AHRQ Comparative Effectiveness Review Surveillance Program

CER #32:
Diagnosis and Treatment of Obstructive Sleep Apnea in Adults

Original release date:
August 8, 2011

Surveillance Report:
July, 2012

Key Findings:

- The conclusions for KQ2 (phased testing), KQ3 (preoperative screening), KQ6 (compliance predictors), and KQ7 (improving compliance) are still considered valid
- The KQ1 (diagnosis) conclusion regarding questionnaires is possibly out of date
- The conclusions for KQ4 (long-term outcomes) are still considered valid but additional studies are available
- The KQ5 (treatment) conclusions regarding CPAP are possibly out of date
- CPAP adverse events due to malfunctions should be reviewed systematically

Summary Decision

This CER's priority for updating is **Medium**

Authors:
Susanne Hempel, PhD
Sydne Newberry, PhD
Jennifer Schneider Chafen, MS, MD
Margaret Maglione, MPP
Aneesa Motala, BA
Jody Larkin, MLIS
Paul Shekelle, MD, PhD

None of the investigators has any affiliations or financial involvement that conflicts with the material presented in this report.

Acknowledgments

The authors gratefully acknowledge the following individuals for their contributions to this project:

Subject Matter Experts

Donna Arand, PhD
Kettering Sleep Disorders Center
Dayton, Ohio

Edward Grandi
American Sleep Apnea Association
Washington, District of Columbia

Jon-Erik Holty, MD
University of California, San Francisco
San Francisco, California

Edmundo Rubio, MD
Clinic Sleep Center
Roanoke, Virginia

Michael J. Sateia, MD
The Geisel School of Medicine at Dartmouth
Hanover, New Hampshire

Contents

Diagnosis and Treatment of Obstructive Sleep Apnea in Adults

1. Introduction

Comparative Effectiveness Review (CER) #32, Diagnosis and Treatment of Obstructive Sleep Apnea in Adults, was released in August 2011.[1] It was therefore due for a surveillance assessment in February, 2012. At that time, we contacted experts involved in the original CER and subject experts to get their opinions as to whether the conclusions had changed and need to be updated. We also conducted an update electronic literature search. During the assessment two articles published in JAMA[2, 3] received considerable media attention prompting us to update the search. Every month since the CER's original release, we received any FDA updates on the included treatments and tests.

2. Methods

2.1 Literature Searches

Using the search strategy employed for the original report, we conducted a limited literature search of Medline for the years 2010 to April 2012. This search included five high-profile general medical interest journals (Annals of Internal Medicine, British Medical Journal, Journal of the American Medical Association, Lancet, and the New England Journal of Medicine) and five specialty journals (Sleep, Chest, American Journal of Respiratory and Critical Care Medicine, Journal of Applied Physiology, and the European Respiratory Journal). The specialty journals were selected according to the search volume of publications on the topic in the last 30 years and most highly represented among the reference for the original report. Appendix A includes the search methodology for this topic.

2.2 Study selection

We used the same inclusion and exclusion criteria as the original CER. We screened the titles and abstracts and obtained full text copies of publications accordingly.

2.3 Expert Opinion

We shared the conclusions of the original report with 9 experts in the field (including the original project leader, suggested field experts, original technical expert panel (TEP) members, and peer reviewers) for their assessment of the need to update the report and their recommendations of any relevant new studies; 5 matter experts responded. Appendix C shows the questionnaire matrix that was sent to the experts.

2.4 Check for qualitative and quantitative signals

After abstracting the study conditions and findings for each new included study into an evidence table, we assessed whether the new findings provided a signal according to the Ottawa Method and/or the RAND Method, suggesting the need for an update. The criteria are listed in the table below.[4, 5]

	Ottawa Method
	Ottawa Qualitative Criteria for Signals of Potentially Invalidating Changes in Evidence
A1	Opposing findings: A pivotal trial or systematic review (or guidelines) including at least one new trial that characterized the treatment in terms opposite to those used earlier.
A2	Substantial harm: A pivotal trial or systematic review (or guidelines) whose results called into question the use of the treatment based on evidence of harm or that did not proscribe use entirely but did potentially affect clinical decision making.
A3	A superior new treatment: A pivotal trial or systematic review (or guidelines) whose results identified another treatment as significantly superior to the one evaluated in the original review, based on efficacy or harm.
	Criteria for Signals of Major Changes in Evidence
A4	Important changes in effectiveness short of "opposing findings"
A5	Clinically important expansion of treatment
A6	Clinically important caveat
A7	Opposing findings from discordant meta-analysis or nonpivotal trial
	Quantitative Criteria for Signals of Potentially Invalidating Changes in Evidence
B1	A change in statistical significance (from nonsignificant to significant)
B2	A change in relative effect size of at least 50 percent
	RAND Method Indications for the Need for an Update
1	Original conclusion is still valid and this portion of the original report does not need updating
2	Original conclusion is possibly out of date and this portion of the original report may need updating
3	Original conclusion is probably out of date and this portion of the original report may need updating
4	Original conclusion is out of date

2.5 Compilation of Findings and Conclusions

For this assessment we constructed a summary table that included the key questions, the original conclusions, and the findings of the new literature search, the expert assessments, and any FDA reports that pertained to each key question. To assess the conclusions in terms of the evidence that they might need updating, we used the 4-category scheme described in the table above for the RAND Method.

In making the decision to classify a CER conclusion into one category or another, we used the following factors when making our assessments:

- If we found no new evidence or only confirmatory evidence and all responding experts assessed the CER conclusion as still valid, we classified the CER conclusion as still valid.

- If we found some new evidence that might change the CER conclusion, and /or a minority of responding experts assessed the CER conclusion as having new evidence that might change the conclusion, then we classified the CER conclusion as possibly out of date.
- If we found substantial new evidence that might change the CER conclusion, and/or a majority of responding experts assessed the CER conclusion as having new evidence that might change the conclusion, then we classified the CER conclusion as probably out of date.
- If we found new evidence that rendered the CER conclusion out of date or no longer applicable, we classified the CER conclusion as out of date. Recognizing that our literature searches were limited, we reserved this category only for situations where a limited search would produce prima facie evidence that a conclusion was out of date, such as the withdrawal of a drug or surgical device from the market, a black box warning from FDA, etc.

2.6 Determining Priority for Updating

We used the following two criteria in making our final conclusion for this CER:

- How much of the CER is possibly, probably, or certainly out of date?
- How out of date is that portion of the CER? For example, would the potential changes to the conclusions involve refinement of original estimates or do the potential changes mean some therapies are no longer favored or may not exist? Is the portion of the CER that is probably or certainly out of date an issue of safety (a drug withdrawn from the market, a black box warning) or the availability of a new drug within class (the latter being less of a signal to update than the former)?

3. Results

3.1 Search

The literature search identified 226 titles. After title and abstract review, we further reviewed the full text of 47 journal articles. The remaining titles were rejected because they clearly did not meet inclusion criteria for any of the review questions. In addition to the electronic database searches, we followed up suggestions from the topic experts for studies not already included in the original report. We reference-mined articles that met inclusion criteria as well as systematic reviews identified by the literature searches to identify additional articles that may have been published since the publication of the report.

Thus, 86 articles went on to full text review. Of these, 51 articles were rejected because they did not meet the inclusion criteria of the original report. The remaining articles reporting on 35 studies were abstracted into evidence tables stratified by key question (Appendix B) for this assessment.[2, 3, 6-38] New pertinent studies were identified for key question 1 (diagnosis), key question 4 (long-term outcomes), key question 5 (treatment), and key question 7 (improving compliance).

3.2 Expert Opinion

Key question 1: One expert indicated that more evidence is now available for the Berlin questionnaire although the evidence may still be low, and more evidence is available for the STOP-Bang questionnaire meaning the evidence may not be insufficient anymore to draw definitive conclusions. The five experts were in agreement that none of the other conclusions changed based on new evidence.

Key Question 3: One expert indicated that there is new evidence, one was not sure and three thought the conclusions are still valid.

Key question 5: One expert thought there is new evidence for outcomes not included in the conclusion for the comparison of continuous positive airway pressure and control. One expert was not sure whether the comparison of CPAP and sham CPAP and the comparison of other treatments were still valid, all other experts thought they were. One expert also indicated that not all pertinent studies were included in the original report.

Key Question 2, 4, 6 and 7: All five experts thought the conclusions are still valid.

3.3 Identifying qualitative and quantitative signals

Table 1 shows the original key questions, the conclusions of the original report, the results of the literature and drug database searches, the experts' assessments, the recommendations of the Southern California Evidence-based Practice Center (SCEPC) regarding the need for update, and qualitative signals.

Table 1: Summary Table

Conclusions From CER Executive Summary	RAND Literature Search	FDA/ Health Canada/MHRA (UK)	Expert Opinion EPC Investigator Other Experts	Conclusion from SCEPC
Key Question 1. How do different available tests compare in their ability to diagnose sleep apnea in adults with symptoms suggestive of disordered sleep? How do these tests compare in different subgroups of patients based on: race, sex, body mass index, existing non-insulin-dependent diabetes mellitus, existing cardiovascular disease, existing hypertension, clinical symptoms, previous stroke, or airway characteristics? PSG devices are classified as Type I monitors. Portable monitors are classified as either Type II, which record all the same information as PSG; Type III, which do not differentiate between whether the patient is asleep or awake, but have at least two respiratory channels (two airflow channels or one airflow and one effort channel); or Type IV, which fail to fulfill criteria for Type III monitors but usually record more than two bioparameters.				
Comparison of Portable Devices and Polysomnography				
The strength of evidence is moderate, among 15 quality A, 45 quality B, and 39 quality C studies, that Type III and Type IV monitors may have the ability to accurately predict AHI suggestive of OSA with high positive likelihood ratios and low negative likelihood ratios for various AHI cutoffs in PSG. Type III monitors perform better than Type IV monitors at AHI cutoffs of 5, 10, and 15 events/hr. Analysis of difference versus average analyses plots suggest that substantial differences in the measured AHI may be encountered between PSG and both Type III and Type IV monitors. Large differences compared with in-laboratory PSG cannot be excluded for all portable monitors. The evidence is insufficient to adequately compare specific monitors to each other.	The search identified 4 new studies. One assessed a home respiratory polygraph (Type III monitor; sensitivity 73%, specificity 77%). One assessed pulse transit time (Type IV monitor) and concluded that the data does not allow to use it as a screening tool. One study concluded that oximetry and nasal flow have equivalent accuracy, and 1 study concluded that a device recording oximetry and nasal pressure areas has high diagnostic utility.	None relevant	All 5 experts thought this conclusion was still valid.	The conclusions are still valid but additional studies are available.
No recent studies compared Type II monitors with PSG. A prior Technology Assessment of home diagnosis of OSA concluded that "based on [three quality B studies], type II monitors [used at home] may identify AHI suggestive of OSA with high positive likelihood ratios and low negative likelihood ratios," though "substantial differences in the [measurement of] AHI may be encountered between type II monitors and facility-based PSG."	No new studies were identified.	None relevant	All 5 experts thought this conclusion was still valid.	The conclusions are still valid.
Comparison of Questionnaires and Polysomnography				
Of the six studies reviewed (one quality A, one quality B, four quality C), *the strength of evidence is low* among three studies supporting the use of the Berlin questionnaire in screening for sleep	We identified 10 new studies. Six used the Berlin questionnaire in various patient populations. Four studies assessed the STOP-	None relevant	4/5 experts thought the conclusions do not change but one expert indicated that more evidence is now available for the	The conclusions are possibly out of date There is now more evidence

