Oxygen vs non-invasive ventilation (NIV) • oxygen vs no treatment
Types of study to be included	• Published NMAs and IPDs will be considered for inclusion. • RCTs only • Parallel or crossover to be included • Minimum duration of follow-up 1 months
Other exclusion criteria	None
Context	-
Primary outcomes (critical outcomes)	• Generic or disease specific quality of life measures (continuous) • Mortality (dichotomous) Outcomes will be separated into short term (latest follow-up to 6 months) and long term (latest follow-up beyond 6 months)
Secondary outcomes (important outcomes)	• Sleepiness scores (continuous, e.g. Epworth) • Apnoea-Hypopnoea index or respiratory disturbance index (continuous) • Oxygen desaturation index (continuous) • Daytime pO2 (continuous) • Daytime pCO2 (continuous) • Daytime bicarbonate (continuous) • Nocturnal transcutaneous CO_2 control (continuous) • Nocturnal oximetry (continuous) • Minor adverse effects of treatment (rates or dichotomous) • Adherence (continuous) • Driving outcomes (continuous) • Neurocognitive outcomes (continuous) • Pulmonary artery pressure by TTE (continuous) • Patient preference (continuous) • Impact on co-existing conditions: ○ HbA1c for diabetes (continuous) ○ Cardiovascular events for cardiovascular disease (dichotomous) ○ 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)
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.
Risk of bias (quality) assessment	Risk of bias will be assessed using the appropriate checklist as described in Developing NICE guidelines: the manual. For Intervention reviews • 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.
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.
Analysis of sub-groups	Subgroups that will be investigated if heterogeneity is present: • 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 • People who are non-compliant to CPAP/NIV vs people who are compliant to CPAP/ NIV • Those who cannot use oral devices (mandibular advancement splints) for example in edentulous patients vs people who are able to use oral devices	
Type and method of review	☒	Intervention
	☐	Diagnostic
	☐	Prognostic
	☐	Qualitative
	☐	Epidemiologic
	☐	Service Delivery
	☐	Other (please specify)
Language	English	
Country	England	
Anticipated or actual start date	NA – not registered on PROSPERO	
Anticipated completion date	NA – not registered on PROSPERO	
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	
Review team members	From the National Guideline Centre: Carlos Sharpin, Guideline lead Sharangini Rajesh, Senior systematic reviewer Audrius Stonkus, Systematic reviewer Madelaine Zucker, 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	
Funding sources/sponsor	This systematic review is being completed by the National Guideline Centre which receives funding from NICE.	

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.
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
Other registration details	NA – not registered
Reference/URL for published protocol	NA – not registered
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.
Keywords	-
Details of existing review of same topic by same authors	NA
Additional information	-
Details of final publication	www.nice.org.uk

Table 8: 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).[33] **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 7 oxygen therapy

This literature search strategy was used for the following review;

- What is the clinical and cost effectiveness of oxygen therapy adjunctive to ventilatory support for people who do not fulfil long term oxygen therapy (LTOT) criteria for managing obstructive sleep apnoea/hypopnoea syndrome, obesity hypoventilation syndrome and COPD-OSAHS overlap syndrome?

The literature searches for this review are detailed below and complied with the methodology outlined in Developing NICE guidelines: the manual.[33]

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 9: 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 Observational studies
Embase (OVID)	1974 – 6 July 2020	Exclusions Randomised controlled trials Systematic review studies Observational 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.	Oxygen Inhalation Therapy/
29.	((oxygen or O2) adj3 (therap* or administrat* or supplement*)).ti,ab.
30.	28 or 29
31.	27 and 30
32.	randomized controlled trial.pt.
33.	controlled clinical trial.pt.
34.	randomi#ed.ti,ab.
35.	placebo.ab.
36.	randomly.ti,ab.
37.	Clinical Trials as topic.sh.
38.	trial.ti.
39.	or/32-38
40.	Meta-Analysis/
41.	exp Meta-Analysis as Topic/
42.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
43.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
44.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
45.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
46.	(search* adj4 literature).ab.
47.	(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.
48.	cochrane.jw.
49.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
50.	or/40-49
51.	Epidemiologic studies/
52.	Observational study/
53.	exp Cohort studies/

54.	(cohort adj (study or studies or analys* or data)).ti,ab.
55.	((follow up or observational or uncontrolled or non randomi#ed or epidemiologic*) adj (study or studies or data)).ti,ab.
56.	((longitudinal or retrospective or prospective or cross sectional) and (study or studies or review or analys* or cohort* or data)).ti,ab.
57.	Controlled Before-After Studies/
58.	Historically Controlled Study/
59.	Interrupted Time Series Analysis/
60.	(before adj2 after adj2 (study or studies or data)).ti,ab.
61.	exp case control studies/
62.	case control*.ti,ab.
63.	Cross-sectional studies/
64.	(cross sectional and (study or studies or review or analys* or cohort* or data)).ti,ab.
65.	or/51-64
66.	31 and (39 or 50 or 65)

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.	oxygen therapy/
27.	((oxygen or O2) adj3 (therap* or administrat* or supplement*)).ti,ab.
28.	26 or 27
29.	25 and 28

30.	random*.ti,ab.
31.	factorial*.ti,ab.
32.	(crossover* or cross over*).ti,ab.
33.	((doubl* or singl*) adj blind*).ti,ab.
34.	(assign* or allocat* or volunteer* or placebo*).ti,ab.
35.	crossover procedure/
36.	single blind procedure/
37.	randomized controlled trial/
38.	double blind procedure/
39.	or/30-38
40.	systematic review/
41.	meta-analysis/
42.	(meta analy* or metanaly* or metaanaly* or meta regression).ti,ab.
43.	((systematic* or evidence*) adj3 (review* or overview*)).ti,ab.
44.	(reference list* or bibliograph* or hand search* or manual search* or relevant journals).ab.
45.	(search strategy or search criteria or systematic search or study selection or data extraction).ab.
46.	(search* adj4 literature).ab.
47.	(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.
48.	cochrane.jw.
49.	((multiple treatment* or indirect or mixed) adj2 comparison*).ti,ab.
50.	or/40-49
51.	Clinical study/
52.	Observational study/
53.	family study/
54.	longitudinal study/
55.	retrospective study/
56.	prospective study/
57.	cohort analysis/
58.	follow-up/
59.	cohort*.ti,ab.
60.	58 and 59
61.	(cohort adj (study or studies or analys* or data)).ti,ab.
62.	((follow up or observational or uncontrolled or non randomi#ed or epidemiologic*) adj (study or studies or data)).ti,ab.
63.	((longitudinal or retrospective or prospective or cross sectional) and (study or studies or review or analys* or cohort* or data)).ti,ab.
64.	(before adj2 after adj2 (study or studies or data)).ti,ab.
65.	or/51-57,60-64
66.	exp case control study/
67.	case control*.ti,ab.
68.	cross-sectional study/
69.	(cross sectional and (study or studies or review or analys* or cohort* or data)).ti,ab.
70.	or/65-69
71.	29 and (39 or 50 or 70)

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: [Oxygen Inhalation Therapy] explode all trees
#9.	((oxygen or O2) near/3 (therap* or administrat* or supplement*)):ti,ab
#10.	(or #8-#9)
#11.	#7 AND #10

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 10: 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 11: 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

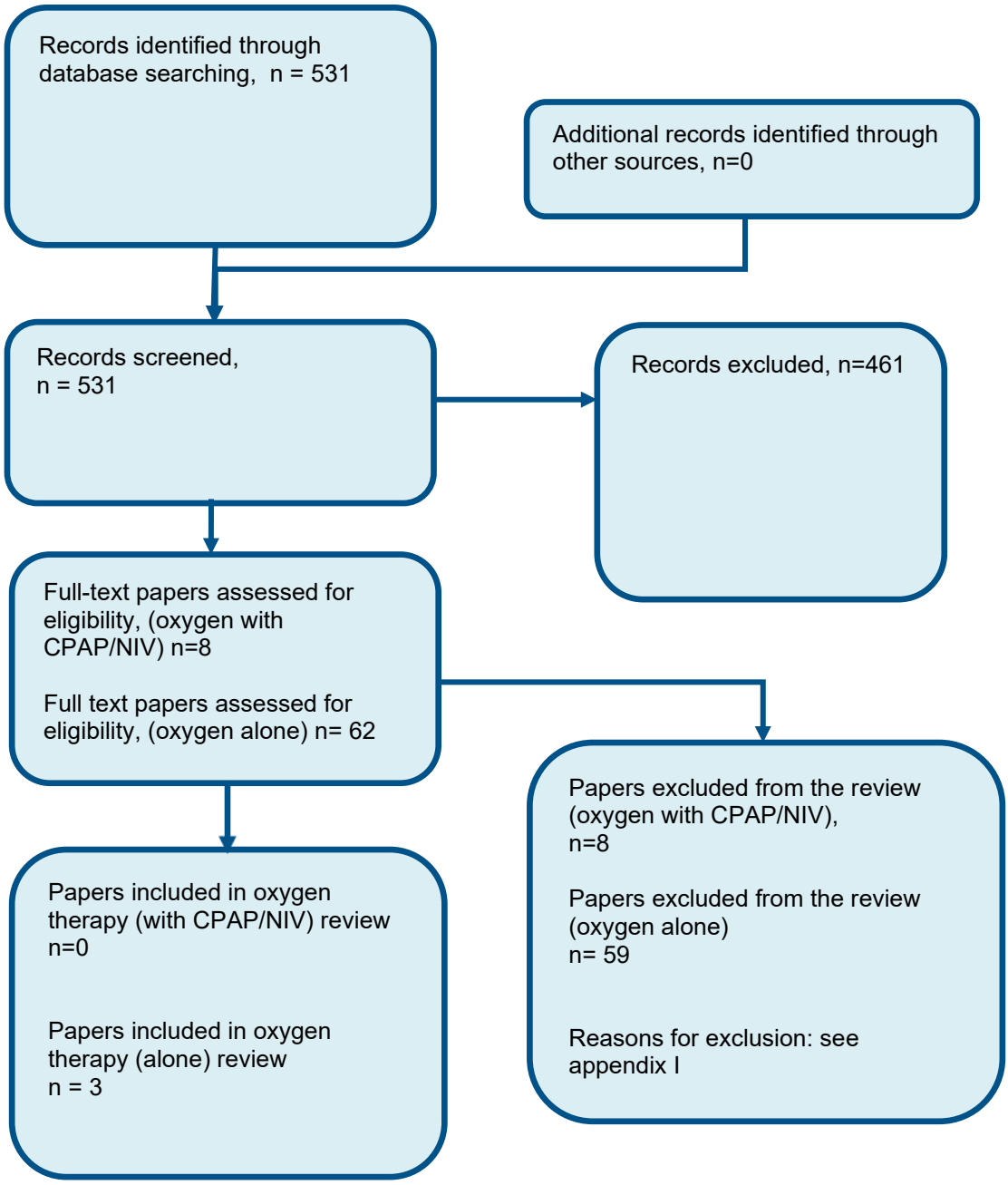
Embase (Ovid) search terms

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 oxygen therapy



Appendix D: Clinical evidence tables

D.1 Clinical evidence for Oxygen therapy (alone) compared to other/no interventions

Study (subsidiary papers)	**HeartBEAT trial: Gottlieb 2014[10] (Lewis 2017[21])**
Study type	RCT (Patient randomised; Parallel)
Number of studies (number of participants)	(n=318)
Countries and setting	Conducted in USA; Setting: four participating medical centres in the USA
Line of therapy	Unclear
Duration of study	Intervention + follow up: 12 weeks
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Moderate
Subgroup analysis within study	Not applicable
Inclusion criteria	Patients with an apnoea hypopnea index (AHI) between 15-50 events/hour were eligible for randomisation. Age 45-75, Berlin Questionnaire score 2 or 3 and established coronary heart disease (prior myocardial infarction or coronary artery revascularisation or angiographically documented >70% stenosis of a major coronary artery). Or 3 or more of the following established cardiovascular risk factors: hypertension (SBP >140 or DBP >90 or use of antihypertensive medication), diabetes mellitus, obesity (BMI over 30 kg/m^2), dyslipidaemia (total cholesterol >240 mg/dl, LDL >160 mg/dl, HDL <45 mg/dl, or taking lipid-lowering medication).

