QUERI

The Comparative Effectiveness, Harms, and Cost of Care Models for the Evaluation and Treatment of Obstructive Sleep Apnea (OSA): A Systematic Review

September 2016

Prepared for:
Department of Veterans Affairs
Veterans Health Administration
Quality Enhancement Research Initiative
Health Services Research & Development Service
Washington, DC 20420

Prepared by:
Evidence-based Synthesis Program (ESP)
Minneapolis VA Health Care System
Minneapolis, MN
Timothy J. Wilt, MD, MPH, Director
Nancy Greer, PhD, Program Manager

Investigators:
Principal Investigators:
Ken M. Kunisaki, MD, MS
Wajahat Khalil, MD

Co-Investigators:
Erin Koffel, PhD
Loretta Pfannes, MD

Research Associates:
Eva Koeller, BA
Roderick MacDonald, MS

VA HEALTH CARE Defining EXCELLENCE in the 21st Century

PREFACE

The VA Evidence-based Synthesis Program (ESP) was established in 2007 to provide timely and accurate syntheses of targeted healthcare topics of particular importance to clinicians, managers, and policymakers as they work to improve the health and healthcare of Veterans. QUERI provides funding for four ESP Centers, and each Center has an active University affiliation. Center Directors are recognized leaders in the field of evidence synthesis with close ties to the AHRQ Evidence-based Practice Centers. The ESP is governed by a Steering Committee comprised of participants from VHA Policy, Program, and Operations Offices, VISN leadership, field-based investigators, and others as designated appropriate by QUERI/HSR&D.

The ESP Centers generate evidence syntheses on important clinical practice topics. These reports help:

- Develop clinical policies informed by evidence;
- Implement effective services to improve patient outcomes and to support VA clinical practice guidelines and performance measures; and
- Set the direction for future research to address gaps in clinical knowledge.

The ESP disseminates these reports throughout VA and in the published literature; some evidence syntheses have informed the clinical guidelines of large professional organizations.

The ESP Coordinating Center (ESP CC), located in Portland, Oregon, was created in 2009 to expand the capacity of QUERI/HSR&D and is charged with oversight of national ESP program operations, program development and evaluation, and dissemination efforts. The ESP CC establishes standard operating procedures for the production of evidence synthesis reports; facilitates a national topic nomination, prioritization, and selection process; manages the research portfolio of each Center; facilitates editorial review processes; ensures methodological consistency and quality of products; produces “rapid response evidence briefs” at the request of VHA senior leadership; collaborates with HSR&D Center for Information Dissemination and Education Resources (CIDER) to develop a national dissemination strategy for all ESP products; and interfaces with stakeholders to effectively engage the program.

Comments on this evidence report are welcome and can be sent to Nicole Floyd, ESP CC Program Manager, at Nicole.Floyd@va.gov.

Recommended citation: Kunisaki KM, Khalil W, Koffel E, Pfannes L, Koeller E, MacDonald R, Greer, N, Wilt TJ. The Comparative Effectiveness, Harms, and Cost of Care Models for the Evaluation and Treatment of Obstructive Sleep Apnea (OSA): A Systematic Review. VA ESP Project #09-009; 2016.

This report is based on research conducted by the Evidence-based Synthesis Program (ESP) Center located at the **Minneapolis VA Health Care System, Minneapolis, MN**, funded by the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, Quality Enhancement Research Initiative. The findings and conclusions in this document are those of the author(s) who are responsible for its contents; the findings and conclusions do not necessarily represent the views of the Department of Veterans Affairs or the United States government. Therefore, no statement in this article should be construed as an official position of the Department of Veterans Affairs. No investigators have any affiliations or financial involvement (*eg*, employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties) that conflict with material presented in the report.

TABLE OF CONTENTS

RESULTS

SUMMARY AND DISCUSSION

FIGURES

TABLES

ABSTRACT

BACKGROUND

Rising requests for obstructive sleep apnea (OSA) diagnostic and treatment services may make traditional in-person evaluation processes unnecessarily expensive and inefficient.

PURPOSE

To assess the comparative effectiveness, harms, and costs of care models for OSA evaluation and treatment.

DATA SOURCES

MEDLINE (Ovid) and CINAHL searched for studies published in English language between January 2000 and May 2016 with hand searching of reference lists of related systematic reviews and included studies.

STUDY SELECTION

Four randomized controlled trials (RCTs) and 4 observational studies evaluated case-finding and care provided by nonsleep specialist practitioners versus sleep specialist physicians (SSP). No reports evaluated electronic versus interactive (*eg,* in-person or telephone) consultation. Twenty-seven reports (3 RCTs for *titration,* 22 RCTs for *treatment* and 2 cohort studies) assessed in-home autotitrating continuous positive airway pressure (APAP) technology versus standard continuous positive airway pressure (CPAP) titrated by in-lab polysomnogram (PSG) as options for PAP titration or treatment.