Conclusions From CER Executive Summary	RAND Literature Search	FDA/ Health Canada/MHRA (UK)	Expert Opinion EPC Investigator Other Experts	Conclusion from SCEPC
definitive conclusions concerning the use of the STOP, STOP-Bang, ASA Checklist, Epworth Sleepiness Scale, and Hawaii Sleep questionnaires to screen for sleep apnea because each questionnaire was assessed in only a single study.	assessed the Epworth Sleepiness Scale, 1 assessed an Obstructive Sleep Apnea/Hypopnea Syndrome (OSAHS) score, and 1 assessed the STOP questionnaire.		more evidence is available for the STOP-Bang questionnaire meaning the evidence may not be insufficient anymore to draw conclusions.	questionnaires possibly changing the strength of evidence evaluation.
Clinical Prediction Rules and Polysomnography				
The strength of evidence is low among seven studies (three quality A, three quality B, and one quality C) that some clinical prediction rules may be useful in the prediction of a diagnosis of OSA. Ten different clinical prediction rules have been described. Nine clinical prediction rules have been used for the prediction of a diagnosis of OSA (using different criteria). The oropharyngeal morphometric model gave near perfect discrimination (area under the curve [AUC] = 0.996) to predict the diagnosis of OSA, and the pulmonary function data model had 100 percent sensitivity with 84 percent specificity to predict diagnosis of OSA. The remaining models reported lower diagnostic sensitivities and specificities. Each model was deemed useful to predict the diagnoses of OSA by the individual study authors. However, while all the models were internally validated, external validation of these predictive rules has not been conducted in the vast majority of the studies.	One new study was identified evaluating the Dixon model. The sensitivity and specificity was 75 and 57%.	None relevant	All 5 experts thought this conclusion was still valid.	The conclusion is still valid but an additional study is available.
Key Question 2. How does phased testing (screening tests or battery followed by full test) compare to full testing alone?				
The strength of evidence is insufficient to determine the utility of phased testing, followed by full testing when indicated, to diagnose sleep apnea, as only one study that met our inclusion criteria investigated this question. This prospective quality C study did not fully analyze the phased testing, thus the sensitivity and specificity of 6 the phased strategy could not be calculated due to a verification bias; not all participants received PSG (full) testing.	No new studies were identified.	None relevant	All 5 experts thought this conclusion was still valid.	The conclusions are still valid.

Conclusions From CER Executive Summary	RAND Literature Search	FDA/ Health Canada/MHRA (UK)	Expert Opinion EPC Investigator Other Experts	Conclusion from SCEPC
Key Question 3. What is the effect of preoperative screening for sleep apnea on surgical outcomes?				
The strength of evidence is insufficient regarding postoperative outcomes with mandatory screening for sleep apnea. Two quality C prospective studies assessed the effect of preoperative screening for sleep apnea on surgical outcomes. One study found no significant differences in outcomes between patients undergoing bariatric surgery who had mandatory PSG or PSG based on clinical parameters. The second study found that general surgery patients willing to undergo preoperative PSG were more likely to have perioperative complications, particularly cardiopulmonary complications, possibly suggesting that patients willing to undergo PSG are more ill than other patients.	No new studies were identified.	None relevant	One expert indicated that there is new evidence, one was not sure, 3 thought the conclusions were still valid.	The conclusions are still valid.
Key Question 4. In adults being screened for obstructive sleep apnea, what are the relationships between apnea-hypopnea index or oxygen desaturation index, and other patient characteristics with respect to long-term clinical and functional outcomes?				
The strength of evidence is high from four studies (three quality A, one quality B) indicating that an AHI >30 events/hr is an independent predictor of all-cause mortality; although one study found that this was true only in men under age 70. All other outcomes were analyzed by only one or two studies. Thus, only a *low strength of evidence* exists that a high AHI (>30 events/hr) is associated with incident diabetes. This association, however, may be confounded by obesity, which may result in both OSA and diabetes. *The strength of evidence is insufficient* regarding the association between AHI and other clinical outcomes. The two studies of cardiovascular mortality did not have consistent findings, and the two studies of hypertension had unclear conclusions. One study of nonfatal cardiovascular disease found a significant association with baseline AHI (as they did for cardiovascular mortality). One study each found no association between AHI and stroke or long-term quality of life.	Two additional studies were identified. A further analysis of the Wisconsin Sleep Cohort Study reported that sleep disordered breathing is associated with increased cancer mortality and all-cause mortality. A further analysis of the Zaragoza Sleep Cohort study in patients with sleep-disordered breathing concluded that compared with participants without obstructive sleep apnea, the presence of obstructive sleep apnea was associated with increased adjusted risk of incident hypertension but treatment with CPAP therapy was associated with a lower risk of hypertension.	None relevant	All 5 experts thought this conclusion was still valid.	The conclusions are still valid but additional studies have been published, one reporting on a new outcome (cancer mortality).

Conclusions From CER Executive Summary	RAND Literature Search	FDA/ Health Canada/MHRA (UK)	Expert Opinion EPC Investigator Other Experts	Conclusion from SCEPC
Key Question 5. What is the comparative effect of different treatments for obstructive sleep apnea in adults? **a. Does the comparative effect of treatments vary based on presenting patient characteristics, severity of obstructive sleep apnea, or other pretreatment factors? Are any of these characteristics or factors predictive of treatment success?** **• Characteristics: age, sex, race, weight, bed partner, airway, other physical characteristics, and specific comorbidities** **• Obstructive sleep apnea severity or characteristics: baseline questionnaire (and similar tools) results, formal testing results (including hypoxemia levels), baseline quality of life, positional dependency** **• Other: specific symptoms** **b. Does the comparative effect of treatments vary based on the definitions of obstructive sleep apnea used by study investigators?**				
Comparison of Continuous Positive Airway Pressure and Control				
There are 22 trials (11 each of quality B and C) that provide sufficient evidence supporting large improvements in sleep measures with continuous positive airway pressure (CPAP) compared with control. There is only weak evidence that demonstrated no consistent benefit in improving quality of life, neurocognitive measures, or other intermediate outcomes. Despite no evidence or weak evidence for an effect of CPAP on clinical outcomes, given the large magnitude of effect on the intermediate outcomes AHI and ESS, the strength of evidence that CPAP is an effective treatment to alleviate sleep apnea signs and symptoms was rated *moderate*.	We identified 5 new RCTs. Three assessed sleepiness and found an improvement. Two RCTs assessed hypertension; the large multicenter study did not report statistically significant reductions for hypertension or cardiovascular events. Three RCTs reported blood pressure reductions. Two RCTs assessed neurological or neurocognitive outcomes, only 1 showed an improvement. Quality of life was assessed in 1 RCT, no improvement was reported.	Three MAUDE Adverse Event Reports for CPAP have been submitted since the publication of the report. A patient experienced failure of a new air supply hose and woke up gasping for air; a patient reported a mask air leak which may have resulted in excessive dryness in the right eye, triggering recurrent corneal erosion; and a patient was hospitalized with hypercarbia due to a defective machine.	4/5 experts thought the conclusions are still valid, 1 expert thought there is new evidence for outcomes not mentioned in the conclusion (mortality, cardiovascular events) and that some studies have been missed that should have been included.	The conclusion regarding clinical outcomes is possibly out of date. New studies are now available, including a large multicenter study, some of the new studies report on additional clinical outcomes not previously covered, and the adverse event signals should be followed up.
Comparison of CPAP and Sham CPAP				
There are 24 trials (5 quality A, 13 quality B, 6 quality C) that provide sufficient evidence supporting large improvements in sleep measures with CPAP compared with sham CPAP, but weak evidence of possibly no difference between CPAP and sham CPAP in improving quality of life, neurocognitive measures, or other intermediate outcomes. Despite no evidence or weak evidence for an effect of CPAP on clinical outcomes, given the large magnitude of effect on the intermediate outcomes of AHI, ESS, and arousal index, the strength of evidence that CPAP is an effective treatment for the relief of signs and symptoms of sleep apnea was rated *moderate*.	Three new RCTs were identified. One reported a reduction in hypertension, one in metabolic syndrome and one in postprandial lipidemia. Two reported reductions in sleepiness and one improved quality of life.	See above	4/5 experts thought the conclusions are still valid, 1 expert was not sure.	The overall conclusions are still valid but new studies reporting clinical improvements are available.

Conclusions From CER Executive Summary	RAND Literature Search	FDA/ Health Canada/MHRA (UK)	Expert Opinion EPC Investigator Other Experts	Conclusion from SCEPC
Comparison of Oral and Nasal CPAP				
Three small trials (one quality B, two quality C) with inconsistent results preclude any substantive conclusions concerning the efficacy of oral (or full face mask) versus nasal CPAP in improving compliance in patients with OSA. Largely due to small sample size, the reported effect estimates in the studies reviewed were generally imprecise. Thus, overall, *the strength of evidence is insufficient* regarding differences in compliance or other outcomes between oral and nasal CPAP.	No new studies were identified.	See above	All 5 experts thought this conclusion was still valid.	The conclusions are still valid.
Comparison of Autotitrating CPAP and Fixed CPAP				
The strength of evidence is moderate that autotitrating CPAP (autoCPAP) and fixed pressure CPAP result in similar levels of compliance (hours used per night) and treatment effects for patients with OSA. Twenty-one studies (1 quality A, 10 quality B, 10 quality C) comprising an experimental population of over 800 patients provided evidence that autoCPAP reduces sleepiness as measured by ESS by approximately 0.5 points more than fixed CPAP. The two devices were found to result in similar compliance and changes in AHI from baseline, quality of life, and most other sleep study measures. However, there is also evidence that minimum oxygen saturation improves more with fixed CPAP than with autoCPAP, although by only about one percent. Evidence is limited regarding the relative effect of fixed CPAP and autoCPAP on blood pressure. There were no data on objective clinical outcomes.	Two new RCTs were identified showing therapeutic equivalence.	See above	All 5 experts thought this conclusion was still valid.	The conclusions are still valid but additional studies are available.
Comparison of Bilevel CPAP and Fixed CPAP				
The strength of evidence is insufficient regarding any difference in compliance or other outcomes between bilevel CPAP and fixed CPAP. Five small, highly clinically heterogeneous trials (one quality B, four quality C) with largely null findings did not support any substantive differences in the efficacy of bilevel CPAP versus fixed CPAP in the treatment of patients with OSA.	The search identified 1 new study. The RCT concluded there was no statistically significant difference in adherence between the auto bilevel and CPAP groups but patients with poor initial CPAP exposure may still achieve an acceptable long-term	See above	All 5 experts thought this conclusion was still valid.	The conclusions are still valid but an additional study is available.