Study (subsidiary papers)	HeartBEAT trial: Gottlieb 2014[10] (Lewis 2017[21])
Exclusion criteria	The following patients were excluded; patients with an AHI >50, an o2 saturation of less than 85%, central apnoea index above 5, diagnosed heart failure with left ventricular ejection fraction <35% or New York Heart Association Class ≥2, poorly controlled hypertension (SBP >170 or DBP >110), poorly controlled diabetes (HbA1c >9.0%), myocardial infarction, stroke or coronary revascularization procedure within 3 months, resting oxyhaemoglobin saturation <90%, severe chronic insomnia with reported usual sleep duration <4 hours per night, severe sleepiness with an Epworth Sleepiness Scale score ≥16 or report of falling, asleep while driving within the previous 2 years, pregnancy or a plan to become pregnant within 6 months, smoking in the bedroom by the participant or anyone sharing a bedroom with the participant, current use of supplemental oxygen, current or past use of a positive airway pressure device or surgery for treatment of sleep apnoea, any uncontrolled medical problem that the investigator felt would significantly impair ability to participate in the study examinations, inability or unwillingness to provide informed consent.
Recruitment/selection of patients	Patients were recruited from cardiology practices in the four participating medical centres in the USA.
Age, gender and ethnicity	Age - Mean (SD): CPAP = 63.5 (7.0), NSO = 62.9 (7.3). HLSE = 63.1(7.7). Gender (M:F): 209/109. Ethnicity: CPAP = 84% white, 7% black, 9% other HLSE = 83% white, 11% black, 6% other NSO = 74% white, 20% black, 0% other
Further population details	1. BMI: BMI of 30 2 kg/m2 or more. Co-existing conditions: (CAD or cardiovascular disease risk factors). 3. High risk occupation group: Not stated / Unclear 4. Sleepiness: ESS <9
Indirectness of population	Serious indirectness: mixed severity AHI moderate to severe (AHI 15-50)Mean AHI moderate severity.
Interventions	(n=106) Intervention 1: oxygen therapy. A stationary oxygen concentrator (EverFlo, Philips Respironics) was used to provide nightly treatment with oxygen at a rate of 2 litres per minute through a nasal cannula. Duration 12 weeks. Concurrent medication/care: All participants received standardized education in habits that promote improved sleep quality and reduce cardiovascular risk, including advice on diet and exercise. This information was based on guidelines from the American Heart Association. (n=106) Intervention 2: CPAP. CPAP group received a CPAP device with automatic adjustment (REMstar

Study (subsidiary papers)	HeartBEAT trial: Gottlieb 2014[10] (Lewis 2017[21])
	Auto CPAP, Philips Respironics), set at a pressure range of 4 to 20 cm of water for 7 days and then reset to the best fixed pressure for each patient, defined as the 90th percentile of the pressure distribution generated through automatic adjustment during those 7 days. Duration 12 weeks. Concurrent medication/care: All participants received standardized education in habits that promote improved sleep quality and reduce cardiovascular risk, including advice on diet and exercise. This information was based on guidelines from the American Heart Association. (n=106) Intervention 3: No treatment/sham treatment - No active treatment/sham treatment. The education in healthy sleep habits included suggestions for maintaining a regular sleep schedule, avoiding alcohol near bedtime, and maintaining sleep duration of 7-8 hours per night. Subjects were also provided with education on a heart-healthy lifestyle, including recommendations for weight loss (for overweight and obese subjects), healthy diet, regular exercise, smoking cessation, and medication adherence. Duration 12 weeks. Concurrent medication/care: All participants received standardized education in habits that promote improved sleep quality and reduce cardiovascular risk, including advice on diet and exercise. This information was based on guidelines from the American Heart Association.
Funding	Equipment / drugs provided by industry

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OXYGEN THERAPY versus CPAP

Protocol outcome 1: Quality of life at >1 month
- Actual outcome for Moderate: SF36 – score 0-100; higher values indicate better quality of life; vitality at 12 weeks; Group 1: mean 49.8 (SD 9); n=101, Group 2: mean 51.8 (SD 11.1); n=99
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 5, Reason: 2 had adverse events, 3 dissatisfaction; Group 2 Number missing: 7, Reason: 5 dissatisfaction, 1 coexisting illness, 1 lost to Follow-up
- Actual outcome for Moderate: SF36 - mental component score at 12 weeks; Group 1: mean 51.9 (SD 10.1); n=100, Group 2: mean 52.6 (SD 10); n=99
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 6, Reason: 1 reason not reported, 2 had adverse events, 3 dissatisfaction; Group 2 Number missing: 7, Reason: 5 dissatisfaction, 1 coexisting illness, 1 lost to Follow-up
- Actual outcome for Moderate: SF36 - physical component score at 12 weeks; Group 1: mean 44.1 (SD 10.5); n=100, Group 2: mean 44.6 (SD 10.2); n=99

Study (subsidiary papers)	HeartBEAT trial: Gottlieb 2014[10] (Lewis 2017[21])
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness, Comments: mixed severity AHI; Group 1 Number missing: 6, Reason: 1 reason not reported, 2 had adverse events, 3 dissatisfaction; Group 2 Number missing: 7, Reason: 5 dissatisfaction, 1 coexisting illness, 1 lost to Follow-up - Actual outcome for Moderate: PHQ-9 (depression score) at 12 weeks; Group 1: mean 4.2 (SD 4); n=101, Group 2: mean 3.5 (SD 4.1); n=99; the patient health questionnaire - 9 , a 9-item instrument with scores ranging 0-27, Top=High is poor outcome Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 5, Reason: 2 had adverse events, 3 dissatisfaction; Group 2 Number missing: 7, Reason: 5 dissatisfaction, 1 coexisting illness, 1 lost to Follow-up Protocol outcome 2: Mortality at >1 month - Actual outcome for Moderate: mortality at 12 weeks; Group 1: 0/101, Group 2: 0/99 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness, Comments: mixed severity AHI; Group 1 Number missing: 5, Reason: 2 had adverse events, 3 dissatisfaction; Group 2 Number missing: 7, Reason: 5 dissatisfaction, 1 coexisting illness, 1 lost to Follow-up Protocol outcome 3: Sleepiness score at >1 month - Actual outcome for Moderate: sleepiness summary score at 12 weeks; 0-12 with higher scores indicating more sleepiness; Group 1: mean 4.6 (SD 2.8); n=100, Group 2: mean 4.3 (SD 2.7); n=96 Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low; Indirectness of outcome: Serious indirectness, Comments: mixed severity population; Baseline details: baseline values of ESS differ (mean values = 8.1 in the CPAP group and 9.7 in the oxygen group); Group 1 Number missing: 6, Reason: 2 had adverse events, 3 dissatisfaction, 1 lost to Follow-up; Group 2 Number missing: 10, Reason: 5 dissatisfaction, 1 coexisting illness, 2 lost to Follow-up, 2 reason not reported Protocol outcome 4: Minor adverse effects of Treatment at >1 month - Actual outcome for Moderate: Atrial fibrillation at 12 weeks; Group 1: 2/101, Group 2: 1/99 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness, Comments: mixed severity AHI; Group 1 Number missing: 5, Reason: 2 had adverse events, 3 dissatisfaction; Group 2 Number missing: 7, Reason: 5 dissatisfaction, 1 coexisting illness, 1 lost to Follow-up Protocol outcome 5: Driving outcomes at >1 month - Actual outcome for Moderate: no of motor vehicle accidents at 12 weeks; Group 1: 1/101, Group 2: 0/99 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness, Comments: mixed severity AHI; Group 1 Number missing: 5, Reason: 2 had adverse events, 3 dissatisfaction; Group 2 Number missing: 7, Reason: 5 dissatisfaction, 1 coexisting illness, 1 lost to Follow-up Protocol outcome 6: CV events at >1 month	

Study (subsidiary papers)	HeartBEAT trial: Gottlieb 2014[10] (Lewis 2017[21])

- Actual outcome for Moderate: cardiovascular complications at 12 weeks; Group 1: 0/101, Group 2: 0/99
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 5, Reason: 2 had adverse events, 3 dissatisfaction; Group 2 Number missing: 7, Reason: 5 dissatisfaction, 1 coexisting illness, 1 lost to Follow-up

Protocol outcome 7: Systolic BP at >1 month
- Actual outcome for Moderate: 24 hour mean systolic BP at 12 weeks; Group 1: mean 126.9 mmHg (SD 16.5); n=94, Group 2: mean 123.4 mmHg (SD 12.8); n=90
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 12, Reason: 2 had adverse events, 3 dissatisfaction 7 no reason reported; Group 2 Number missing: 16, Reason: 5 dissatisfaction, 1 coexisting illness, 1 lost to Follow-up, 9 reason not reported
- Actual outcome for Moderate: daytime mean systolic BP at 12 weeks; Group 1: mean 130.2 mm hg (SD 17); n=95, Group 2: mean 126.8 mm hg (SD 12.9); n=93
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - ; Indirectness of outcome: Serious indirectness, Comments: mixed severity population; Group 1 Number missing: 11, Reason: 2 had adverse events, 3 dissatisfaction, 6 reason not reported; Group 2 Number missing: 13, Reason: 5 dissatisfaction, 1 coexisting illness, 7 reason not reported

Protocol outcome 8: Nocturnal oximetry at Define
- Actual outcome for Moderate: ODI change score at 12 weeks; Group 1: mean -19.6 (SD 13.6); n=101, Group 2: mean -17.2 (SD 13.6); n=99
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - ; Indirectness of outcome: Serious indirectness, Comments: mixed severity population; Group 1 Number missing: 5, Reason: 2 had adverse events, 3 dissatisfaction; Group 2 Number missing: 7, Reason: 5 dissatisfaction, 1 coexisting illness, 1 lost to Follow-up

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OXYGEN THERAPY versus NO ACTIVE TREATMENT/SHAM TREATMENT

Protocol outcome 1: Quality of life at >1 month
- Actual outcome for Moderate: SF36 - vitality at 12 weeks; Group 1: mean 49.8 (SD 9); n=101, Group 2: mean 49.5 (SD 9.4); n=100
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness, Comments: mixed severity AHI; Group 1 Number missing: 5, Reason: 2 had adverse events, 3 dissatisfaction; Group 2 Number missing: 6, Reason: 1 withdrew, 1 administrative reasons, 3 lost to Follow-up, 1 reason not reported
- Actual outcome for Moderate: SF36 - mental component score at 12 weeks; Group 1: mean 51.9 (SD 10.1); n=100, Group 2: mean 49.7 (SD 11); n=99
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness, Comments: mixed severity AHI; Group 1 Number missing: 6, Reason: 2 had adverse events, 3 dissatisfaction, reason not reported; Group 2 Number missing: 7, Reason: 1 withdrew, 1 administrative reasons, 3 lost to Follow-up, 2 reason not reported

Study (subsidiary papers)	HeartBEAT trial: Gottlieb 2014[10] (Lewis 2017[21])