DATA EXTRACTION

Two investigators independently extracted study data, rated individual study risk of bias, and assessed overall strength of evidence.

DATA SYNTHESIS/RESULTS

No studies assessed the diagnostic accuracy of non-sleep-specialist nurses for case finding and referral. One retrospective study reported good agreement between a primary care pulmonologist and a SSP on what sleep test to order for patients referred by their family physician. Patient-centered outcomes were infrequently and inconsistently reported. When reported there was no significant difference in clinical outcomes between OSA treated by primary care/nurses and SSPs (moderate strength of evidence for quality of life). Sleep symptom scores were similar between groups (moderate strength of evidence). Treatment compliance was similar between patients treated by SSPs and those not (moderate strength of evidence). Strength of evidence for access to care and adverse events was insufficient.

Few studies compared patient-centered outcomes between in-lab CPAP *titration* and at-home APAP *titration*. In limited reporting, study groups were generally similar on quality of life (moderate strength of evidence) and cognitive symptoms. Some differences were noted for resource utilization and patient preference. Sleep measures, blood pressure, adverse events, and

compliance/adherence were generally similar (moderate strength of evidence for Epworth Sleepiness Scale (ESS) scores and low strength of evidence for compliance).

Twenty-three studies compared treatment with CPAP versus APAP. Few studies reported patient-centered outcomes other than quality of life and patient preference for one treatment approach over another. Quality of life, assessed with the SF-36, was generally similar (moderate strength of evidence). Patient preference was generally similar or favored APAP. Strength of evidence was insufficient for access to care. Post-treatment ESS scores were generally similar for the 2 treatment approaches (moderate strength of evidence). Adverse events were mild and similar for APAP and CPAP (low strength of evidence). Compliance was similar (moderate strength of evidence).

LIMITATIONS

Studies were limited to English language and those published in the United States, Canada, Western Europe, Australia, or New Zealand. Few studies assessed non-sleep specialist case-finding or care and none assessed electronic consultations. Clinical outcomes were infrequently reported.

CONCLUSIONS

Among patients suspected of having OSA, primary care providers and sleep-specialist nurses provide similar outcomes to SSPs, although many outcomes were inconsistently reported. At-home APAP provides similar outcomes to CPAP titrated in the PSG laboratory. No evidence addressed electronic consultation for the management of known or suspected OSA. Future studies are needed to determine which patients derive the most benefit from treatment and should be prioritized for testing and treatment, whether newer models of care with less reliance on SSP time (either through utilization of other types of providers or electronic consultation) result in similar outcomes to traditional models, and, if effective, how such models should be implemented.

EXECUTIVE SUMMARY

INTRODUCTION

Obstructive sleep apnea (OSA) is a chronic condition that results from repeated closure of the upper airway during sleep resulting in reduced airflow (hypopnea) or complete airflow cessation (apnea) leading to cyclic sleep disruption. Patients with OSA frequently experience excessive daytime sleepiness and decreased quality of life. However, not all individuals have excessive daytime sleepiness and symptoms are not required to make a diagnosis or obtain treatment. OSA has also been associated with a higher risk of myocardial infarction, heart failure, stroke, and cognitive decline. Continuous positive airway pressure (CPAP) effectively reduces apneas and hypopneas in most patients with OSA, improves blood pressure, and – particularly in those with symptoms of excessive daytime sleepiness – improves quality of life and sleep symptoms. CPAP use is also associated with a reduced risk of motor vehicle accidents.

The estimated prevalence of mild to severe OSA in the United States (2007-2010 data) among 30- to 70-year-olds is 34% for men and 17% for women. Despite the data associating OSA with consequences to health and quality of life and the conclusions from other guideline groups, many persons with OSA remain undiagnosed. In 2010, among 1.8 million US Veterans receiving outpatient care at 136 Veterans Affairs (VA) facilities, 37.4% had a body mass index (BMI) ≥ 30 kg/m^2. Given the strong relationship between high BMI and OSA, these data suggest that a substantial portion of US Veterans are at high risk for OSA. A recent analysis of Veterans Administration Informatics and Computing Infrastructure (VINCI) data between 2000 and 2010 showed that among 9.8 million Veterans, the age-adjusted prevalence of diagnosed sleep apnea was 0.4% in 2000 and had increased to 3.0% in 2010 (a relative increase of 650%). As awareness of OSA by patients and providers continues to increase, and as BMI continues to increase in the US and globally, healthcare systems such as the VA need to develop strategies to manage the increasing demand for sleep services. One strategy is to target screening and testing to those most likely to derive benefit from OSA treatment (*eg*, those with significant unexplained daytime sleepiness), as suggested by the American College of Physicians (ACP). Another strategy is to improve efficiency within healthcare systems by implementing innovative, less resource-intense models of care for OSA.