Conclusions From CER Executive Summary	RAND Literature Search	FDA/ Health Canada/MHRA (UK)	Expert Opinion EPC Investigator Other Experts	Conclusion from SCEPC
Largely due to small sample sizes, the studies mostly had imprecise estimates of the comparative effects.	clinical outcome. Improvements in functional outcomes, sleepiness and fatigue complaints were comparable.			
Comparison of Flexible Bilevel CPAP and Fixed CPAP				
The strength of evidence is insufficient regarding the relative merits of flexible bilevel CPAP and fixed CPAP as there was only one quality B study that investigated this comparison. This study found that flexible bilevel CPAP may yield increased compliance (use ≥4 hr/night) compared with fixed CPAP.	No new studies were identified.	See above	All 5 experts thought this conclusion was still valid.	The conclusions are still valid.
Comparison of C-Flex™ and Fixed CPAP				
No statistically significant differences in compliance or other outcomes were found between C-Flex and fixed CPAP. *The strength of evidence is low* for this finding because of the mixed quality (Bs and Cs) of the four primary studies.	One new RCT was identified showing no superiority of C-Flex over fixed CPAP. In addition, 1 study comparing A-Flex with fixed CPAP was identified. The RCT showed equivalency but non-superiority (except for average leak values) in efficacy, adherence, and functional outcomes.	See above	All 5 experts thought this conclusion was still valid.	The conclusions regarding C-Flex are still valid but we identified a study comparing another next generation device (A-Flex CPAP).
Comparison of Humidification in CPAP				
The strength of evidence is insufficient to determine whether there is a difference in compliance or other outcomes between positive airway pressure treatment with and without humidification. Five trials examined different aspects of humidified CPAP treatment for patients with OSA. While some studies reported a benefit of added humidity in CPAP treatment in improving patient compliance, this effect was not consistent across all the studies. Overall, the studies were clinically heterogeneous, small, and of quality B (three studies) or C (two studies).	One new small crossover RCT was identified reporting that the addition of heated humidification decreases nasal resistance and mucosal inflammation.	One MAUDE Adverse Event Report indicated that a patient lost his sense of smell while using a CPAP humidifier.	All 5 experts thought this conclusion was still valid.	The conclusions are still valid, but an additional study is available and the adverse event signal should be followed up.
Comparison of Mandibular Advancement Devices and No Treatment or Inactive Oral Devices				
The strength of evidence is moderate to show that the use of mandibular advancement devices (MAD) improves sleep apnea signs and symptoms. Five trials (four quality B, one quality	No new studies were identified.	None relevant	All 5 experts thought this conclusion was still valid.	The conclusions are still valid.

Conclusions From CER Executive Summary	RAND Literature Search	FDA/ Health Canada/MHRA (UK)	Expert Opinion EPC Investigator Other Experts	Conclusion from SCEPC
C) compared MAD with no treatment, using a variety of different types of MAD, and found significant improvements with MAD in AHI, ESS, and other sleep study measures. Any differences in quality of life measures or neurocognitive tests were equivocal between treatment groups. No trial evaluated objective clinical outcomes. Another five trials (four quality B, one quality C) compared the effects of MAD with inactive oral devices and reported similar findings.				
Comparison of Different Oral Devices				
The strength of evidence is insufficient to draw conclusions with regard to the relative efficacy of different types of oral MAD in patients with OSA because the reviewed studies were generally small, and each was concerned with a unique comparison. Five studies (four quality B, one quality C) with unique comparisons found little to no differences between different types and methods of use of MAD or other oral devices in sleep study or sleepiness measures. No study evaluated objective clinical outcomes. Only one study evaluated compliance; no significant differences were observed. One trial found that a greater degree of mandibular advancement resulted in an increased number of patients achieving an AHI <10 events/hr; however, the mean AHI was similar between treatment groups.	No new studies were identified.	None relevant	All 5 experts thought this conclusion was still valid.	The conclusions are still valid.
Comparison of Mandibular Advancement Devices and CPAP				
The strength of evidence is moderate that CPAP is superior to MAD in improving sleep study measures. Ten mostly quality B trials overall found that CPAP resulted in greater reductions in AHI and arousal index, and increases in minimum oxygen saturation. The evidence regarding the relative effects on ESS were too heterogeneous to allow conclusions. In a single study, patients were more compliant with MAD than CPAP (hours used per night and nights used). No study evaluated objective clinical outcomes. *The strength of evidence is insufficient* to address	No new studies were identified.	None relevant	All 5 experts thought this conclusion was still valid.	The conclusions are still valid.

Conclusions From CER Executive Summary	RAND Literature Search	FDA/ Health Canada/MHRA (UK)	Expert Opinion EPC Investigator Other Experts	Conclusion from SCEPC
which patients might benefit most from either treatment.				
Comparison of Surgery and Control				
The strength of evidence is insufficient to evaluate the relative efficacy of surgical interventions for the treatment of OSA. Six trials and one nonrandomized prospective study with unique interventions compared surgery with control treatment for the management of patients with OSA. Three studies were rated quality A, one quality B, and three quality C. The results were inconsistent across studies as to which outcomes were improved with surgery compared with no or sham surgery.	No new studies were identified.	None relevant	All 5 experts thought this conclusion was still valid.	The conclusions are still valid.
Comparison of Surgery and CPAP				
The strength of evidence is insufficient to determine the relative merits of surgical treatments versus CPAP. Of 12 studies (1 quality A, 11 quality C) comparing surgical modalities with CPAP, only two were RCTs, and they compared CPAP with uvulopalatopharyngoplasty (UPPP), removal of the soft tissue at the back of the throat, the uvula, and soft palate. While one of these trials found that CPAP resulted in a higher mortality benefit, the other found no difference between groups. Due to the heterogeneity of interventions and outcomes examined, the variability of findings across studies, and the inherent bias of all but one study regarding which patients received surgery, it is not possible at this time to draw useful conclusions comparing surgical interventions with CPAP in the treatment of patients with OSA. The quality A trial was the only unbiased comparison of surgery and CPAP (patients had previously received neither treatment). It did not find statistically significant differences in ESS and quality of life measures between patients with mild to moderate OSA who had temperature-controlled radiofrequency tissue volume reduction of the soft palate and those who had CPAP at 2 months followup. Likewise, the other trial,	No new studies were identified	None relevant	All 5 experts thought this conclusion was still valid.	The conclusions are still valid.

Conclusions From CER Executive Summary	RAND Literature Search	FDA/ Health Canada/MHRA (UK)	Expert Opinion EPC Investigator Other Experts	Conclusion from SCEPC
comparing maxillomandibular advancement osteotomy and CPAP, did not find statistically significant differences in AHI and ESS in patients with severe OSA. For the nonrandomized studies, comparisons between surgery and CPAP are difficult to interpret since baseline patient characteristics (including sleep apnea severity) differed significantly between groups, particularly in regards to what previous treatments patients had. The reported findings on sleep study and quality of life measures were heterogeneous across studies.				
Comparison of Surgery and Mandibular Advancement Devices				
The strength of evidence is insufficient regarding the relative merit of MAD versus surgery in the treatment of OSA, as there was only one study (quality B) that examined this question. A statistically significant improvement in AHI was observed in the MAD group compared with the surgery group. No study evaluated objective clinical outcomes.	No new studies were identified.	None relevant	All 5 experts thought this conclusion was still valid.	The conclusions are still valid.
Comparison of Other Treatments				
The strength of evidence is low to show that some intensive weight loss programs may be effective treatment for OSA in obese patients. Three trials (one quality A, two quality B) compared weight loss interventions with control interventions. All three trials found significant relative reductions in AHI with diet. Other outcomes were inconsistent. *The strength of evidence is insufficient* to determine the effects of other potential treatments for OSA. Twentyone studies evaluated other interventions including atrial overdrive pacing, eight different drugs, palatal implants, oropharyngeal exercises, a tongue-retaining device, a positional alarm, combination tongue-retaining device and positional alarm, bariatric surgery, nasal dilator strips, acupuncture, and auricular plaster. All of these interventions were evaluated by one or two studies only. The findings were heterogeneous. No study evaluated objective	Four new studies were identified. One RCT tested a nasal expiratory device and reported reductions in AHI and sleepiness compared to sham. One additional study on acetazolamide was identified and reported positive results during an altitude sojourn while CPAP was discontinued. One RCT investigated valsartan and reported a fourfold higher decrease in blood pressure than CPAP One RCT showed that ondansetron and fluoxetine reduces AHI..	None relevant	4/5 experts thought this conclusion was still valid, 1 was not sure	The conclusions are still valid but new studies are available.

Conclusions From CER Executive Summary	RAND Literature Search	FDA/ Health Canada/MHRA (UK)	Expert Opinion EPC Investigator Other Experts	Conclusion from SCEPC
clinical outcomes.				
Key Question 6. In OSA patients prescribed nonsurgical treatments, what are the associations of pretreatment patient-level characteristics with treatment compliance?				
Across five studies (one quality A, one quality B, three quality C), *the strength of evidence is moderate* that more severe OSA as measured by higher AHI is associated with greater compliance with CPAP use. Each study measured compliance differently, including thresholds of 1, 2, or 3 hours of use per night or as a continuous variable, and undefined "objective compliance" measured by the device. *The strength of evidence is moderate* that a higher ESS score is also associated with improved compliance. *There are low strengths of evidence* that younger age, snoring, lower CPAP pressure, higher BMI, higher mean oxygen saturation, and the sleepiness domain on the Grenoble Sleep Apnea Quality of Life test are each possible independent predictors of compliance. It is important to note, however, that selective reporting, particularly of nonreporting of nonsignificant associations, cannot be ruled out. The heterogeneity of analyzed and reported potential predictors greatly limits these conclusions. Differences across studies as to which variables were independent predictors may be due to the adjustment for different variables, in addition to differences in populations, outcomes, CPAP machines, and CPAP training and followup. One quality C study of mandibular advancement devices failed to identify potential predictors of compliance.	No new studies were identified.	None relevant	All 5 experts thought this conclusion was still valid.	The conclusions are still valid.
Key Question 7. What is the effect of interventions to improve compliance with device (positive airway pressure, oral appliances, positional therapy) use on clinical and intermediate outcomes?				
The strength of evidence is low that some specific adjunct interventions may improve CPAP compliance, but studies are heterogeneous and no general type of intervention (e.g., education, telemonitoring) was more promising than others. The 18 trials (two quality A, eight quality B, and	One study comparing CPAP and A-Flex CPAP (comfort feature) was identified but did not show superiority in efficacy, adherence or functional outcomes.	None relevant	All 5 experts thought this conclusion was still valid.	The conclusions are still valid but an additional study is available.