- Actual outcome for Moderate: SF36 - physical component score at 12 weeks; Group 1: mean 44.1 (SD 10.5); n=100, Group 2: mean 42.9 (SD 9.3); n=99
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 6, Reason: 2 had adverse events, 3 dissatisfaction, 1 reason not reported; Group 2 Number missing: 7, Reason: 1 withdrew, 1 administrative reasons, 3 lost to Follow-up, 2 reason not reported
- Actual outcome for Moderate: PHQ-9 (depression score) at 12 weeks; Group 1: mean 4.2 (SD 4); n=101, Group 2: mean 5.1 (SD 4); n=100, the patient health questionnaire - 9 , a 9-item instrument with scores ranging 0-27, High is poor outcome Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness, Comments: mixed severity AHI; Group 1 Number missing: 5, Reason: 2 had adverse events, 3 dissatisfaction; Group 2 Number missing: 6, Reason: 5 dissatisfaction, 1 coexisting illness, 1 lost to Follow-up

Protocol outcome 2: Mortality at >1 month
- Actual outcome for Moderate: mortality at 12 weeks; Group 1: 0/101, Group 2: 0/101
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - ; Indirectness of outcome: Serious indirectness ; Baseline details: difference in sleepiness and patients using beta blockers; Group 1 Number missing: 5, Reason: 2 had adverse events, 3 dissatisfaction; Group 2 Number missing: 5, Reason: 1 withdrew, 1 administrative reasons, 3 lost to Follow-up

Protocol outcome 3: Sleepiness score at >1 month
- Actual outcome for Moderate: sleepiness summary score at 12 weeks; Group 1: mean 4.6 (SD 2.8); n=100, Group 2: mean 5.6 (SD 2.9); n=98; sleepiness summary score 0-12 Top=High is poor outcome
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low; Indirectness of outcome: Serious indirectness, Comments: mixed severity population; Group 1 Number missing: 6, Reason: 2 had adverse events, 3 dissatisfaction, 1 lost to Follow-up; Group 2 Number missing: 10, Reason: 5 dissatisfaction, 1 coexisting illness, 4 lost to Follow-up

Protocol outcome 4: Minor adverse effects of Tx at >1 month
- Actual outcome for Moderate: Atrial fibrillation at 12 weeks; Group 1: 2/101, Group 2: 1/101
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 5, Reason: 2 had adverse events, 3 dissatisfaction; Group 2 Number missing: 5, Reason: 1 withdrew, 1 administrative reasons, 3 lost to Follow-up

Protocol outcome 5: Driving outcomes at >1 month
- Actual outcome for Moderate: no of motor vehicle accidents at 12 weeks; Group 1: 1/101, Group 2: 0/101
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness, Comments: mixed severity AHI; Group 1 Number missing: 5, Reason: 2 had adverse events, 3 dissatisfaction; Group 2 Number missing: 5, Reason: 1 withdrew, 1 administrative reasons, 3 lost to Follow-up

Study (subsidiary papers)	HeartBEAT trial: Gottlieb 2014[10] (Lewis 2017[21])
Protocol outcome 6: CV events at >1 month - Actual outcome for Moderate: cardiovascular complications at 12 weeks; Group 1: 0/101, Group 2: 4/101 (one of each: unstable angina, myocardial infarction, percutaneous coronary intervention for worsening angina and stroke) Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness, Comments: mixed severity AHI; Group 1 Number missing: 5, Reason: 2 had adverse events, 3 dissatisfaction; Group 2 Number missing: 5, Reason: 1 withdrew, 1 administrative reasons, 3 lost to Follow-up Protocol outcome 7: Systolic BP at >1 month - Actual outcome for Moderate: 24 hour mean systolic BP at 12 weeks ; Group 1: mean 126.9 mmHg (SD 16.5); n=94, Group 2: mean 124.7 mmHg (SD 16.4); n=97 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness, Comments: mixed severity AHI; Group 1 Number missing: 12, Reason: 2 had adverse events, 3 dissatisfaction, 7 reason not reported; Group 2 Number missing: 9, Reason: 1 withdrew, 1 administrative reasons, 3 lost to Follow-up, 4 reason not reported - Actual outcome for Moderate: daytime mean systolic BP at 12 weeks ; Group 1: mean 130.2 mm hg (SD 17); n=95, Group 2: mean 128 mm hg (SD 16); n=97 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness, Comments: mixed severity population; Group 1 Number missing: 11, Reason: 2 had adverse events, 3 dissatisfaction, 6 reason not reported; Group 2 Number missing: 9, Reason: 3 lost to Follow-up, 1 withdrew, 1 administrative reasons, 4 reason not reported Protocol outcome 8: Nocturnal oximetry - Actual outcome for Moderate: ODI change score at 12 weeks; Group 1: mean -19.6 (SD 13.6); n=101, Group 2: mean -1.5 (SD 13.8); n=101 Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: Serious indirectness, Comments: mixed severity population; Group 1 Number missing: 5, Reason: 2 had adverse events, 3 dissatisfaction; Group 2 Number missing: 5, Reason: 3 lost to Follow-up, 1 withdrew, 1 administrative reasons	
Protocol outcomes not reported by the study	AHI/RDI at >1 month; daytime PO2 at >1 month; daytime PCO2 at >1 month; Neurocognitive outcomes at >1 month; Patient preference at >1 month; HbA1c at >1 month; adherence at Define; Pulmonary artery pressure by TTE >1 month; Nocturnal transcutaneous CO2 control at >1 month

Study	Phillips 1990[42]
Study type	RCT (Patient randomised; Crossover: no washout)
Number of studies (number of participants)	1 (n=8)
Countries and setting	Conducted in USA; Setting: sleep apnoea laboratory in the university of Kentucky college of medicine
Line of therapy	Unclear
Duration of study	Intervention time: 1 month
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Mild
Subgroup analysis within study	Not applicable
Inclusion criteria	AHI of equal to or over 5 and at least one of the following; daytime hypersomnolence with a mean sleep latency of ≤10 minutes on multiple sleep latency testing, hypertension with a mean of at least 5 measurements of either systolic blood pressure >150mm hg and/or diastolic blood pressure ≥95 mm hg, significant cardiac arrhythmias, including marked sinus arrhythmia, sinus bradycardia, frequent (>6 per hour0 premature ventricular contractions, one or more sinus pauses > 2 seconds associated with apnoea or hypoxemia, or supraventricular tachycardia.
Exclusion criteria	AHI ≥40, initial MSLT <5 mins, mean low Sa02 ≥8 %. lung disease, subjects with ≥20% central apnoea's or hypopnoeas were excluded.
Recruitment/selection of patients	Participants were recruited from patients studied in the sleep apnoea laboratory at the university of Kentucky college of medicine either through routine clinic referral or as part of an ongoing study of sleep disordered breathing in the elderly
Age, gender and ethnicity	Age - Mean (SD): 57 (13.6). Gender (M:F): 8/0. Ethnicity: not reported

Study	Phillips 1990[42]
Further population details	1. BMI: Not stated / Unclear 2. Co-existing conditions: HTN 3. High risk occupation group: Not stated / Unclear 4. Sleepiness: Not stated / Unclear
Indirectness of population	Serious indirectness: includes a mix of mild, moderate and severe OSA
Interventions	(n=8) Intervention 1: oxygen therapy. Nocturnal nasal o2 at 4 LPM nightly for one month. Duration 1 month. Concurrent medication/care: All subjects underwent the 3 treatments for 1-month duration including nasal o2, nasal compressed air and CPAP. All subjects received CPAP last. Indirectness: No indirectness (n=8) Intervention 2: CPAP. Subjects received 1 month of nasal CPAP. Levels of nasal CPAP were established initially in the lab and then adjusted at home by a respiratory therapist based on behaviour during napping. The CPAP levels ranged between 2.5 and 12.5cm H20. Duration 1 month. Concurrent medication/care: All subjects underwent a total of 3 treatments for 1-month duration including nasal o2, nasal compressed air and CPAP. All subjects received CPAP last. Indirectness: No indirectness (n=8) Intervention 3: No treatment/sham treatment - No active treatment/sham treatment. subjects received nasal compressed air nightly for 1 month. Duration 1 month. Concurrent medication/care: All subjects underwent a total of 3 treatments for 1 month duration including nasal o2, nasal compressed air and CPAP. All subjects received CPAP last. Indirectness: No indirectness
Funding	Academic or government funding

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OXYGEN THERAPY versus CPAP

Protocol outcome 1: Sleepiness score at >1 month
- Actual outcome for Moderate: Stanford sleepiness score at 1 month; Group 1: mean 2.5 (SD 0.6); n=8, Group 2: mean 2.5 (SD 0.8); n=8
Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0

Protocol outcome 2: AHI/RDI at >1 month
- Actual outcome for Moderate: AHI at 1 month; Group 1: mean 16.8 (SD 9.1); n=8, Group 2: mean 3 (SD 2.5); n=8
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low,

Study	Phillips 1990[42]
Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 3: Neurocognitive outcomes at >1 month - Actual outcome for Moderate: Attention (2&7 test) at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: digit symbol at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: long-term storage at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: consistent retrieval at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: Benton visual retention at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: Rey Figure - copy at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: Rey Figure - immediate recall at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: Rey Figure - delayed recall at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: finger tapping - dominant hand at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: finger tapping - non-dominant hand at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 4: Systolic BP at >1 month - Actual outcome for Moderate: systolic BP at 1 month; Group 1: mean 139.6 mmHg (SD 14.7); n=8, Group 2: mean 140.8 (SD 12.7); n=8	

Study	Phillips 1990[42]
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 5: Nocturnal oximetry - Actual outcome for Moderate: mean low Sa02 at 1 month; Group 1: mean 95.9 % (SD 0.8); n=8, Group 2: mean 93.7 % (SD 2.5); n=8 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: OXYGEN THERAPY versus NO ACTIVE TREATMENT/SHAM TREATMENT Protocol outcome 1: Sleepiness score at >1 month - Actual outcome for Moderate: Stanford sleepiness score at 1 month; Group 1: mean 2.5 (SD 0.6); n=8, Group 2: mean 2.9 (SD 0.8); n=8 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 2: AHI/RDI at >1 month - Actual outcome for Moderate: AHI at 1 month; Group 1: mean 16.8 (SD 9.1); n=8, Group 2: mean 22.1 (SD 16.1); n=8 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 3: Neurocognitive outcomes at >1 month - Actual outcome for Moderate: Attention (2&7 test) at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: digit symbol at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: long-term storage at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: consistent retrieval at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: Benton visual retention at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low,	

Study	Phillips 1990[42]
Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: Rey Figure - copy at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: Rey Figure - immediate recall at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: Rey Figure - delayed recall at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: finger tapping - dominant hand at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 - Actual outcome for Moderate: finger tapping - non-dominant hand at 1 month; Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; Group 1 Number missing: 0; Group 2 Number missing: 0 Protocol outcome 4: Systolic BP at >1 month - Actual outcome for Moderate: systolic BP at 1 month; Group 1: mean 139.6 mmHg (SD 14.7); n=8, Group 2: mean 144.6 (SD 16.7); n=8 Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - High; Indirectness of outcome: Serious indirectness ; 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; daytime PO2 at >1 month; daytime PCO2 at >1 month; Minor adverse effects of Treatment at >1 month; Driving outcomes at >1 month; Patient preference at >1 month; HbA1c at >1 month; CV events at >1 month; adherence at Define; Pulmonary artery pressure by TTE at Define; Nocturnal transcutaneous CO2 control at Define

Appendix E: Forest plots

E.1 Oxygen therapy (alone) compared to CPAP - moderate OSAHS)