The traditional model of OSA evaluation and care relies upon primary care providers to refer patients with suspected OSA to a sleep specialist physician (SSP) for consultation. After an initial consultative visit, the SSP can order formal, in-lab polysomnogram (PSG) for diagnostic purposes and for those with confirmed OSA, a PSG for titration of CPAP pressures. The SSP would typically initiate CPAP at the pressure suggested by the titration PSG, and then the patient would follow up with the SSP at regular intervals for assessment of treatment compliance and efficacy. Given the rapidly rising requests for OSA diagnostic and treatment services, this traditional model is increasingly viewed as unnecessarily expensive and inefficient for the evaluation and treatment of patients at high risk of OSA. Recent data also indicate a decreasing supply of SSPs to care for patients with known or suspected OSA.

Therefore, new models of OSA care have been proposed and implemented. These new models include home sleep testing (HST) for diagnostic purposes, followed by treatment with an autotitrating CPAP (APAP) device which has internal algorithms to adjust CPAP pressure to keep the airway open during sleep. These models reduce PSG-associated costs and logistical

barriers, yet typically still include consultation and follow up with a SSP. Other proposed models would reduce reliance on SSPs by including non-SSP providers such as nurses or primary care physicians to provide the bulk of OSA diagnosis and treatment.

Although several studies have been conducted to test some of these new models, systematic reviews are lacking. The Minneapolis VA's Evidence-based Synthesis Program (ESP) Center, in partnership with topic nominators and a Technical Expert Panel (TEP), was commissioned to systematically review the evidence regarding the comparative effectiveness, harms, and cost of these new models of OSA evaluation and treatment.

We addressed the following key questions (KQs):

Key Question 1. For adults with suspected OSA, what are the effectiveness/harms/resource utilization of case finding and care provided by practitioners who are not sleep physicians (including PCPs, PAs, NPs, technologists, nurses, and respiratory therapists), compared to case finding and care provided by sleep specialist physicians?

Key Question 2. For adults with suspected OSA, what are the effectiveness/harms/resource utilization of electronic consultation versus interactive (*eg*, in-person, telephone) consultation?

Key Question 3. For adults diagnosed with OSA, what are the effectiveness/harms/resource utilization (including cost avoidance) of using in-home autotitrating continuous positive airway pressure (APAP) technology compared to standard continuous positive airway pressure (CPAP) titrated by in-lab PSG?

METHODS

Data Sources and Searches

We searched MEDLINE (Ovid) and CINAHL for articles published between 2000 and May 2016. We obtained additional articles by hand searching the reference lists of related systematic reviews and included studies.

Study Selection

Abstracts were independently reviewed in duplicate by investigators and research associates. We included studies of any design, published in English, that reported on OSA care or case finding in adults with suspected or diagnosed OSA and took place in North America, Europe, Australia, or New Zealand. For KQ1 we excluded studies that did not include a comparison of a supervised practitioner or non-specialist licensed independent practitioner (*eg*, primary care physician, physician's assistant, nurse, technologist, or respiratory therapist) to a SSP. We excluded studies evaluating the role of dentists and anesthesiologists. For KQ2 we excluded studies that did not compare an electronic initial consultation without patient contact to an interactive initial consultation. For KQ3 we excluded studies that did not compare the use of APAP to CPAP for *titration* or *treatment* of OSA. Furthermore, in studies of *titration* we only included articles in which the APAP was used at home and CPAP was manually titrated in a lab. We also excluded studies if they did not report any of our outcomes of interest.

Full-text reports of studies identified as potentially eligible based on abstract review were obtained for further review. Each article was independently reviewed by 2 investigators or research associates with disagreements settled by a third.

Data Abstraction and Risk of Bias Assessment

Study characteristics (location, setting, intervention groups, follow-up, aim of study, treatments, inclusion/exclusion criteria, and patient characteristics) as well as intermediate and clinical outcomes were extracted onto evidence tables by one investigator or research associate and verified by another. Trained research methodologists rated the risk of bias of individual studies as low, moderate, or high risk.

We assessed strength of evidence for the following outcomes: access to care, quality of life, compliance (hours of use per night), and adverse events. Strength of evidence was rated as high, moderate, low, or insufficient based on precision, consistency, directness, and risk of bias of the individual studies.