Conclusions From CER Executive Summary	RAND Literature Search	FDA/ Health Canada/MHRA (UK)	Expert Opinion EPC Investigator Other Experts	Conclusion from SCEPC
eight quality C) had inconsistent effects across a wide variety of interventions. Studies generally had small sample sizes with less than 1 year of followup. Compared with usual care, several interventions were shown to significantly increase hours of CPAP use per night in some studies. These included intensive support or literature (designed for patient education), cognitive behavioral therapy (given to patients and their partners), telemonitoring, and a habit-promoting audio-based intervention. However, the majority of studies did not find a significant difference in CPAP compliance between patients who received interventions to promote compliance with device use and those who received usual care. No study of nurseled care (which was not focused primarily on compliance) showed an effect on compliance rates.				

Legend: AHI: Apnea Hypopnea Index in events/hour of sleep, CPAP: Continuous Positive Airway Pressure

References

1. Balk EM, Moorthy D, Obadan NO, et al. Diagnosis and Treatment of Obstructive Sleep Apnea in Adults. Comparative Effectiveness Review No. 32. (Preparted by Tufts Evidence-based Practice Center under Contract No. 290-2007-10055-1). AHRQ Publication No. 11-EHC052-EF Agency for Healthcare Research and Quality. Rockville, MD: July 2011.

2. Barbe F, Duran-Cantolla J, Sanchez-de-la-Torre M, et al. Effect of continuous positive airway pressure on the incidence of hypertension and cardiovascular events in nonsleepy patients with obstructive sleep apnea: a randomized controlled trial. JAMA. 2012 May 23;307(20):2161-8. PMID 22618923.

3. Marin JM, Agusti A, Villar I, et al. Association between treated and untreated obstructive sleep apnea and risk of hypertension. JAMA. 2012 May 23;307(20):2169-76. PMID 22618924.

4. Shekelle PG, Newberry SJ, Maglione M, et al. Assessment of the Need to Update Comparative Effectiveness Reviews: Report of an Initial Rapid Program Assessment (2005-2009) (Prepared by the Southern California Evidence-based Practice Center). Rockville, MD: Agency for Healthcare Research and Quality; October 2009.

5. Shojania KG, Sampson M, Ansari MT, et al. How quickly do systematic reviews go out of date? A survival analysis. Ann Intern Med. 2007 Aug 21;147(4):224-33. PMID 17638714.

6. Bakker J, Campbell A, Neill A. Randomized controlled trial comparing flexible and continuous positive airway pressure delivery: effects on compliance, objective and subjective sleepiness and vigilance. Sleep. 2010 Apr;33(4):523-9. PMID 20394322.

7. Berry RB, Kryger MH, Massie CA. A novel nasal expiratory positive airway pressure (EPAP) device for the treatment of obstructive sleep apnea: a randomized controlled trial. Sleep. 2011 Apr;34(4):479-85. PMID 21461326.

8. Bloch KE, Latshang TD, Nussbaumer-Ochsner Y, et al. Equivalence of autocpap and fixed cpap in the long-term treatment of obstructive sleep apnea syndrome: a randomized, controlled multicente [Abstract]. Chest; 2010. p. 886a.

9. Chouchou F, Sforza E, Celle S, et al. Pulse t time in screening sleep disordered breathing in elderly population: the PROOF-SYNAPSE stuc Sleep. 2011 Aug;34(8):1051-9. PMID 2180466

10. Duran-Cantolla J, Aizpuru F, Montserrat JI al. Continuous positive airway pressure as treat for systemic hypertension in people with obstru sleep apnoea: randomised controlled trial. BMJ 2010;341:c5991. PMID 21106625.

11. Epstein MD, Segal LN, Ibrahim SM, et al. Snoring and the risk of obstructive sleep apnea patients with pulmonary embolism. Sleep. 2010 Aug;33(8):1069-74. PMID 20815188.

12. Koutsourelakis I, Vagiakis E, Perraki E, et Nasal inflammation in sleep apnoea patients usi CPAP and effect of heated humidification. Euro Respiratory Journal. 2011 Mar;37(3):587-94. P 20595158.

13. Kushida CA, Berry RB, Blau A, et al. Posit airway pressure initiation: a randomized contro trial to assess the impact of therapy mode and titration process on efficacy, adherence, and outcomes. Sleep. 2011 Aug;34(8):1083-92. PM 21804670.

14. Masa JF, Corral J, Pereira R, et al. Therape decision-making for sleep apnea and hypopnea syndrome using home respiratory polygraphy: a multicentric study. American Journal of Respira & Critical Care Medicine. 2011 Oct 15;184(8):9 71. PMID 21737584.

15. Parra O, Sanchez-Armengol A, Bonnin M, Early treatment of obstructive apnoea and strok outcome: a randomised controlled trial. Europea Respiratory Journal. 2011 May;37(5):1128-36. 20847081.

16. Pepin J-L, Tamisier R, Barone-Rochette G, Comparison of continuous positive airway pres and valsartan in hypertensive patients with sleep apnea. American Journal of Respiratory & Criti Care Medicine. 2010 Oct 1;182(7):954-60. PM 20522795.

17. Phillips CL, Yee BJ, Marshall NS, et al. Continuous positive airway pressure reduces postprandial lipidemia in obstructive sleep apnea: a randomized, placebo-controlled crossover trial. American Journal of Respiratory & Critical Care Medicine. 2011 Aug 1;184(3):355-61. PMID 21527567.

18. Prasad B, Radulovacki M, Olopade C, et al. Prospective trial of efficacy and safety of ondansetron and fluoxetine in patients with obstructive sleep apnea syndrome. Sleep. 2010 Jul;33(7):982-9. PMID 20614859.

19. Rofail LM, Wong KKH, Unger G, et al. Comparison between a single-channel nasal airflow device and oximetry for the diagnosis of obstructive sleep apnea. Sleep. 2010 Aug;33(8):1106-14. PMID 20815194.

20. Sert Kuniyoshi FH, Zellmer MR, Calvin AD, et al. Diagnostic accuracy of the Berlin Questionnaire in detecting sleep-disordered breathing in patients with a recent myocardial infarction. Chest. 2011 Nov;140(5):1192-7. PMID 21596794.

21. Sharma SK, Agrawal S, Damodaran D, et al. CPAP for the metabolic syndrome in patients with obstructive sleep apnea. New England Journal of Medicine. 2011 Dec 15;365(24):2277-86. PMID 22168642.

22. Silva GE, Vana KD, Goodwin JL, et al. Identification of patients with sleep disordered breathing: comparing the four-variable screening tool, STOP, STOP-Bang, and Epworth Sleepiness Scales. J Clin Sleep Med. 2011 Oct 15;7(5):467-72. PMID 22003341.

23. Farney RJ, Walker BS, Farney RM, et al. The STOP-Bang equivalent model and prediction of severity of obstructive sleep apnea: relation to polysomnographic measurements of the apnea/hypopnea index. J Clin Sleep Med. 2011 Oct 15;7(5):459-65B. PMID 22003340.

24. Powell ED, Gay PC, Ojile JM, et al. A pilot study assessing adherence to auto-bilevel following a poor initial encounter with CPAP. J Clin Sleep Med. 2012 Feb 15;8(1):43-7. PMID 22334808.

25. Nussbaumer-Ochsner Y, Latshang TD, Ulrich S, et al. Patients with obstructive sleep apnea syndrome benefit from acetazolamide during an altitude sojourn: a randomized, placebo-controlled, double-blind trial. Chest. 2012 Jan;141(1):131-8. PMID 21659435.

26. Thurtell MJ, Bruce BB, Rye DB, et al. The Berlin questionnaire screens for obstructive sleep apnea in idiopathic intracranial hypertension. J Neuroophthalmol. 2011 Dec;31(4):316-9. PMID 21537196.

27. Srijithesh PR, Shukla G, Srivastav A, et al. Validity of the Berlin Questionnaire in identifying obstructive sleep apnea syndrome when administered to the informants of stroke patients. J Clin Neurosci. 2011 Mar;18(3):340-3. PMID 21237648.

28. Gantner D, Ge JY, Li LH, et al. Diagnostic accuracy of a questionnaire and simple home monitoring device in detecting obstructive sleep apnoea in a Chinese population at high cardiovascular risk. Respirology. 2010 Aug;15(6):952-60. PMID 20624255.

29. Chung F, Subramanyam R, Liao P, et al. High STOP-Bang score indicates a high probability of obstructive sleep apnoea. Br J Anaesth. 2012 May;108(5):768-75. PMID 22401881.

30. Ong TH, Raudha S, Fook-Chong S, et al. Simplifying STOP-BANG: use of a simple questionnaire to screen for OSA in an Asian population. Sleep Breath. 2010 Dec;14(4):371-6. PMID 20419474.

31. Friedman M, Wilson MN, Pulver T, et al. Screening for obstructive sleep apnea/hypopnea syndrome: subjective and objective factors. Otolaryngol Head Neck Surg. 2010 Apr;142(4):531-5. PMID 20304273.

32. Kolotkin RL, LaMonte MJ, Walker JM, et al. Predicting sleep apnea in bariatric surgery patients. Surg Obes Relat Dis. 2011 Sep-Oct;7(5):605-10. PMID 21684219.

33. Drager LF, Pedrosa RP, Diniz PM, et al. The effects of continuous positive airway pressure on prehypertension and masked hypertension in men with severe obstructive sleep apnea. Hypertension. 2011 Mar;57(3):549-55. PMID 21242462.

34. Ryan CM, Bayley M, Green R, et al. Influence of continuous positive airway pressure on outcomes of rehabilitation in stroke patients with obstructive sleep apnea. Stroke. 2011 Apr;42(4):1062-7. PMID 21372306.

35. Lozano L, Tovar JL, Sampol G, et al. Continuous positive airway pressure treatment in sleep apnea patients with resistant hypertension: a randomized, controlled trial. J Hypertens. 2010 Oct;28(10):2161-8. PMID 20577130.

36. Nieto FJ, Peppard PE, Young T, et al. Sleep disordered breathing and cancer mortality: results from the Wisconsin Sleep Cohort Study. Am J Respir Crit Care Med. 2012 May 20PMID 22610391.

37. Masa JF, Corral J, Pereira R, et al. Effectiveness of home respiratory polygraphy for the diagnosis of sleep apnoea and hypopnoea syndrome. Thorax. 2011 Jul;66(7):567-73. PMID 21602541.

38. Bakker J, Campbell A, Neill A. Randomised controlled trial of auto-adjusting positive airway pressure in morbidly obese patients requiring high therapeutic pressure delivery. J Sleep Res. 2011 Mar;20(1 Pt 2):233-40. PMID 20545835.