Figure 2: Mortality

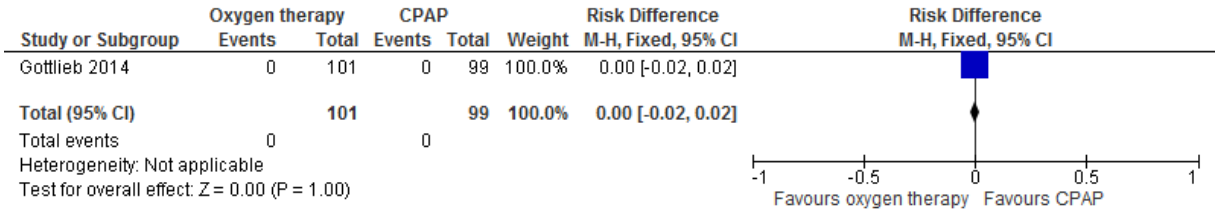


Figure 3: Atrial Fibrillation

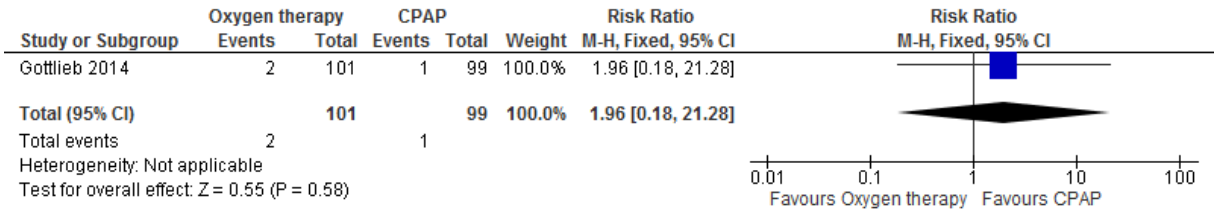


Figure 4: Cardiovascular complications

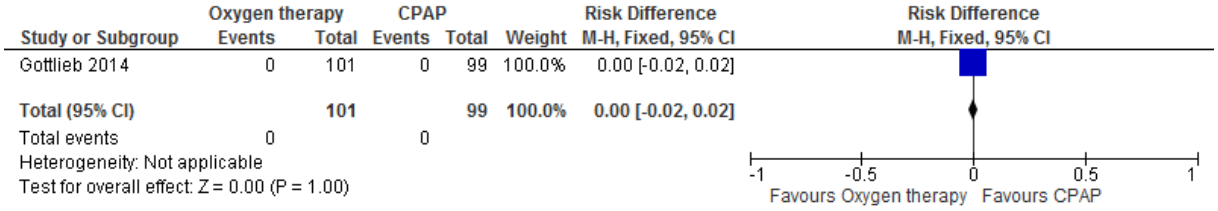


Figure 5: Number of motor vehicle accidents (12 weeks)

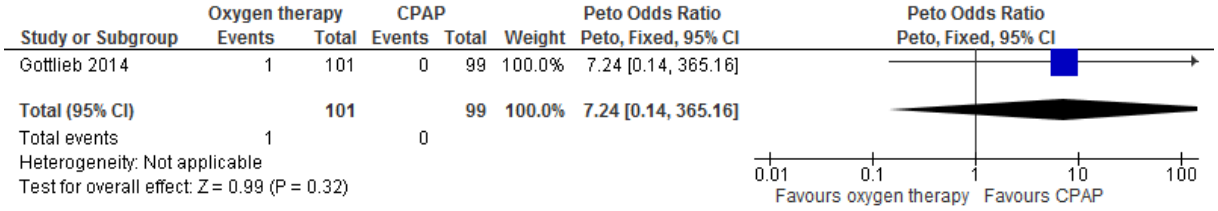


Figure 6: 24 hour mean systolic BP

Study or Subgroup	oxygen therapy Mean	SD	Total	CPAP Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Gottlieb 2014	126.9	16.5	94	123.4	12.8	90	100.0%	3.50 [-0.76, 7.76]
Total (95% CI)			**94**			**90**	**100.0%**	**3.50 [-0.76, 7.76]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.61 (P = 0.11)

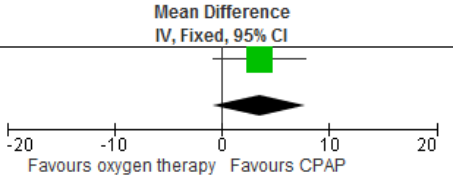


Figure 7: SF36 Physical, 0-100 (higher is better)

Study or Subgroup	Oxygen therapy Mean	SD	Total	CPAP Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Lewis 2017	44.1	10.5	100	44.6	10.2	99	100.0%	-0.50 [-3.38, 2.38]
Total (95% CI)			**100**			**99**	**100.0%**	**-0.50 [-3.38, 2.38]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.34 (P = 0.73)

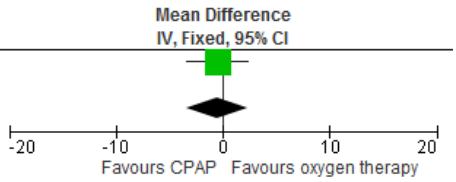


Figure 8: SF36 Mental, 0-100 (higher is better)

Study or Subgroup	Oxygen therapy Mean	SD	Total	CPAP Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Lewis 2017	51.9	10.1	100	52.6	10	99	100.0%	-0.70 [-3.49, 2.09]
Total (95% CI)			**100**			**99**	**100.0%**	**-0.70 [-3.49, 2.09]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.49 (P = 0.62)

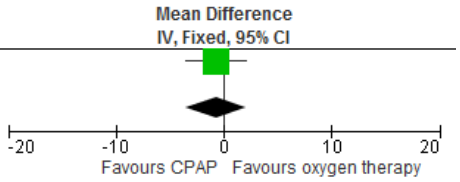


Figure 9: SF36 vitality, 0-100 (higher is better)

Study or Subgroup	Oxygen therapy Mean	SD	Total	CPAP Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Lewis 2017	49.8	9	101	51.8	11.1	99	100.0%	-2.00 [-4.80, 0.80]
Total (95% CI)			**101**			**99**	**100.0%**	**-2.00 [-4.80, 0.80]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.40 (P = 0.16)

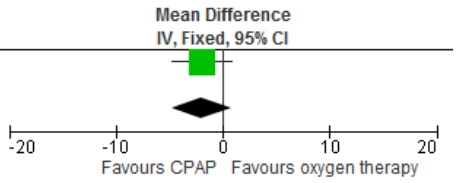


Figure 10: PHQ-9 Depression, 0-27, lower is better

Study or Subgroup	Oxygen therapy Mean	SD	Total	CPAP Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Lewis 2017	4.2	4	101	3.5	4.1	99	100.0%	0.70 [-0.42, 1.82]
Total (95% CI)			**101**			**99**	**100.0%**	**0.70 [-0.42, 1.82]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.22 (P = 0.22)

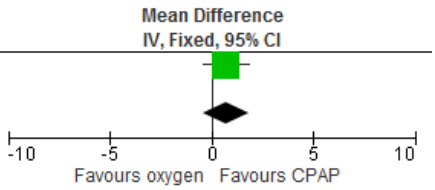


Figure 11: ODI change score

Study or Subgroup	Oxygen therapy Mean	SD	Total	CPAP Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Lewis 2017	19.6	13.6	101	17.2	13.6	99	100.0%	2.40 [-1.37, 6.17]
Total (95% CI)			**101**			**99**	**100.0%**	**2.40 [-1.37, 6.17]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.25 (P = 0.21)

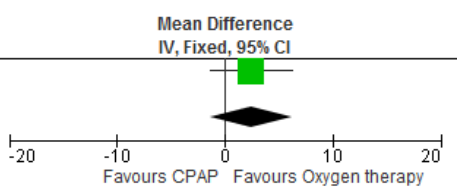


Figure 12: Daytime mean systolic BP

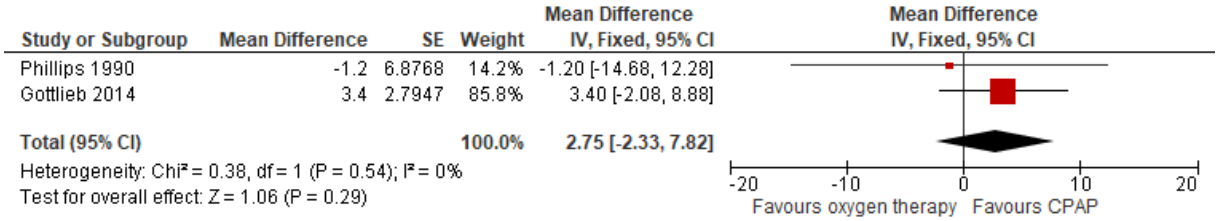


Figure 13: AHI (lower is better)

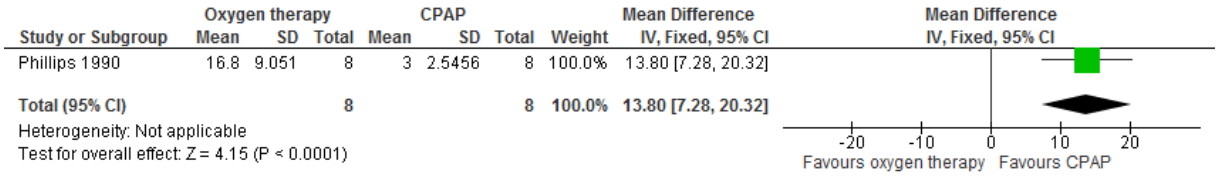


Figure 14: Stamford sleepiness score, 0-24, higher is worse

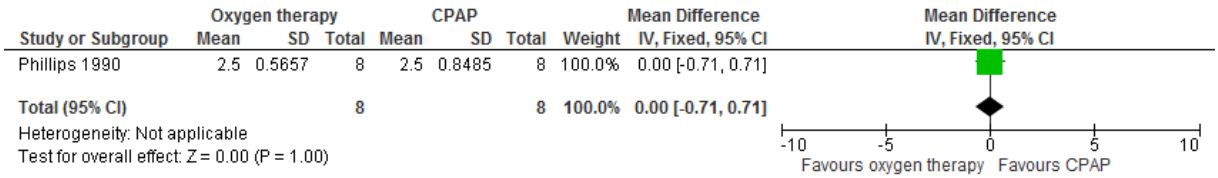


Figure 15: Attention (2&7 test)

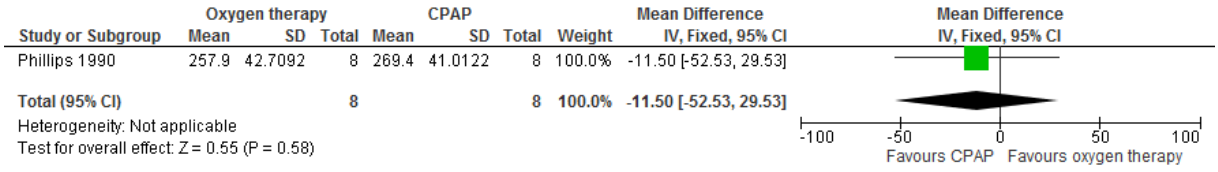


Figure 16: Digit Symbol

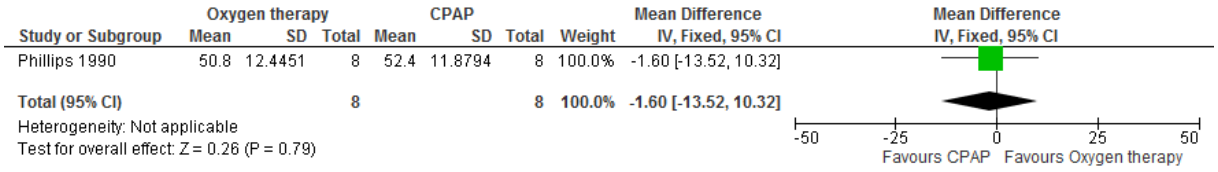


Figure 17: Selective reminding – (long term storage)

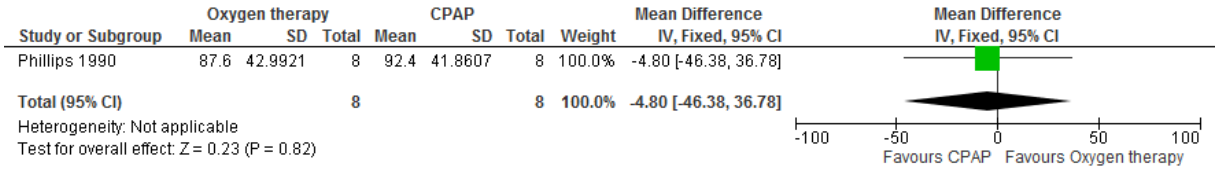


Figure 18: Selective reminding (consistent retrieval)

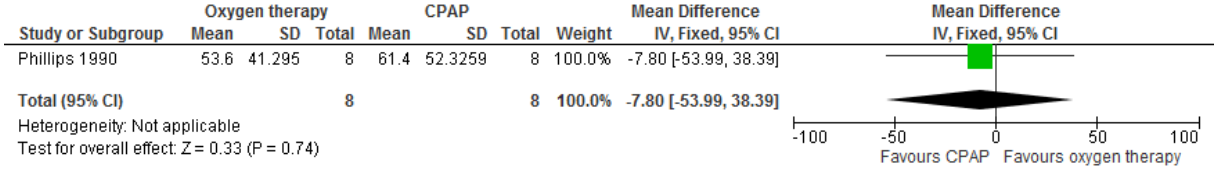


Figure 19: Benton visual retention

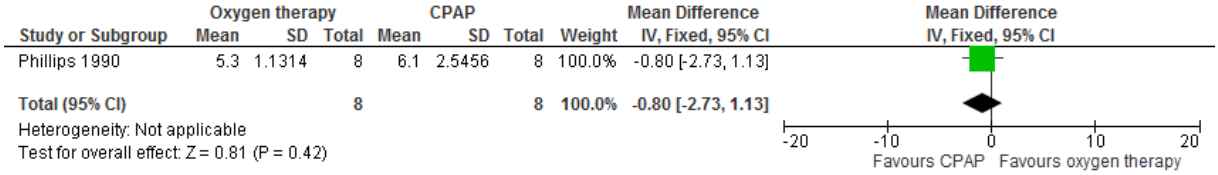


Figure 20: Rey figure (copy)

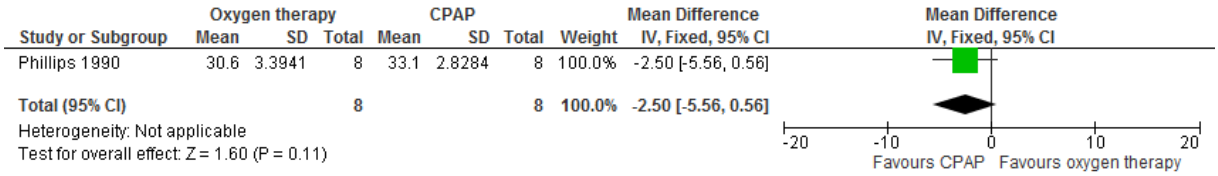


Figure 21: Rey Figure (immediate recall)

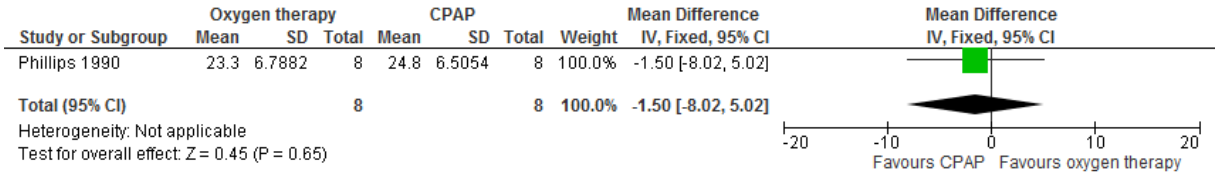


Figure 22: Rey figure (delayed recall)

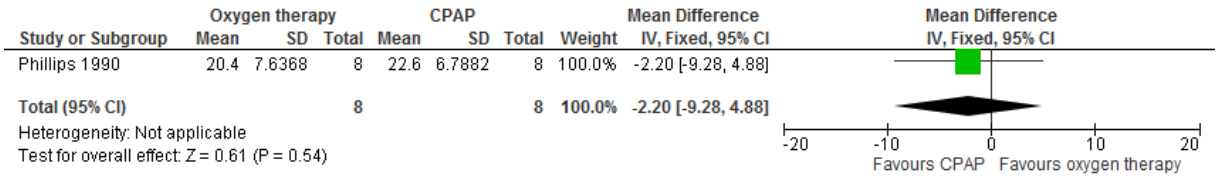


Figure 23: Finger tapping (dominant hand)

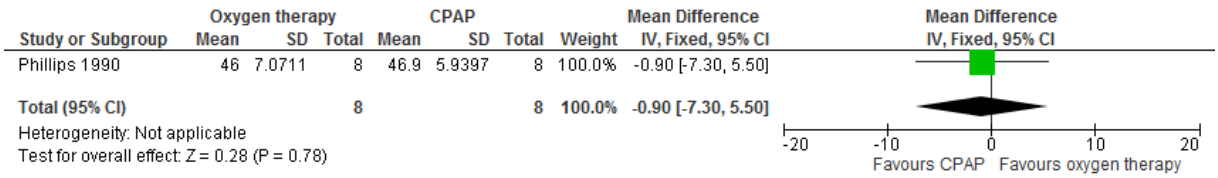
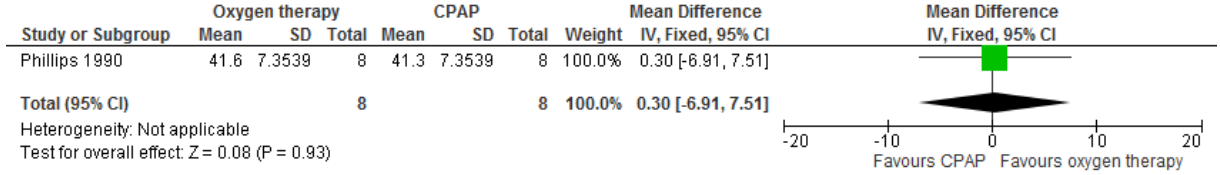


Figure 24: Finger tapping (non-dominant hand)



E.2 Oxygen therapy (alone) compared to no treatment/placebo - moderate OSAHS

Figure 25: Mortality

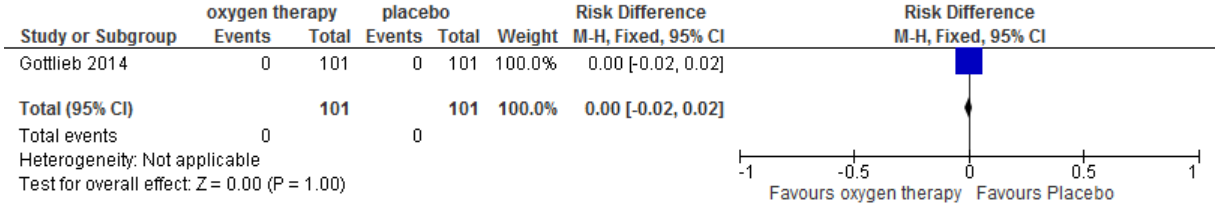


Figure 26: Atrial Fibrillation

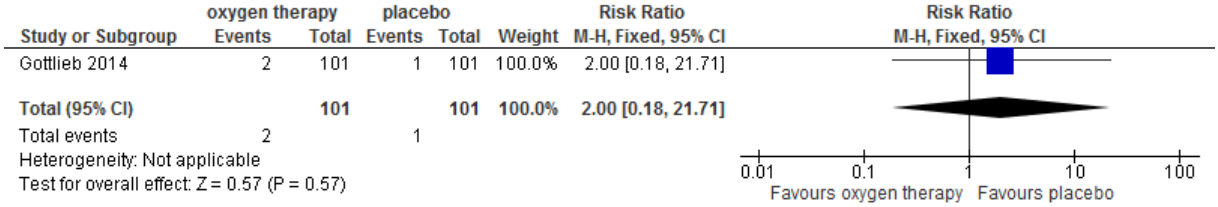


Figure 27: Cardiovascular complications

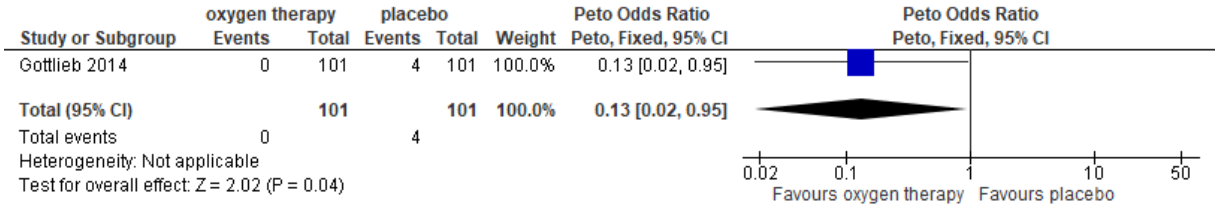


Figure 28: Number of motor vehicle accidents (12 weeks)

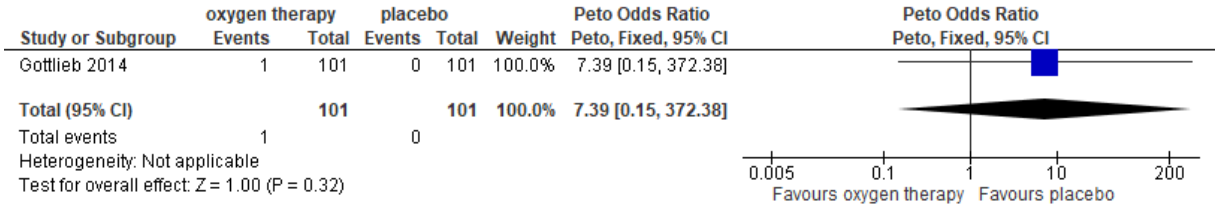


Figure 29: 24-hour mean systolic BP

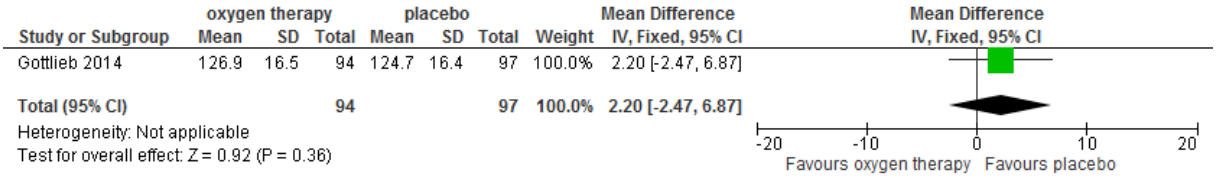


Figure 30: SF36 Physical, 0-100 (higher is better)

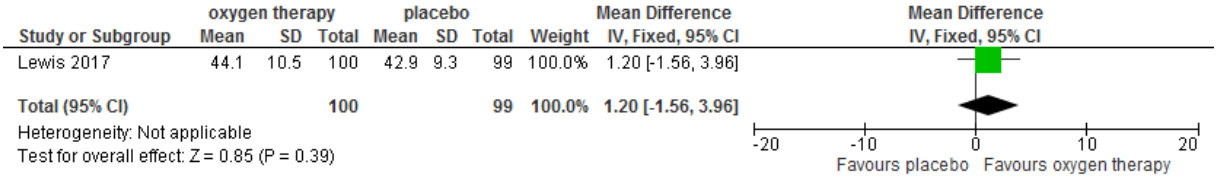


Figure 31: SF36 Mental, 0-100 (higher is better)