Appendices

Appendix A: Search Methodology

Appendix B: Evidence Tables

Appendix C: Questionnaire Matrix

Appendix A. Search Methodology

Journal ranking by number of sleep apnea publications

SLEEP	2271
CHEST	1144
AMERICAN JOURNAL OF RESPIRATORY AND CRITICAL CARE MEDICINE	834
JOURNAL OF APPLIED PHYSIOLOGY	544
EUROPEAN RESPIRATORY JOURNAL	534
SLEEP MEDICINE	455
JOURNAL OF SLEEP RESEARCH	426
THORAX	378
AMERICAN REVIEW OF RESPIRATORY DISEASE	350
OTOLARYNGOLOGY HEAD AND NECK SURGERY	320
SLEEP AND BREATHING	309
LARYNGOSCOPE	296
CIRCULATION	242
JOURNAL OF HYPERTENSION	211
RESPIRATORY PHYSIOLOGY NEUROBIOLOGY	185
NEUROLOGY	176
RESPIRATION	174
REVUE DES MALADIES RESPIRATOIRES	174
INTERNATIONAL JOURNAL OF PEDIATRIC OTORHINOLARYNGOLOGY	171

Search strategy

(replication of search employed for original report)

1 exp Sleep Apnea Syndromes/ or exp Sleep Apnea, Obstructive
2 exp Airway Resistance/
3 exp snoring/
4 Upper airway resistance syndrome.mp.
5 Respiratory disturbance.mp.
6 obstructive sleep apn?ea.mp.
7 or/1-6
8 randomized controlled trial.pt.
9 controlled clinical trial.pt.
10 randomized controlled trials/
11 Random Allocation/
12 Double-blind Method/
13 Single-Blind Method/
14 clinical trial.pt.
15 Clinical Trials.mp. or exp Clinical Trials/
16 (clinic$ adj25 trial$).tw.
17 ((singl$ or doubl$ or trebl$ or tripl$) adj (mask$ or blind$)).tw.
18 Placebos/
19 placebo$.tw.
20 random$.tw.
21 trial$.tw.
22 (latin adj square).tw.
23 Comparative Study.tw. or Comparative Study.pt.
24 exp Evaluation studies/
25 Follow-Up Studies/
26 Prospective Studies/
27 (control$ or prospectiv$ or volunteer$).tw.
28 Cross-Over Studies/
29 Or/8-28
30 exp Positive-Pressure Respiration/ or exp Continuous Positive Airway Pressure/
31 exp Intermittent Positive-Pressure Ventilation/ or exp Ventilators, Mechanical/ or exp Masks/
32 general surgery/ or neurosurgery/ or otolaryngology/ or surgery, plastic/ or thoracic surgery/
33 Surgical Procedures, Operative/
34 oral appliances.mp.
35 exp Physical Therapy Modalities/ or exp Exercise Therapy/
36 positional therapy.mp.
37 exp Weight Loss/
38 exp Exercise/ or exp Exercise Therapy/
39 exp Therapeutics/
40 exp anesthesia/ or pre-operative screening/ or Anesthetic agents/
41 sleep apnea, obstructive/th
42 *tonsillectomy/
43 or/30-42
44 exp Polysonography/
45 exp Oximetry
46 exp Monitoring, physiologic/
47 pulse transit time.mp.
48 exp Monitoring, Ambulatory/
49 peripheral aterial tonometry.mp.
50 exp Questionnaires/
51 Diagnostic Tests, Routine/
52 (Epworth OR Stanford OR Berlin OR Pittsburgh OR scale).af.
53 (friedman OR surgical OR standing).mp.

54 STOP-Bang.af.
55 sleep apnea, obstructive/di
56 or/44-55
57 exp Laboratory Techniques/ and Procedures/
58 56 or 57
59 exp “sensitivity and specificity”/
60 exp predictive value of tests
61 exp ROC CURVE
62 exp mass screening
63 exp diagnosis/
64 exp reproducibility of results/
65 exp false negative reactions/ OR false positive reactions/
66 predictive value.tw.
67 (sensitivity or specificity).tw.
68 accuracy.tw.
69 screen$.tw.
70 diagno$.tw.
71 roc.tw.
72 reproducib$.tw.
73 (false positive or false negative).tw.
74 likelihood ratio.tw.
75 accuracy.tw.
76 di.fs.
77 or/59-76
78 7 and 29 and 43
79 limit 78 to English language
80 limit 79 to humans
81 79 and humans.sh.
82 80 or 81
83 remove duplicates from 82
84 7 and 43
85 84 not 83
86 limit 85 to English Language
87 limit 86 to humans
88 86 and humans.sh.
89 87 OR 88
90 remove duplicates from 89
91 limit 90 to (addresses or bibliography or biography or case reports or comment or congresses or consensus development conference or dictionary or directory or festschrift or in vitro or interactive tutorial or interview or lectures or legal cases or legislation or news or newspaper article or overall or patient education handout or periodical index or portraits or "scientific integrity review" or twin study)
92 90 not 91
93 7 and 29 and 58
94 limit 93 to English language
95 limit 94 to humans
96 94 and humans.sh.
97 95 or 96
98 remove duplicates from 97
99 98 not (83 or 92)
100 7 and 58 and 77
101 limit 100 to English language
102 limit 101 to humans
103 101 and humans.sh.
104 102 or 103
105 104 not (83 or 92 or 99)
106 remove duplicates from 105

107 limit 7 to (guideline or meta analysis or practice guideline)
108 7 and Cochrane database or systematic reviews.jn.
109 107 or 108
110 remove duplicates from 109
111 110 not (83 or 92 or 99 or 106)
112 83 or 92 or 99 or 106 or 111
113 7 and 29
114 113 not 112
115 (ep or co or mo).fs.
116 (incidence or longitudinal studies or prospective studies or survival analysis or follow-up studies or logistic models or proportional hazards models or linear models or regression analysis).sh.
117 exp patient compliance/ or exp medication adherence/ or exp treatment refusal/
118 or/115-117
119 114 and 118
120 limit 119 to (English language and humans)
121 112 or 120
122 exp orthodontic appliances, removable/
123 palate, soft/su or pharynx/su or uvula/su or sleep apnea syndromes/su
124 sleep apnea syndromes/pc
125 or/ 122-124
126 7 and 29 and 125
127 126 not (112 or 119)
128 limit 127 to (English language and humans)
129 121 or 128
130 limit 129 to yr = "2010 – Current"

Citations were limited to those in these journals:
Annals of Internal Medicine
New England Journal of Medicine
Journal of the American Medical Association
Lancet
British Medical Journal
Sleep
Chest
American Journal of Respiratory and Critical Care Medicine
Journal of Applied Physiology
European Respiratory Journal

Latest search date: 4/17/2012
Retrieved citations: 226

Appendix B. Evidence Tables

Evidence Table Key Question 1. How do different available tests compare in their ability to diagnose sleep apnea in adults with symptoms suggestive of disordered sleep? How do these tests compare in different subgroups of patients based on: race, sex, body mass index, existing non-insulin-dependent diabetes mellitus, existing cardiovascular disease, existing hypertension, clinical symptoms, previous stroke, or airway characteristics?

Study Description	Test (Type)	Population	Finding
Comparison of Portable Devices and Polysomnography			
Masa, 2011 (Therapeutic decision-making...)[14], Masa, 2011 (Effectiveness of home..)[37]; Spain	Home Respiratory Polygraph (Type III monitor)	N=366, patients with intermediate or high sleep apnea-hypopnea syndrome suspicion	Therapeutic decisions had a sensitivity of 73%, a specificity of 77%, an agreement level (sum of true positives and negatives) of 76%. Sensitivity, specificity, and agreement increased in patients with higher apnea-hypopnea index scores. Home respiratory polygraphy is an alternative to polysomnography in patients with suspected sleep apnea hypopnea syndrome
Chouchou, 2011[9]; France	Pulse Transit Time measurement (Type IV monitor)	N=780, healthy elderly volunteers free of cardiac and neurologic disease (mean age 68.6, SD 1,0)	ROC curves for apnea-hypopnea index ≥15 defined an area under the curve of 0.67 and a cutoff point to autonomic arousal index 32.3 events per hours. Sensitivity 70.5%, specificity 55%. Specificity but not sensitivity increased when predicting scores ≥30.
Gantner, 2010[28]; Shanghai	ApneaLink device (oximetry, nasal pressure recordings; Type IV monitor), Berlin Questionnaire	N=143, patients with high cardiovascular morbidity	ApneaLink recordings of oximetry and nasal pressure areas had high diagnostic utility.
Rofail, 2010[19]; Australia	Oximetry (Type IV monitor), nasal airflow	N=105, sleep clinic patients with suspected obstructive sleep apnea	Nasal flow and oximetry have equivalent (high) accuracy for diagnosing obstructive sleep apnea in the home setting.
Comparison of Questionnaires and Polysomnography			
Epstein, 2010[11]; USA	Berlin Questionnaire	N=23, patients with pulmonary embolism	Sensitivity 100%, specificity 91%, positive predictive value 92% and negative predictive value 100%
Sert Kuniyoshi, 2011[20]; USA	Berlin Questionnaire	N=99, patients with recent myocardial infarction	The sensitivity was 0.68, the specificity 0.34, positive predictive value 0.68, negative predictive value 0.50, positive likelihood ratio 1.27, negative likelihood ratio 0.68, overall diagnostic accuracy was 63%.
Srijithesh, 2011[27]; India	Berlin Questionnaire adapted for caregiver	N=121, acute stroke patients	Sensitivity was 67, specificity 56%, positive predictive value 63%, negative predictive value 59%.

Study Description	Test (Type)	Population	Finding
Thurtell, 2011[26]; USA	Berlin Questionnaire	N=30, patients with newly diagnosed idiopathic intracranial hypertension	The sensitivity and specificity were 83% and 58%, the positive predictive value was 75%; a low-risk score identifies patients who are unlikely to have obstructive sleep apnea, polysomnography should be considered in those with a high-risk score.
Friedman, 2010[31]; USA	Berlin Questionnaire, Epworth Sleepiness Scale, Obstructive Sleep Apnea/Hypopnea Syndrome (OSAHS) score, visual analog scale for snoring, new screening method incorporating subjective and objective factors	N=223, patients referred to sleep laboratory	In predicting a high apnea hypopnea index, the sensitivity and specificity, were 0.615 and 0.226 for the Berlin, 0.863 and 0.468 for OSAHS score, and 0.82 and 0.834 for new screening method.
Gantner, 2010[28]; Shanghai	Berlin Questionnaire, ApneaLink device (oximetry, nasal pressure recordings)	N=143, patients with high cardiovascular morbidity	The Berlin questionnaire had low overall diagnostic accuracy.
Chung, 2012[29]; Canada	STOP-Bang	N=746, patients scheduled for inpatient surgery	The predicted probabilities for moderate/severe obstructive sleep apnea increased from 0.36 to 0.60 as the test score increased from 3 to 7 and 8.
Farney, 2011[23]; USA	STOP-Bang modification	N=1426, patients undergoing diagnostic polysomnography	The STOP-Bang model may be useful to categorize obstructive sleep apnea severity, triage patients for diagnostic evaluation or exclude them.
Ong, 2010[30]; Singapore	STOP-Bang, shortened STOP-Bang	N=348, patients undergoing diagnostic polysomnography	Sensitivity was 86, 93, and 96 for different cutoffs with negative predictive values of 85, 93 for moderate and severe obstructive sleep apnea.
Silva, 2011[22]; USA	STOP, STOP-Bang, Epworth Sleepiness Scales, 4-Variable screening tool	N=4,770, patients at risk for sleep disordered breathing	STOP-Bang questionnaire had higher sensitivity to predict moderate-to-severe (87%) and severe (70%) sleep disordered breathing, while the 4-Variable screening tool had higher specificity to predict moderate –to-severe and severe sleep disordered breathing.
Clinical Prediction Rules and Polysomnography			
Kolotkin, 2011[32]; USA	Dixon model	N=310, bariatric surgery patients	A sensitivity of 75% and specificity of 57% was seen in this patient group

Note: ROC: Receiver operating characteristic curves

Evidence Table Key Question 4. In adults being screened for obstructive sleep apnea, what are the relationships between apnea-hypopnea index or oxygen desaturation index, and other patient characteristics with respect to long-term clinical and functional outcomes?