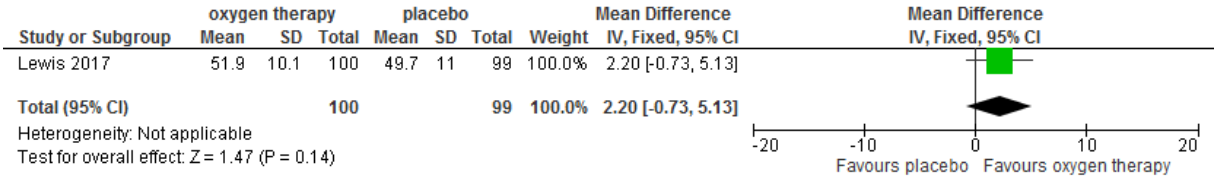


Figure 32: SF36 vitality, 0-100 (higher is better)

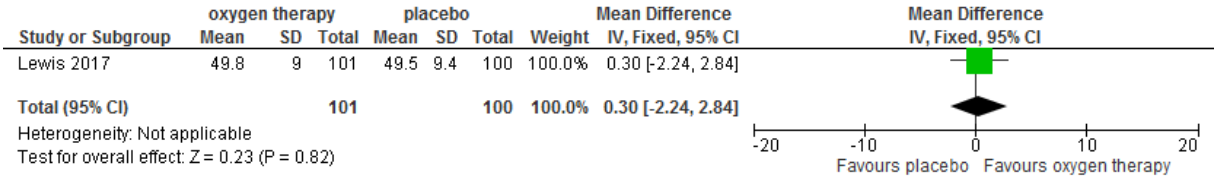


Figure 33: PHQ-9 Depression, 0-27, (lower is better)

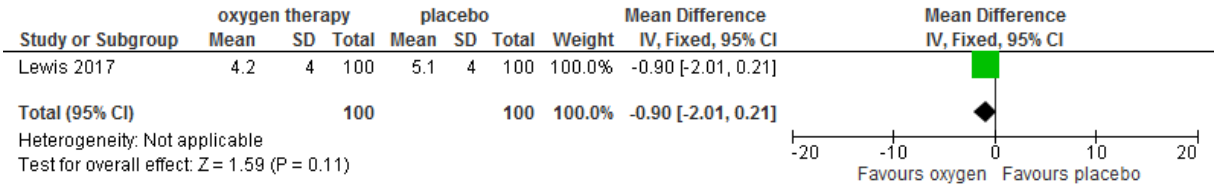


Figure 34: ODI change score, (lower is better)

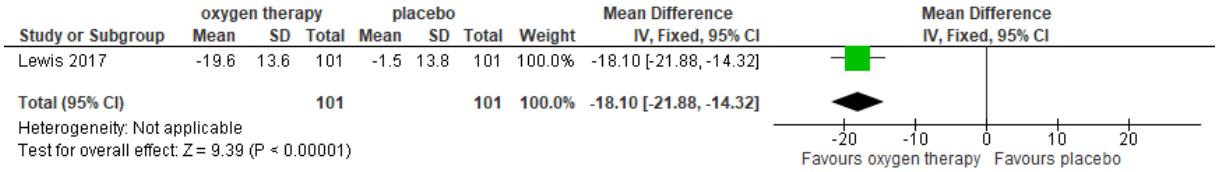


Figure 35: Daytime mean systolic BP

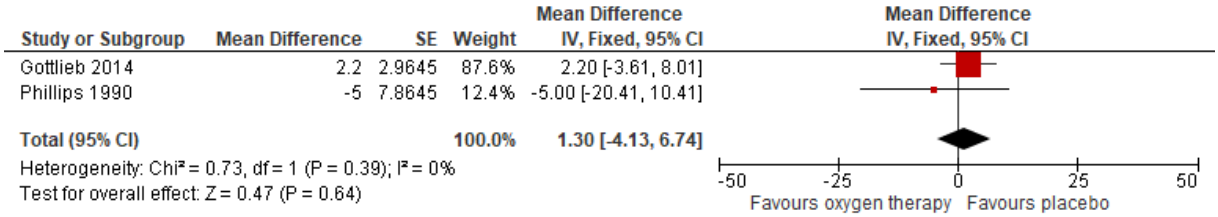


Figure 36: AHI (lower is better)

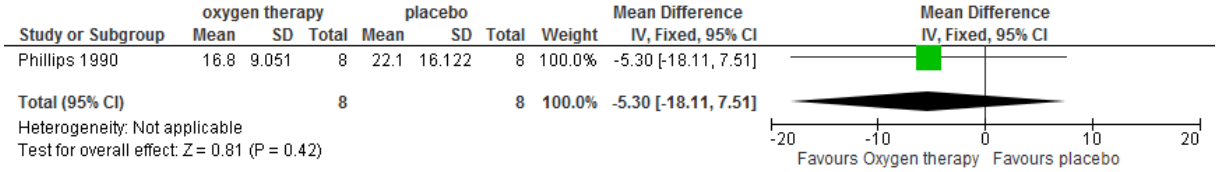


Figure 37: Stamford sleepiness score, 0-24, higher is worse

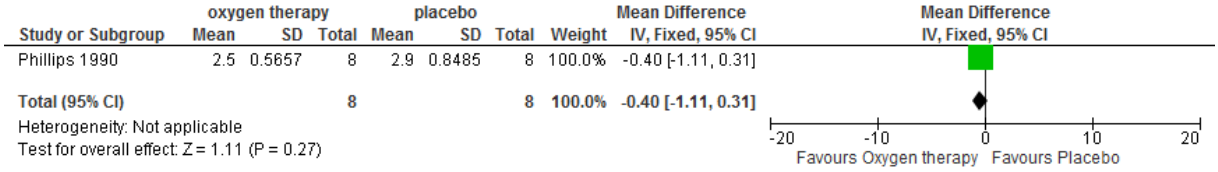


Figure 38: Attention (2&7 test)

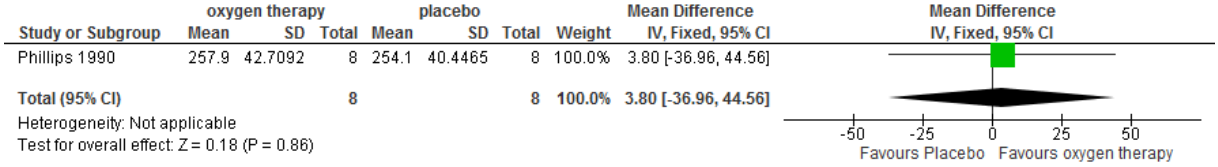


Figure 39: Digit Symbol

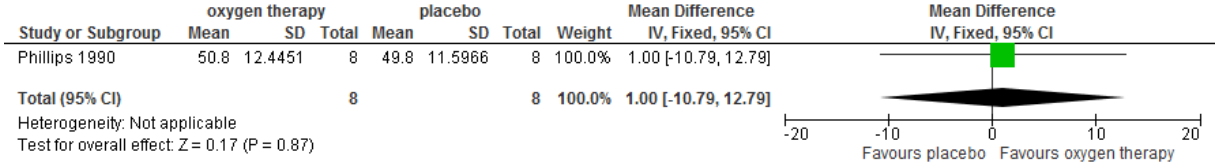


Figure 40: Selective reminding – (long term storage)

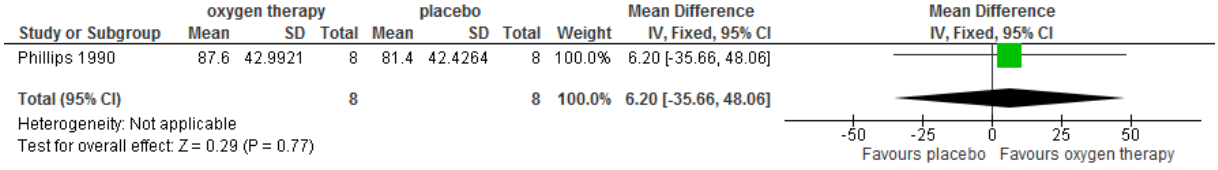


Figure 41: Selective reminding (consistent retrieval)

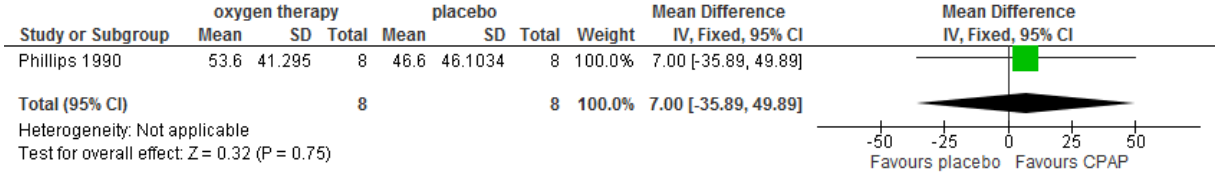


Figure 42: Benton visual retention

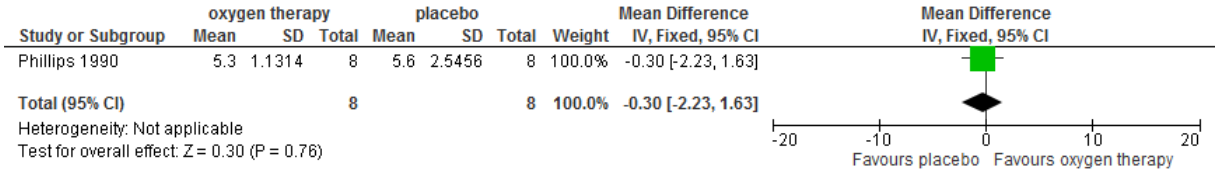


Figure 43: Rey figure (copy)

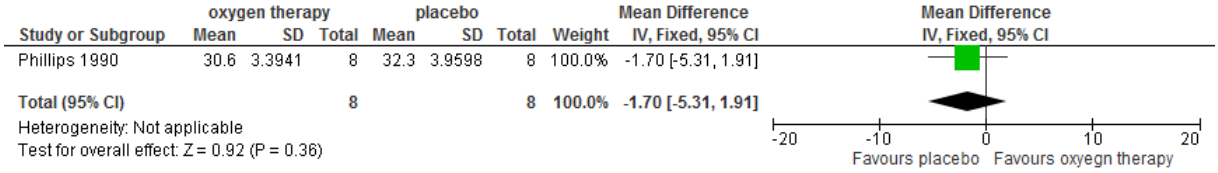


Figure 44: Rey figure (immediate recall)

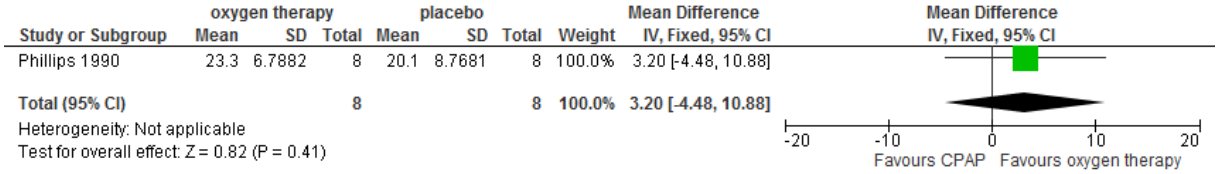


Figure 45: Rey figure (delayed recall)

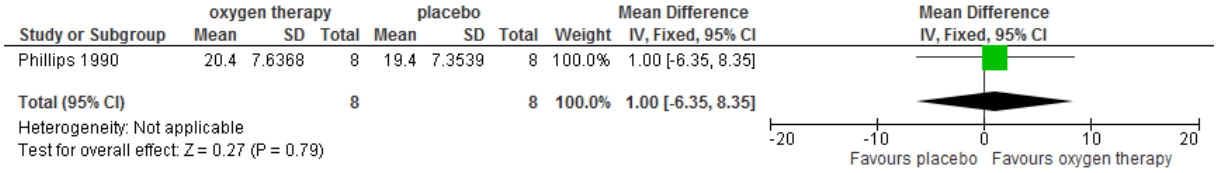


Figure 46: Finger tapping (dominant hand)

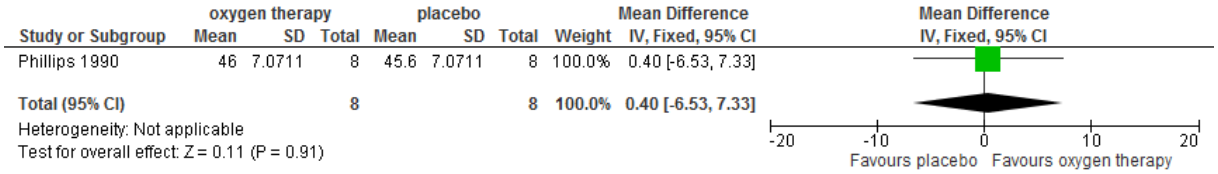
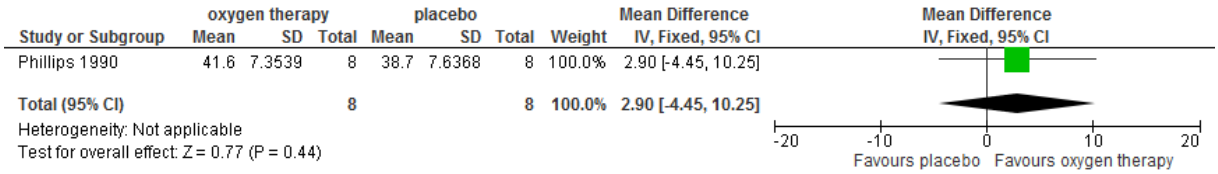


Figure 47: Finger tapping (non-dominant hand)



Appendix F: GRADE tables

Table 12: Clinical evidence profile: Oxygen therapy versus CPAP - moderate OSAHS

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Oxygen therapy versus CPAP	Control	Relative (95% CI)	Absolute		
Mortality (follow-up mean 12 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	None	0/101 (0%)	0%	RD 0.00 (-0.02 to 0.02)	0 events	⊕OOO VERY LOW	CRITICAL
Atrial Fibrillation (follow-up mean 12 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	2/101 (2%)	1%	RR 1.96 (0.18 to 21.28)	10 more per 1000 (from 8 fewer to 203 more)	⊕OOO VERY LOW	IMPORTANT
cardiovascular complications (follow-up mean 12 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	None	0/101 (0%)	0%	RD 0.00 (-0.02 to 0.02)	0 events	⊕OOO VERY LOW	IMPORTANT
number of motor vehicle accidents[5] (follow-up mean 12 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	None	1/101 (0.99%)	0%	OR 7.24 (0.14 to 365.16)	10 more (from 20 fewer to 40 more per 1000)	⊕OOO VERY LOW	IMPORTANT

24-hour mean systolic BP (follow-up mean 12 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	None	94	90	-	MD 3.5 higher (0.76 lower to 7.76 higher)	⊕OOO VERY LOW	IMPORTANT
SF36 - physical (follow-up mean 12 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	100	99	-	MD 0.5 lower (3.38 lower to 2.38 higher)	⊕OOO VERY LOW	CRITICAL
SF36 - mental (follow-up mean 12 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	None	100	99	-	MD 0.7 lower (3.49 lower to 2.09 higher)	⊕OOO VERY LOW	CRITICAL
SF36 - vitality (follow-up mean 12; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	No serious inconsistency	None	100	99	-	MD -2 lower (4.80 lower to 0.80 higher)	⊕OOO VERY LOW	CRITICAL
PHQ-9 (depression) (follow-up mean 12 weeks; range of scores: 0-27; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	No serious inconsistency	None	101	99	-	MD 0.7 higher (0.42 lower to 1.82 higher)	⊕OOO VERY LOW	IMPORTANT
ODI change score (follow-up mean 12 weeks; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	No serious inconsistency	None	101	99	-	MD 2.4 higher (1.37 lower to 6.17 higher)	⊕⊕OO LOW	IMPORTANT
Daytime mean systolic BP (follow-up mean 8 weeks; Better indicated by lower values)[6]												

2	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	None	95	93	-	MD 2.75 higher (2.33 lower to 7.82 higher)	⊕OOO VERY LOW	IMPORTANT
AHI (follow-up mean 4 weeks; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 13.8 higher (7.28 to 20.32 higher)	⊕OOO VERY LOW	IMPORTANT
Stamford Sleepiness Score (follow-up mean 4 weeks; range of scores: 1-7; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 0 higher (0.71 lower to 0.71 higher)	⊕OOO VERY LOW	IMPORTANT
Attention (2&7 test) (follow-up mean 4 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 11.5 lower (52.53 lower to 29.53 higher)	⊕OOO VERY LOW	IMPORTANT
Neurocognitive (digit symbol) (follow-up mean 4 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 1.6 lower (13.52 lower to 10.32 higher)	⊕OOO VERY LOW	IMPORTANT
neurocognitive (selective reminding - long term storage) (follow-up mean 4 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 4.8 lower (46.38 lower to 36.78 higher)	⊕OOO VERY LOW	IMPORTANT
neurocognitive (selective reminding - consistent retrieval) (follow-up mean 4 weeks; Better indicated by higher values)												

1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 7.8 lower (53.99 lower to 38.39 higher)	⊕OOO VERY LOW	IMPORTANT
neurocognitive - Benton visual Retention (follow-up mean 4 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	None	8	8	-	MD 0.8 lower (2.73 lower to 1.13 higher)	⊕OOO VERY LOW	IMPORTANT
neurocognitive - (Rey Figure - copy) (follow-up mean 4 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	None	8	8	-	MD 2.5 lower (5.56 lower to 0.56 higher)	⊕OOO VERY LOW	IMPORTANT
neurocognitive - (Rey Figure - immediate recall) (follow-up mean 4 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 1.5 lower (8.02 lower to 5.02 higher)	⊕OOO VERY LOW	IMPORTANT
neurocognitive - (Rey Figure - Delayed recall) (follow-up mean 4 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 2.2 lower (9.28 lower to 4.88 higher)	⊕OOO VERY LOW	IMPORTANT
neurocognitive - finger tapping (dominant hand) (follow-up mean 4 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 0.9 lower (7.3 lower to 5.5 higher)	⊕OOO VERY LOW	IMPORTANT
neurocognitive - finger tapping (non-dominant hand) (follow-up mean 4 weeks; Better indicated by higher values)												

1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 0.3 higher (6.91 lower to 7.51 higher)	⊕OOO VERY LOW	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] The majority of the evidence included an indirect population (downgrade by one increment) or a very indirect population (downgrade by two increments). A mixed severity OSHAS population was included based on mean AHI.
[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs. MID for systolic BP – 5mmhg.; Established MIDs for SF-36 physical/mental- 2/3.. . GRADE default MID (0.5XSD) usedfor all other continuous outcomes.
4. Outcomes 1-11 are from a 3-arm trial, results from the participants receiving the oxygen therapy are used in both the CPAP and placebo/no treatment comparisons
5. Peto odds ratio analysis used as there were zero events in one treatment arm.
6. GIV analysis used as cross over and parallel design RCTS combined

Table 13: Clinical evidence profile: Oxygen therapy versus no treatment/placebo - moderate OSAHS

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Oxygen therapy versus placebo/no treatment**	**Control**	**Relative (95% CI)**	**Absolute**		
Mortality (follow-up mean 12 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	None	0/101 (0%)	0%	RD 0.00 (-0.02 to 0.02)	0 events	⊕OOO VERY LOW	CRITICAL
Atrial fibrillation (follow-up mean 12 weeks)[5]												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	2/101 (2%)	1%	RR 2.00 (0.18 to 21.71)	30 more per 1000 (from 6 fewer to 425 more)	⊕OOO VERY LOW	IMPORTANT
cardiovascular complications[7] (follow-up mean 12 weeks)												

1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	None	0/101 (0%)	4%	OR 0.13 (0.02 to 0.95)	40 fewer per 1000 (from 80 fewer to 0 fewer)	⊕OOO VERY LOW	IMPORTANT
number of motor vehicle accidents[7] (follow-up mean 12 weeks)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	1/101 (0.99%)	0%	OR 7.39 (0.15 to 372.38)	10 more per 100 (from 20 fewer to 40 more)	⊕OOO VERY LOW	IMPORTANT
24-hour mean systolic BP (follow-up mean 12 weeks; Better indicated by lower values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	None	94	97	-	MD 2.2 higher (2.47 lower to 7.94 higher)	⊕OOO VERY LOW	IMPORTANT
SF36 physical (follow-up mean 12 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	None	100	99	-	MD 1.2 higher (1.56 lower to 3.96 higher)	⊕OOO VERY LOW	CRITICAL
SF36 - mental (follow-up mean 12 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	None	100	99	-	MD 2.2 higher (0.73 lower to 5.13 higher)	⊕OOO VERY LOW	CRITICAL
SF36 - vitality (follow-up mean 12 weeks; range of scores: 0-100; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	None	None	101	100	-	MD 0.3 higher (-2.4 lower to 2.84 higher)	⊕OOO VERY LOW	CRITICAL
PHQ-9 (depression) (follow-up mean 12 weeks; range of scores: 0-27; Better indicated by lower values)												

1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	None	101	100	-	MD 0.9 lower (2.01 lower to 0.21 higher)	⊕OOO VERY LOW	CRITICAL
ODI change score (follow-up mean 12 weeks; Better indicated by higher values)												
1	randomised trials	serious[1]	no serious inconsistency	serious[2]	None	None	101	101	-	MD 18.1 higher (14.32 to 21.88 higher)	⊕⊕OO LOW	IMPORTANT
daytime mean systolic BP (follow-up mean 8 weeks; Better indicated by lower values)[8]												
2	randomised trials	serious[1]	no serious inconsistency	serious[2]	serious[3]	None	95	97	-	MD 1.3 higher (4.13 lower to 6.74 higher)	⊕OOO VERY LOW	IMPORTANT
AHI (follow-up mean 4 weeks; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	None	8	8	-	MD 5.3 lower (18.11 lower to 7.51 higher)	⊕OOO VERY LOW	IMPORTANT
Stamford Sleepiness Score (follow-up mean 4 weeks; range of scores: 1-7; Better indicated by lower values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	None	8	8	-	MD 0.4 lower (1.11 lower to 0.31 higher)	⊕OOO VERY LOW	IMPORTANT
Attention (2&7 test) (follow-up mean 4 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 3.8 higher (36.96 lower to 44.56 higher)	⊕OOO VERY LOW	IMPORTANT
Neurocognitive (digit symbol) (follow-up mean 4 weeks; Better indicated by higher values)												

1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 1 higher (10.79 lower to 12.79 higher)	⊕OOO VERY LOW	IMPORTANT
neurocognitive (selective reminding - long term storage) (follow-up mean 4 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 6.2 higher (35.66 lower to 48.06 higher)	⊕OOO VERY LOW	IMPORTANT
neurocognitive (selective reminding - consistent retrieval) (follow-up mean 4 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 7 higher (35.89 lower to 49.89 higher)	⊕OOO VERY LOW	IMPORTANT
neurocognitive - Benton visual Retention (follow-up mean 4 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 0.3 lower (2.23 lower to 1.63 higher)	⊕OOO VERY LOW	IMPORTANT
neurocognitive - (Rey Figure - copy) (follow-up mean 4 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	serious[3]	None	8	8	-	MD 1.7 lower (5.31 lower to 1.91 higher)	⊕OOO VERY LOW	IMPORTANT
neurocognitive - (Rey Figure - Immediate recall) (follow-up mean 4 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 3.2 higher (4.48 lower to 10.88 higher)	⊕OOO VERY LOW	IMPORTANT
neurocognitive - (Rey Figure - Delayed recall) (follow-up mean 4 weeks; Better indicated by higher values)												

1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 1 higher (6.35 lower to 8.35 higher)	⊕OOO VERY LOW	IMPORTANT
neurocognitive - finger tapping (dominant hand) (follow-up mean 4 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 0.4 higher (6.53 lower to 7.33 higher)	⊕OOO VERY LOW	IMPORTANT
neurocognitive - finger tapping (non-dominant hand) (follow-up mean 4 weeks; Better indicated by higher values)												
1	randomised trials	very serious[1]	no serious inconsistency	serious[2]	very serious[3]	None	8	8	-	MD 2.9 higher (4.45 lower to 10.25 higher)	⊕OOO VERY LOW	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: The majority of the evidence included an indirect population (downgrade by one increment) or a very indirect population (downgrade by two increments). A mixed severity OSHAS population was included based on mean AHI.