Study Description	Design	Population	Predictors and Covariates	Outcome	Findings
Marin, 2012[3]; Spain	Prospective cohort study, multivariable analysis	N=1889, Zaragoza Sleep Cohort	Obstructive sleep apnea severity, CPAP therapy, non-modifiable risk factors (e.g., age, sex), blood pressure, Body Mass Index (BMI), modifiable risk factors (alcohol use, smoking status, hyperlipidemia, lipid-lowering drugs, glucose, triglycerides, lipoprotein cholesterol, menopausal status),	New-onset hypertension	Compared with participants without obstructive sleep apnea, the presence of obstructive sleep apnea was associated with increased adjusted risk of incident hypertension; treatment with CPAP therapy was associated with a lower risk of hypertension.
Nieto, 2012[36]; USA	Prospective cohort study, multivariable analysis	N=1,522, Wisconsin Sleep Cohort	Apnea Hypopnea Index range, age, gender, Body Mass Index (BMI), obesity, smoking, physical activity, alcohol use, education, diabetes, waist circumference, sleep duration, CPAP	All-cause mortality, cancer mortality	Baseline sleep disordered breathing is associated with increased cancer mortality in a community-based sample.

Note: CPAP: Continuous Positive Airway Pressure

Evidence Table Key Question 5. What is the comparative effect of different treatments for obstructive sleep apnea in adults?
a. Does the comparative effect of treatments vary based on presenting patient characteristics, severity of obstructive sleep apnea, or other pretreatment factors? Are any of these characteristics or factors predictive of treatment success?
• Characteristics: age, sex, race, weight, bed partner, airway, other physical characteristics, and specific comorbidities
• Obstructive sleep apnea severity or characteristics: baseline questionnaire (and similar tools) results, formal testing results (including hypoxemia levels), baseline quality of life, positional dependency
• Other: specific symptoms
b. Does the comparative effect of treatments vary based on the definitions of obstructive sleep apnea used by study investigators?

Study Description	Design	Population	Intervention and comparator	Outcome	Findings
Comparison of Continuous Positive Airway Pressure (CPAP) and Control					
Barbe, 2012[2]; Spain	Parallel RCT	N=725, patients with apnea hypopnea index >= 20 and Epworth sleepiness scale <= 10	CPAP Vs. no active intervention	Systemic hypertension, cardiovascular events, hospitalization, severity of obstructive sleep apnea, sleepiness (Epworth Sleepiness Scale)	In patients with obstructive sleep apnea without daytime sleepiness, the prescription of CPAP compared with usual care did not result in a statistically significant reduction in the incidence of hypertension or cardiovascular events, however, the study may have had limited power to detect a significant difference. Both groups had changes in sleepiness scores, but these were significantly less in the control group ($p<0.001$)
Drager, 2011[33]; Brazil	Parallel RCT	N=36, male patients with untreated severe obstructive sleep apnea and prehypertension and masked hypertension	CPAP (Respironics) Vs. no treatment	Blood pressure, frequency of prehypertension and masked hypertension	Effective CPAP therapy promotes significant reduction in the frequency of prehypertension and masked hypertension by promoting significant blood pressure reductions.
Lozano, 2010[35]; Spain	Parallel RCT	N=96, patients with obstructive sleep apnea and resistant hypertension	CPAP plus conventional treatment vs. conventional treatment	Systolic and diastolic blood pressure, sleepiness (Epworth Sleepiness Scale)	In patients with resistant hypertension and obstructive sleep apnea, CPAP treatment for 3 months achieves reductions in 24-h blood pressure. The Sleepiness showed no decrease from baseline to the follow-up in the conventional treatment group but improved with CPAP.
Parra, 2011[15]; Spain	Parallel RCT	N=71, stroke patients with apnea hypopnea index >=20	Nasal CPAP plus conventional stroke treatment Vs. conventional stroke treatment only	Neurologic improvement, quality of life, occurrence of new cardiovascular events, mortality	Early use of CPAP appears to accelerate neurological recovery and to delay the appearance of cardiovascular events, although an improvement in survival or quality of life was not shown.

Study Description	Design	Population	Intervention and comparator	Outcome	Findings
Ryan, 2011[34]; Canada	Parallel RCT	N=44, stroke patients with obstructive sleep apnea	CPAP plus standard stroke occupational and physiotherapy vs. standard stroke occupational and physiotherapy alone	Sleep outcomes, sleepiness (Epworth Sleepiness scale, Stanford Sleepiness scale); functional and motor outcomes; neuropsychological outcomes, depression	The treatment in stroke patients undergoing rehabilitation improved functional and motor, but not neurocognitive outcomes. There were significant reductions in sleepiness compared to control.
Comparison of CPAP and Sham CPAP					
Duran-Cantolla, 2010[10]; Spain	Parallel RCT	N=340, patients recently diagnosed with systemic hypertension and apnea hypopnea index > 15	CPAP Vs. sham CPAP (very low pressure)	Blood pressure, sleepiness (Epworth scale), quality of life (EuroQol)	CPAP produced a statistically significant reduction in blood pressure in patients with systemic hypertension and obstructive sleep apnea but the reduction is small and may have uncertain clinical relevance. Sleepiness showed larger improvements in the CPAP group ($p<0.001$). The quality of life improved only in the CPAP group ($p=0.01$)
Sharma, 2011[21]; India	Crossover RCT	N=86, patients with obstructive sleep apnea syndrome	CPAP Vs. sham CPAP (mask containing escape holes)	Metabolic syndrome variables, anthropometric variables, adverse events	In patients with moderate-to-severe obstructive sleep apnea syndrome, 3 months of CPAP therapy lowers blood pressure and partially reverses metabolic abnormalities.
Phillips, 2011[17]; Australia	Crossover RCT	N=38, patients with obstructive sleep apnea in the upper moderate or severe range	CPAP Vs. placebo CPAP (identical appearance)	Postprandial lipidemia, triglycerides concentration, lipid profiles in triglycerides, cholesterol, free fatty acids, urinary catecholamines; sleepiness (Epworth Sleepiness Scale, Functional Outcomes of Sleep Questionnaire); sleep apnea	Treatment of severe obstructive sleep apnea with CPAP improves postprandial triglycerides and total cholesterol levels; effects may reduce the risk for cardiovascular events. Significant differences in sleep apnea events and one out of two sleepiness questionnaires.
Comparison of Autotitrating CPAP and Fixed CPAP					
Bakker, 2011[38]New Zealand	Crossover RCT	N=12, morbidly obese patients requiring high therapeutic pressure delivery	Auto-adjusting PAP Vs. CPAP	Residual apnea hypopnea index, pressure, leaks, compliance, comfort, sleepiness (Epworth Sleepiness Scale), treatment preference	Both therapies substantially reduced the apnea hypopnea index.
Bloch, 2010[8]; Switzerland	Parallel RCT	N=105, obstructive sleep apnea patients	AutoCPAP Vs. fixed CPAP	Sleep resistance time, Sleepiness (Epworth Sleepiness Scale), quality of	The data demonstrate therapeutic equivalence of autoCPAP and fixed CPAP during at least one year in terms

Study Description	Design	Population	Intervention and comparator	Outcome	Findings
				life (SF-36), apnea hypopnea index, blood pressure, oxygen saturation, treatment adherence	of improving symptoms, quality of life, breathing disturbances, vigilance, and blood pressure.
Comparison of Bilevel CPAP and Fixed CPAP					
Powell, 2012[24]; USA	Parallel RCT	N=51, patients with poor initial CPAP experience	Auto-bilevel Vs. CPAP	Subjective daytime functioning, functional impact of sleepiness, sleepiness, fatigue severity, impact on daily living	Patients with a poor initial CPAP exposure may still achieve an acceptable long-term clinical outcome; both groups demonstrated comparable significant improvements in functional outcomes, sleepiness, and fatigue complaints over the treatment period.
Comparison of C-Flex™ and Fixed CPAP					
Bakker, 2010[6]; New Zealand	Parallel RCT	N=76, patients with severe obstructive sleep apnea	C-Flex Vs. CPAP	Sleep latency, reaction time and number of lapses, comfort	In patients with severe obstructive sleep apnea both CPAP and C-Flex resulted in substantial improvements in sleepiness, vigilance, and quality of life; neither treatment appeared superior.
Kushida, 2011[13]; USA	Parallel RCT	N=168, patients with moderate to severe obstructive sleep apnea	A-Flex, pressure automatically adjusted, during inhaling as well as exhaling, comfort feature Vs. first A-Flex then fixed pressure Vs. fixed CPAP CPAP	Residual apnea, key polysomnography variables, functional outcomes, subjective sleepiness, vigilance, attitudes toward positive airway pressure, acceptance of therapy	A-Flex shows equivalency, but non-superiority (except for average leak values), in efficacy, adherence, and functional outcomes compared to CPAP after 3 and 6 months.
Comparison of Humidification in CPAP					
Koutsourelakis, 2011[12]; Greece	Cross-over RCT	N=20, patients with obstructive sleep apnea	CPAP plus heated humidification Vs. CPAP plus sham-heated humidification	Nasal symptom questionnaire, anterior rhinomanometry, nasal lavage, nasal mucosa biopsies	Nasal obstruction of obstructive sleep apnea patients on CPAP treatment is inflammatory in origin and the addition of heated humidification decreases nasal resistance and mucosal inflammation.
Comparison of Other Treatments					
Berry, 2011[7]; USA	Parallel RCT	N=250, patients with obstructive sleep apnea, apnea hypopnea index >=10	Nasal expiratory positive airway pressure device Vs. sham device (similar in appearance)	Oxygenation, sleep architecture, impact of position and sleep stage, reduction in apnea hypopnea index, subjective sleepiness, adverse	The device significantly reduced the apnea hypopnea index and improved subjective daytime sleepiness compared to the sham treatment in patients with mild to severe obstructive sleep apnea

Study Description	Design	Population	Intervention and comparator	Outcome	Findings
				events	with excellent adherence.
Nussbaumer-Ochsner, 2012[25]; USA	Crossover RCT	N=45, patients with obstructive sleep apnea during an altitude sojourn	Acetazolamide Vs. placebo	Polysomnography variables, sleep, blood pressure, pulse oximetry, body weight, bigilance, mountain sickness, perceived insomnia, sleepiness, side effects	In patients with obstructive sleep apnea discontinuing CPAP during an altitude sojourn, acetazolamide improves oxygenation, breathing disturbances, and sleep quality by stimulating ventilation. Patiehts may benefit from acetazolamide at altitude if CPAP therapy is not feasible.
Pepin, 2010[16]; France	Crossover RCT	N=23, hypertensive patients with obstructive sleep apnea	CPAP Vs. valsartan	Blood pressure, biological parameters	Although the blood pressure decrease was significant with CPAP treatment, valsartan induced a fourfold higher decrease in mean 24-hour blood pressure than CPAP in untreated hypertensive patients with obstructive sleep apnea.
Prasad, 2010[18]; USA	Parallel RCT	N=25, adult patients with apnea hypopnea index > 10	Ondansetron and Fluoxetine vs. Ondansetron only vs Fluoxetine only Vs. placebo	Apnea hypopnea reduction, oximetry, sleep architecture	Combined treatment with ondansetron and fluoxetine is well-tolerated and reduces apnea hypopnea index values, yielding a potentially therapeutic response in some subjects with obstructive sleep apnea.