[3] Downgraded by 1 increment if the confidence interval crossed one MID or by 2 increments if the confidence interval crossed both MIDs. MID for systolic BP – 5mmhg. Established MIDs for SF-36 physical/mental- 2/3. AHI- different severity groups, likely true MCID will vary, qualitatively considered in decision making throughout. . GRADE default MID (0.5XSD) used for all other continuous outcomes.

4. No events were reported in the control group while one event occurred in the O2 group. However, the absolute effect was not estimable on GRADE.

5.. Outcomes 1-11 are from a 3-arm trial, results from the participants receiving the oxygen therapy are used in both the CPAP and placebo/no treatment comparisons

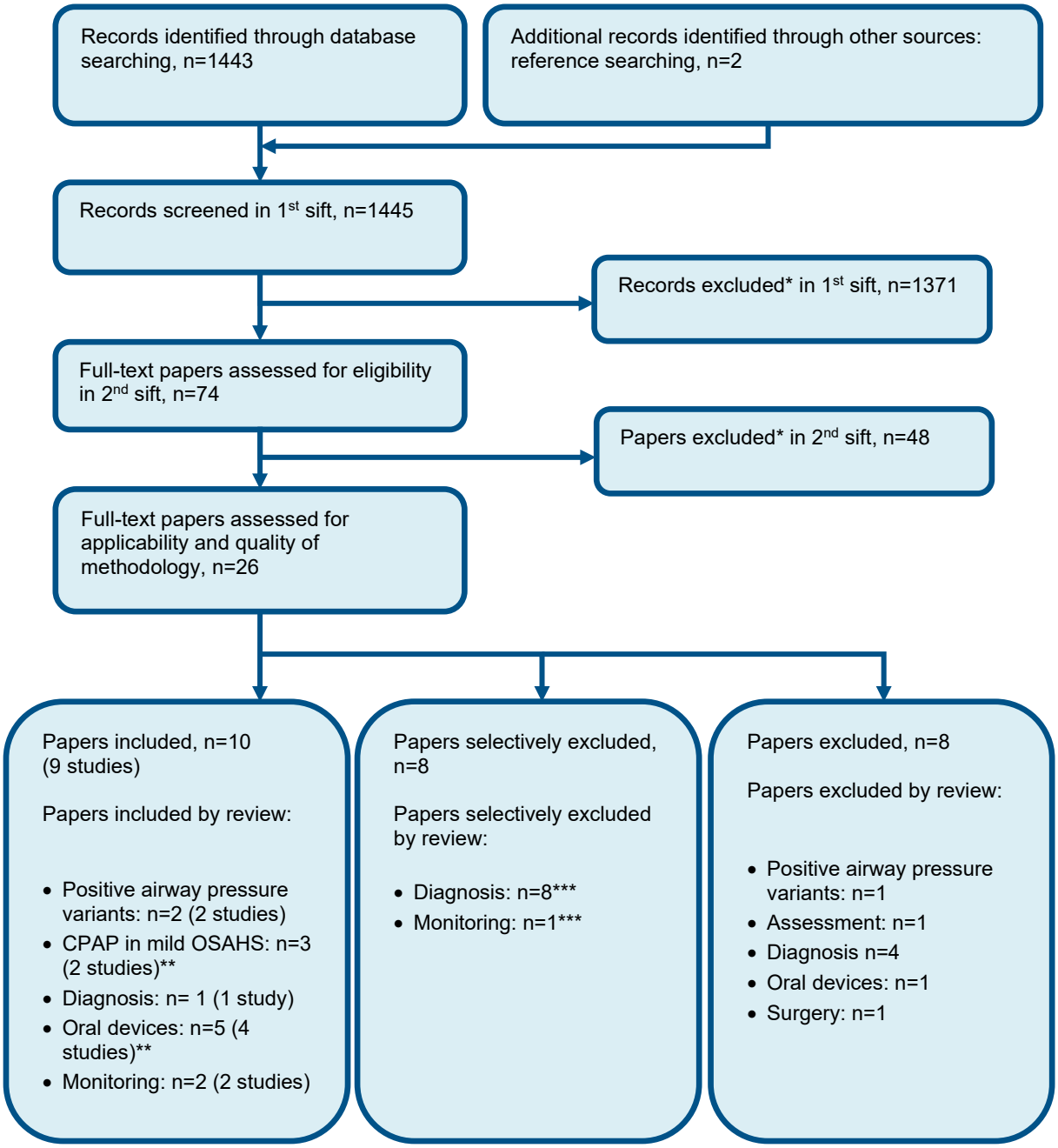
6. Outcome includes atrial fibrillation and 1 episode of unspecified tachyarrhythmia requiring hospitalisation in the placebo group

7. Peto odds ratio analysis used as there were zero events in one treatment arm.

8. GIV analysis used as cross over and parallel design RCTS combined

Appendix G: Health economic evidence selection

Figure 48: 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: Excluded studies

H.1 Excluded clinical studies

Table 14: Studies excluded from the clinical review for oxygen therapy (with CPAP/NIV)

Author	Exclusion reason
Bordier 2015[4]	Oxygen not given alongside CPAP/ non-invasive ventilation (NIV)
Bordier 2016[2]	Systematic review, references checked
Gottlieb 2014[10]	Oxygen not given alongside CPAP/ non-invasive ventilation (NIV)
Lewis 2017[21]	Oxygen not given alongside CPAP/ non-invasive ventilation (NIV)
Masa 2016[25]	Non-randomised study
Mehta 2013[27]	Systematic review, references checked
Murase 2016[32]	Wrong interventions
Turnbull 2019[58]	Oxygen not given alongside CPAP/ non-invasive ventilation (NIV)

Table 15: Studies excluded from the clinical review for oxygen therapy alone

Author	Exclusion reason
Bardwell 2007[1]	Less than minimum duration
Bordier 2013[3]	Conference abstract
Bordier 2015[4]	Includes patients with central sleep Apnoea
Bordier 2016[2]	Systematic review – references screened
Frohnhofen 1998[5]	Conference abstract
Frohnhofen 1998[6]	Not available in English
George 2018[7]	Conference abstract
Gold 1986[8]	Cross over study with no inclusion criteria or washout period and non-randomised
Gottlieb 2013[9]	Conference abstract
Hagenah 1996[11]	Not available in English
Hollier 2012[12]	Conference abstract
Hollier 2012[13]	Conference abstract
Hollier 2014[14]	Less than minimum duration
Hubatsch 2014[15]	Conference abstract
Janssens 2014[16]	Review article
Javaheri 1999[17]	Not guideline condition
Kasai 2010[18]	Incorrect interventions
Kempf 1991[19]	Not available in English
Landry 2017[20]	Wrong population – oxygen not given alone
Lewis 2014[22]	Conference abstract
Loredo 2006[23]	Less than minimum duration
Marrone 1992[24]	Less than minimum duration
Masa 2016[25]	Inappropriate study design
Mayos 2001[26]	Not available in English
Mehta 2012[28]	Systematic review - refs screened

Author	Exclusion reason
Mehta 2013[27]	Systematic review - refs screened
Mills 2006[29]	Less than minimum duration
Mostafavi 2017[30]	Less than minimum duration
Murase 2015[31]	Conference abstract
Murase 2016[32]	Not guideline condition
NCT 2010[34]	Trial website
NCT 2011[35]	Trial website
NCT 2013[37]	Trial website
NCT 2010 [36]	trials web page – citation only
Norman 2006[38]	Less than minimum duration
Pagel 2011[39]	Conference abstract
Patel 2013[40]	Conference abstract
Petousi 2018[41]	Conference abstract
Pokorski 2000[43]	Less than minimum duration
Prezerakos 2014[44]	Conference abstract
Quiroga 2016[45]	Conference abstract
Roffe 2014[46]	Systematic review. Not in English
Sasayama 2006[47]	Not review population
Scherz 2014[48]	Not guideline condition. Not in English
Selim 2018[49]	Wrong population
Shafazand 2014[50]	Inappropriate comparison. Incorrect interventions
Shigemitsu 2007[51]	Not guideline condition
Staniforth 1998[52]	Not guideline condition
Sugimura 2016[53]	Unavailable
Tan 2017[55]	Conference abstract
Tan 2019[54]	conference abstract – citation only
Teramoto 2003[56]	Less than minimum duration
Turnbull 2017[57]	Conference abstract
Turnbull 2018[59]	Conference abstract
Turnbull 2019[58]	Less than minimum duration
Ulrich 2013[60]	Conference abstract
Walsh 1995[61]	Incorrect interventions. Inappropriate comparison
Wijesinghe 2011[62]	Less than minimum duration
Yamamoto 2007[63]	Study retracted/not available

H.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:

None.

Appendix I: Research recommendations

I.1 Oxygen therapy

Research question:

What is the clinical and cost effectiveness of nocturnal oxygen compared to placebo in people with OSAHS intolerant to CPAP?

Why this is important:

Currently there is no evidence available for oxygen therapy in people with OSAHS intolerant of CPAP. The committee agreed that this was a difficult group to manage and hence there is a need for a robust evidence to inform if oxygen therapy would be effective in this population.

Criteria for selecting high-priority research recommendations:

PICO question	Population: People (16 and older) with OSAHS intolerant of CPAP. Intervention(s): nocturnal oxygen therapy Comparison: placebo Outcome(s): Critical • generic or disease specific quality of life measures (continuous) • mortality (dichotomous) Important • sleepiness scores (continuous, e.g. Epworth) • apnoea-Hypopnoea index or respiratory disturbance index • oxygen desaturation index • daytime pO2 • daytime pCO2 • daytime bicarbonate • nocturnal transcutaneous CO2 control • nocturnal oximetry • minor adverse effects of treatment • adherence • driving outcomes • neurocognitive outcomes • pulmonary artery pressure by Transthoracic echocardiography (TTE) • patient preference • impact on co-existing conditions: • HbA1c for diabetes • cardiovascular events for cardiovascular disease • systolic blood pressure for hypertension
Importance to patients or the population	There is lack of evidence for oxygen therapy in management of people with OSAHS intolerant of CPAP. This research will inform if oxygen therapy would be effective in this group of patients.This would give patients a therapeutic option if benefit is shown and may be easier for some patients to use than CPAP.
Relevance to NICE guidance	Currently there is lack of evidence to inform the recommendations. Further research would inform an update of the guideline.
Relevance to the NHS	A clear recommendation will offer clinicians clearer guidance on use of oxygen therapy in people with OSAHS intolerant to CPAP. There is a cost

	to the overnight use of oxygen, but if this is shown to improve health outcomes, which lead to less prescription of other medication eg antihypertensives, or fewer hospital admissions, this may be a worthwhile treatment.
National priorities	No
Current evidence base	No evidence was available for oxygen therapy in people with OSAHS intolerant to CPAP. The committee made research recommendation specifically for this population as they agreed from their experience that this was a difficult group to manage and further trials on this population could identify if oxygen therapy would be effective.
Equality	The recommendation is unlikely to impact on equality issues.
Study design	Randomised controlled trial with economic analysis.
Feasibility	The time scale will need to be at least 6 months to ensure adequate follow-up so that differences in outcomes can be seen between the groups.
Other comments	None
Importance	High: the research is essential to inform future updates of key recommendations in the guideline.