Note: CPAP: Continuous Positive Airway Pressure

Evidence Table Key Question 7. What is the effect of interventions to improve compliance with device (positive airway pressure, oral appliances, positional therapy) use on clinical and intermediate outcomes?

Study Description	Design	Population	Intervention	Findings
Kushida, 2011[13]; USA	Parallel RCT	N=168, patients with moderate to severe obstructive sleep apnea	A-Flex, pressure automatically adjusted, during inhaling as well as exhaling (comfort feature) Vs. first A-Flex then fixed pressure Vs. fixed CPAP CPAP	A-Flex shows equivalency, but non-superiority (except for average leak values), in efficacy, adherence, and functional outcomes compared to CPAP after 3 and 6 months.

Note: CPAP: Continuous Positive Airway Pressure

Appendix C. Questionnaire Matrix

Surveillance and Identification of Triggers for Updating Systematic Reviews for the EHC Program

Title: **Diagnosis and Treatment of Obstructive Sleep Apnea in Adults**

Conclusions From CER Executive Summary	Is this conclusion almost certainly still supported by the evidence?	Has there been new evidence that may change this conclusion?	Do Not Know
Key Question 1. How do different available tests compare in their ability to diagnose sleep apnea in adults with symptoms suggestive of disordered sleep? How do these tests compare in different subgroups of patients based on: race, sex, body mass index, existing non-insulin-dependent diabetes mellitus, existing cardiovascular disease, existing hypertension, clinical symptoms, previous stroke, or airway characteristics?			
Comparison of Portable Devices and Polysomnography			
PSG devices are classified as Type I monitors. Portable monitors are classified as either Type II, which record all the same information as PSG; Type III, which do not differentiate between whether the patient is asleep or awake, but have at least two respiratory channels (two airflow channels or one airflow and one effort channel); or Type IV, which fail to fulfill criteria for Type III monitors but usually record more than two bioparameters. *The strength of evidence is moderate*, among 15 quality A, 45 quality B, and 39 quality C studies, that Type III and Type IV monitors may have the ability to accurately predict AHI suggestive of OSA with high positive likelihood ratios and low negative likelihood ratios for various AHI cutoffs in	☐	New Evidence:	☐

Conclusions From CER Executive Summary	Is this conclusion almost certainly still supported by the evidence?	Has there been new evidence that may change this conclusion?	Do Not Know
PSG. Type III monitors perform better than Type IV monitors at AHI cutoffs of 5, 10, and 15 events/hr. Analysis of difference versus average analyses plots suggest that substantial differences in the measured AHI may be encountered between PSG and both Type III and Type IV monitors. Large differences compared with in-laboratory PSG cannot be excluded for all portable monitors. The evidence is insufficient to adequately compare specific monitors to each other.			
No recent studies compared Type II monitors with PSG. A prior Technology Assessment of home diagnosis of OSA concluded that "based on [three quality B studies], type II monitors [used at home] may identify AHI suggestive of OSA with high positive likelihood ratios and low negative likelihood ratios," though "substantial differences in the [measurement of] AHI may be encountered between type II monitors and facility-based PSG.".	☐	New Evidence:	☐
Comparison of Questionnaires and Polysomnography			
Of the six studies reviewed (one quality A, one quality B, four quality C), *the strength of evidence is low* among three studies supporting the use of the Berlin questionnaire in screening for sleep apnea because of the likely selection biases. *The strength of evidence is insufficient* to draw definitive conclusions concerning the use of	☐	New Evidence:	☐

Conclusions From CER Executive Summary	Is this conclusion almost certainly still supported by the evidence?	Has there been new evidence that may change this conclusion?	Do Not Know
the STOP, STOP-Bang, ASA Checklist, Epworth Sleepiness Scale, and Hawaii Sleep questionnaires to screen for sleep apnea because each questionnaire was assessed in only a single study.			
Clinical Prediction Rules and Polysomnography			
The strength of evidence is low among seven studies (three quality A, three quality B, and one quality C) that some clinical prediction rules may be useful in the prediction of a diagnosis of OSA. Ten different clinical prediction rules have been described. Nine clinical prediction rules have been used for the prediction of a diagnosis of OSA (using different criteria). The oropharyngeal morphometric model gave near perfect discrimination (area under the curve [AUC] = 0.996) to predict the diagnosis of OSA, and the pulmonary function data model had 100 percent sensitivity with 84 percent specificity to predict diagnosis of OSA. The remaining models reported lower diagnostic sensitivities and specificities. Each model was deemed useful to predict the diagnoses of OSA by the individual study authors. However, while all the models were internally validated, external validation of these predictive rules has not been conducted in the vast majority of the studies.	☐	New Evidence:	☐
Key Question 2. How does phased testing (screening tests or battery followed by full test) compare to full testing alone?			

Conclusions From CER Executive Summary	Is this conclusion almost certainly still supported by the evidence?	Has there been new evidence that may change this conclusion?	Do Not Know
The strength of evidence is insufficient to determine the utility of phased testing, followed by full testing when indicated, to diagnose sleep apnea, as only one study that met our inclusion criteria investigated this question. This prospective quality C study did not fully analyze the phased testing, thus the sensitivity and specificity of 6 the phased strategy could not be calculated due to a verification bias; not all participants received PSG (full) testing.	☐	New Evidence:	☐
Key Question 3. What is the effect of preoperative screening for sleep apnea on surgical outcomes?			
The strength of evidence is insufficient regarding postoperative outcomes with mandatory screening for sleep apnea. Two quality C prospective studies assessed the effect of preoperative screening for sleep apnea on surgical outcomes. One study found no significant differences in outcomes between patients undergoing bariatric surgery who had mandatory PSG or PSG based on clinical parameters. The second study found that general surgery patients willing to undergo preoperative PSG were more likely to have perioperative complications, particularly cardiopulmonary complications, possibly suggesting that patients willing to undergo PSG are more ill than other patients.	☐	New Evidence:	☐
Key Question 4. In adults being screened for obstructive sleep apnea, what are the relationships between apnea-hypopnea index or oxygen desaturation index, and other patient characteristics with respect to long-term clinical and functional outcomes?			

Conclusions From CER Executive Summary	Is this conclusion almost certainly still supported by the evidence?	Has there been new evidence that may change this conclusion?	Do Not Know
The strength of evidence is high from four studies (three quality A, one quality B) indicating that an AHI >30 events/hr is an independent predictor of all-cause mortality; although one study found that this was true only in men under age 70. All other outcomes were analyzed by only one or two studies. Thus, only a *lowstrength of evidence* exists that a high AHI (>30 events/hr) is associated with incident diabetes. This association, however, may be confounded by obesity, which may result in both OSA and diabetes. *The strength of evidence is insufficient* regarding the association between AHI and other clinical outcomes. The two studies of cardiovascular mortality did not have consistent findings, and the two studies of hypertension had unclear conclusions. One study of nonfatal cardiovascular disease found a significant association with baseline AHI (as they did for cardiovascular mortality). One study each found no association between AHI and stroke or long-term quality of life.	☐	New Evidence:	☐
Key Question 5. What is the comparative effect of different treatments for obstructive sleep apnea in adults? **a. Does the comparative effect of treatments vary based on presenting patient characteristics, severity of obstructive sleep apnea, or other pretreatment factors? Are any of these characteristics or factors predictive of treatment success?** **• Characteristics: age, sex, race, weight, bed partner, airway, other physical characteristics, and specific comorbidities** **• Obstructive sleep apnea severity or characteristics: baseline questionnaire (and similar tools) results, formal testing results (including hypoxemia levels), baseline quality of life, positional dependency** **• Other: specific symptoms** **b. Does the comparative effect of treatments vary based on the definitions of obstructive sleep apnea used by study investigators?**			

Conclusions From CER Executive Summary	Is this conclusion almost certainly still supported by the evidence?	Has there been new evidence that may change this conclusion?	Do Not Know
With some exceptions for studies of surgical interventions, we reviewed only randomized controlled trials (RCT) of interventions used specifically for the treatment of obstructive sleep apnea (OSA).			
Comparison of Continuous Positive Airway Pressure and Control			
There are 22 trials (11 each of quality B and C) that provide sufficient evidence supporting large improvements in sleep measures with continuous positive airway pressure (CPAP) compared with control. There is only weak evidence that demonstrated no consistent benefit in improving quality of life, neurocognitive measures, or other intermediate outcomes. Despite no evidence or weak evidence for an effect of CPAP on clinical outcomes, given the large magnitude of effect on the intermediate outcomes AHI and ESS, the strength of evidence that CPAP is an effective treatment to alleviate sleep apnea signs and symptoms was rated *moderate*.	☐	New Evidence:	☐
Comparison of CPAP and Sham CPAP			
There are 24 trials (5 quality A, 13 quality B, 6 quality C) that provide sufficient evidence supporting large improvements in sleep measures with CPAP compared 7 with sham CPAP, but weak evidence of possibly no difference between CPAP and sham CPAP in improving quality of life, neurocognitive measures, or other intermediate outcomes. Despite no evidence or weak evidence for an effect of CPAP on clinical outcomes, given the large magnitude of effect on the intermediate	☐	New Evidence:	☐

Conclusions From CER Executive Summary	Is this conclusion almost certainly still supported by the evidence?	Has there been new evidence that may change this conclusion?	Do Not Know
outcomes of AHI, ESS, and arousal index, the strength of evidence that CPAP is an effective treatment for the relief of signs and symptoms of sleep apnea was rated *moderate*.			
Comparison of Oral and Nasal CPAP			
Three small trials (one quality B, two quality C) with inconsistent results preclude any substantive conclusions concerning the efficacy of oral (or full face mask) versus nasal CPAP in improving compliance in patients with OSA. Largely due to small sample size, the reported effect estimates in the studies reviewed were generally imprecise. Thus, overall, *the strength of evidence is insufficient* regarding differences in compliance or other outcomes between oral and nasal CPAP.	☐	New Evidence:	☐
Comparison of Autotitrating CPAP and Fixed CPAP			
The strength of evidence is moderate that autotitrating CPAP (autoCPAP) and fixed pressure CPAP result in similar levels of compliance (hours used per night) and treatment effects for patients with OSA. Twenty-one studies (1 quality A, 10 quality B, 10 quality C) comprising an experimental population of over 800 patients provided evidence that autoCPAP reduces sleepiness as measured by ESS by approximately 0.5 points more than fixed CPAP. The two devices were found to result in similar compliance and changes in	☐	New Evidence:	☐

Conclusions From CER Executive Summary	Is this conclusion almost certainly still supported by the evidence?	Has there been new evidence that may change this conclusion?	Do Not Know
AHI from baseline, quality of life, and most other sleep study measures. However, there is also evidence that minimum oxygen saturation improves more with fixed CPAP than with autoCPAP, although by only about one percent. Evidence is limited regarding the relative effect of fixed CPAP and autoCPAP on blood pressure. There were no data on objective clinical outcomes.			
Comparison of Bilevel CPAP and Fixed CPAP			
The strength of evidence is insufficient regarding any difference in compliance or other outcomes between bilevel CPAP and fixed CPAP. Five small, highly clinically heterogeneous trials (one quality B, four quality C) with largely null findings did not support any substantive differences in the efficacy of bilevel CPAP versus fixed CPAP in the treatment of patients with OSA. Largely due to small sample sizes, the studies mostly had imprecise estimates of the comparative effects.	☐	New Evidence:	☐
Comparison of Flexible Bilevel CPAP and Fixed CPAP			
The strength of evidence is insufficient regarding the relative merits of flexible bilevel CPAP and fixed CPAP as there was only one quality B study that investigated this comparison. This study found that flexible bilevel CPAP may yield increased compliance (use ≥4 hr/night) compared with fixed CPAP.	☐	New Evidence:	☐

Conclusions From CER Executive Summary	Is this conclusion almost certainly still supported by the evidence?	Has there been new evidence that may change this conclusion?	Do Not Know
Comparison of C-Flex™ and Fixed CPAP			
No statistically significant differences in compliance or other outcomes were found between C-Flex and fixed CPAP. *The strength of evidence is low* for this finding because of the mixed quality (Bs and Cs) of the four primary studies.	☐	New Evidence:	☐
Comparison of Humidification in CPAP			
The strength of evidence is insufficient to determine whether there is a difference in compliance or other outcomes between positive airway pressure treatment with and without humidification. Five trials examined different aspects of humidified CPAP treatment for patients with OSA. While some studies reported a benefit of added humidity in CPAP treatment in improving patient compliance, this effect was not consistent across all the studies. Overall, the studies were clinically heterogeneous, small, and of quality B (three studies) or C (two studies).	☐	New Evidence:	☐
Comparison of Mandibular Advancement Devices and No Treatment or Inactive Oral Devices			
The strength of evidence is moderate to show that the use of mandibular advancement devices (MAD) improves sleep apnea signs and symptoms. Five trials (four quality B, one quality C) compared MAD with no treatment, using a variety of different types of MAD, and found significant improvements with MAD in	☐	New Evidence:	☐

Conclusions From CER Executive Summary	Is this conclusion almost certainly still supported by the evidence?	Has there been new evidence that may change this conclusion?	Do Not Know
AHI, ESS, and other sleep study measures. Any differences in quality of life measures or neurocognitive tests were equivocal between treatment groups. No trial evaluated objective clinical outcomes. Another five trials (four quality B, one quality C) compared the effects of MAD with inactive oral devices and reported similar findings.			
Comparison of Different Oral Devices			
The strength of evidence is insufficient to draw conclusions with regard to the relative efficacy of different types of oral MAD in patients with OSA because the reviewed studies were generally small, and each was concerned with a unique comparison. Five studies (four quality B, one quality C) with unique comparisons found little to no differences between different types and methods of use of MAD or other oral devices in sleep study or sleepiness measures. No study evaluated objective clinical outcomes. Only one study evaluated compliance; no significant differences were observed. One trial found that a greater degree of mandibular advancement resulted in an increased number of patients achieving an AHI <10 events/hr; however, the mean AHI was similar between treatment groups.	☐	New Evidence:	☐
Comparison of Mandibular Advancement Devices and CPAP			

Conclusions From CER Executive Summary	Is this conclusion almost certainly still supported by the evidence?	Has there been new evidence that may change this conclusion?	Do Not Know
The strength of evidence is moderate that CPAP is superior to MAD in improving sleep study measures. Ten mostly quality B trials overall found that CPAP resulted in greater reductions in AHI and arousal index, and increases in minimum oxygen saturation. The evidence regarding the relative effects on ESS were too heterogeneous to allow conclusions. In a single study, patients were more compliant with MAD than CPAP (hours used per night and nights used). No study evaluated objective clinical outcomes. *The strength of evidence is insufficient* to address which patients might benefit most from either treatment.	☐	New Evidence:	☐
Comparison of Surgery and Control			
The strength of evidence is insufficient to evaluate the relative efficacy of surgical interventions for the treatment of OSA. Six trials and one nonrandomized prospective study with unique interventions compared surgery with control treatment for the management of patients with OSA. Three studies were rated quality A, one quality B, and three quality C. The results were inconsistent across studies as to which outcomes were improved with surgery compared with no or sham surgery.	☐	New Evidence:	☐
Comparison of Surgery and CPAP			
The strength of evidence is insufficient to determine the relative merits of surgical	☐	New Evidence:	☐

Conclusions From CER Executive Summary	Is this conclusion almost certainly still supported by the evidence?	Has there been new evidence that may change this conclusion?	Do Not Know
treatments versus CPAP. Of 12 studies (1 quality A, 11 quality C) comparing surgical modalities with CPAP, only two were RCTs, and they compared CPAP with uvulopalatopharyngoplasty (UPPP), removal of the soft tissue at the back of the throat, the uvula, and soft palate. While one of these trials found that CPAP resulted in a higher mortality benefit, the other found no difference between groups. Due to the heterogeneity of interventions and outcomes examined, the variability of findings across studies, and the inherent bias of all but one study regarding which patients received surgery, it is not possible at this time to draw useful conclusions comparing surgical interventions with CPAP in the treatment of patients with OSA. The quality A trial was the only unbiased comparison of surgery and CPAP (patients had previously received neither treatment). It did not find statistically significant differences in ESS and quality of life measures between patients with mild to moderate OSA who had temperature-controlled radiofrequency tissue volume reduction of the soft palate and those who had CPAP at 2 months followup. Likewise, the other trial, comparing maxillomandibular advancement osteotomy and CPAP, did not find statistically significant differences in AHI and ESS in patients with severe OSA. For the nonrandomized studies, comparisons			

Conclusions From CER Executive Summary	Is this conclusion almost certainly still supported by the evidence?	Has there been new evidence that may change this conclusion?	Do Not Know
between surgery and CPAP are difficult to interpret since baseline patient characteristics (including sleep apnea severity) differed significantly between groups, particularly in regards to what previous treatments patients had. The reported findings on sleep study and quality of life measures were heterogeneous across studies.			
Comparison of Surgery and Mandibular Advancement Devices			
The strength of evidence is insufficient regarding the relative merit of MAD versus surgery in the treatment of OSA, as there was only one study (quality B) that examined this question. A statistically significant improvement in AHI was observed in the MAD group compared with the surgery group. No study evaluated objective clinical outcomes.	☐	New Evidence:	☐
Comparison of Other Treatments			
The strength of evidence is low to show that some intensive weight loss programs may be effective treatment for OSA in obese patients. Three trials (one quality A, two quality B) compared weight loss interventions with control interventions. All three trials found significant relative reductions in AHI with diet. Other outcomes were inconsistent. *The strength of evidence is insufficient* to determine the effects of other potential treatments for OSA. Twentyone studies	☐	New Evidence:	☐

Conclusions From CER Executive Summary	Is this conclusion almost certainly still supported by the evidence?	Has there been new evidence that may change this conclusion?	Do Not Know
evaluated other interventions including atrial overdrive pacing, eight different drugs, palatal implants, oropharyngeal exercises, a tongue-retaining device, a positional alarm, combination tongueretaining device and positional alarm, bariatric surgery, nasal dilator strips, acupuncture, and auricular plaster. All of these interventions were evaluated by one or two studies only. The findings were heterogeneous. No study evaluated objective clinical outcomes.			
Key Question 6. In OSA patients prescribed nonsurgical treatments, what are the associations of pretreatment patient-level characteristics with treatment compliance?			
Across five studies (one quality A, one quality B, three quality C), *the strength of evidence is moderate* that more severe OSA as measured by higher AHI is associated with greater compliance with CPAP use. Each study measured compliance differently, including thresholds of 1, 2, or 3 hours of use per night or as a continuous variable, and undefined "objective compliance" measured by the device. *The strength of evidence is moderate* that a higher ESS score is also associated with improved compliance. *There are low strengths of evidence* that younger age, snoring, lower CPAP pressure, higher BMI, higher mean oxygen saturation, and the sleepiness domain on the Grenoble Sleep Apnea Quality of Life test are each possible independent predictors of compliance. It is	☐	New Evidence:	☐

Conclusions From CER Executive Summary	Is this conclusion almost certainly still supported by the evidence?	Has there been new evidence that may change this conclusion?	Do Not Know
important to note, however, that selective reporting, particularly of nonreporting of nonsignificant associations, cannot be ruled out. The heterogeneity of analyzed and reported potential predictors greatly limits these conclusions. Differences across studies as to which variables were independent predictors may be due to the adjustment for different variables, in addition to differences in populations, outcomes, CPAP machines, and CPAP training and followup. One quality C study of mandibular advancement devices failed to identify potential predictors of compliance.			
Key Question 7. What is the effect of interventions to improve compliance with device (positive airway pressure, oral appliances, positional therapy) use on clinical and intermediate outcomes?			
The strength of evidence is low that some specific adjunct interventions may improve CPAP compliance, but studies are heterogeneous and no general type of intervention (e.g., education, telemonitoring) was more promising than others. The 18 trials (two quality A, eight quality B, and eight quality C) had inconsistent effects across a wide variety of interventions. Studies generally had small sample sizes with less than 1 year of followup. Compared with usual care, several interventions were shown to significantly increase hours of CPAP use per night in some studies. These included intensive support or literature (designed for	☐	New Evidence:	☐

Conclusions From CER Executive Summary	Is this conclusion almost certainly still supported by the evidence?	Has there been new evidence that may change this conclusion?	Do Not Know
patient education), cognitive behavioral therapy (given to patients and their partners), telemonitoring, and a habit-promoting audio-based intervention. However, the majority of studies did not find a significant difference in CPAP compliance between patients who received interventions to promote compliance with device use and those who received usual care. No study of nurseled care (which was not focused primarily on compliance) showed an effect on compliance rates.			
Are there new data that could inform the key questions that might not be addressed in the conclusions?			