Data Synthesis and Analysis

We summarized findings by Key Question. We were able to pool data for quality of life (SF-36 scores), Epworth Sleepiness Scale (ESS) scores, and compliance (hours per night of use) for KQ1 and KQ3 (*titration* and *treatment* studies).

RESULTS

Results of Literature Search

We reviewed 2,847 abstracts, 2,252 from MEDLINE, and 595 from CINAHL. We excluded 2,493 abstracts and reviewed the full-text of 354 references. During full-text review we excluded 323 articles, leaving 31 eligible for inclusion. Hand searching reference lists of pertinent trials and systematic review identified an additional 3 references.

Summary of Results for Key Questions

Key Question 1

Eight studies (n = 1,401; 4 randomized controlled trials [RCTs]) reported results for KQ1. Sleep physician care was compared to management by primary care in 4 studies (n = 564), sleep-specialist nurses in 3 studies (n = 434), and other non-sleep physicians in one study (n = 403). Patients were generally middle-aged and moderately obese with mild sleepiness and severe sleep apnea as determined by apnea-hypopnea index (AHI).

Case Finding

One retrospective study reported on the ability of a primary care pulmonologist to order the proper sleep test. There was good agreement between the primary care pulmonologist and SSP (kappa = .74) and 93% (89/96) of the referred patients were diagnosed with OSA.

Care

Seven studies reported treatment outcomes in patients being managed by providers other tha SSPs. Three of these studies compared SSP care to primary care (n = 468, 1 RCT), 3 to sleep specialist nurses (n = 434; 3 RCTs), and one compared SSP care to management by a variety of physicians who were not sleep specialists.

Three studies reported quality of life; all found that SF-36 scores were similar between patients being treated by primary care (k = 1) or sleep-specialist nurses (k = 2) as compared to SSPs. Two studies reporting patient satisfaction found overall satisfaction was similar between groups.

CPAP compliance was reported in 7 studies. Six of the 7 studies found no difference in compliance, regardless of measure used, when comparing patients receiving SSP care to those receiving care from non-SSPs. The final study found that patients who were referred for a sleep study by non-SSPs were significantly less compliant, with fewer hours per night and less regular use, than those patients who were referred by SSPs. Cost was reported, in various ways, by 5 of the studies. Three did not report the significance of cost differences between SSP and non-SSP care. Two studies, however, found that nurse-led OSA care was associated with significantly lower costs per patient and within-trial costs.

Five studies reported ESS scores; all found scores were similar between groups receiving care from different providers. AHI was reported by one study. Residual AHI on CPAP was significantly lower in patients referred for PSG by non-SSPs than in patients referred by SSPs (P<.001).

Occurrence of adverse events was similar in the SSP and nurse/primary care groups. Time to initiation of therapy was reported in 2 studies. One found that significantly fewer patients in the primary care group received CPAP within one month of PSG when compared to patients in the SSP group (P = .012). The other reported that while there was no significant difference between groups for satisfaction with time waiting (P = .706), patients receiving nurse-led care were more satisfied with their impression of wait time (P = .004).

Key Question 2

No articles were identified that met inclusion criteria for this question.

Key Question 3

Twenty-seven studies addressed KQ3 and KQ3a, including 4 that compared *titration* with APAP to *titration* with CPAP and 23 that compared *treatment* with APAP to *treatment* with CPAP. The mean age of patients enrolled was 52 years, 80% were male, baseline BMI was 33 kg/m^2, baseline ESS was 13, and baseline AHI was 44 events/hour.

Titration

Of the 4 studies comparing *titration* in-lab with CPAP to *titration* at home with APAP, 3 were RCTs and one was an observational study. Mean ESS scores at baseline ranged from 14.0 to 15.5.

Few studies reported clinical outcomes. Most frequently reported was quality of life with 2 studies finding CPAP *titration* and APAP *titration* to be similar while another found mixed results for different subscales of the SF-36. No study reported access to care.

Intermediate outcomes were more commonly reported. ESS scores and compliance were generally similar for the CPAP *titration* and APAP *titration* groups.

Treatment

Among the 23 studies comparing treatment with CPAP to treatment with APAP, there were 22 RCTs (15 using a crossover design) and one retrospective cohort study. Baseline ESS scores ranged from 6.4 to 17.4.

The most commonly reported clinical outcomes were quality of life (specifically SF-36 scores) and patient preference for one treatment over another. Quality of life was similar for the 2 treatment groups in 7 of 9 studies reporting that outcome. In 7 of 12 studies reporting treatment preference, APAP was preferred. No study reported a significantly higher preference for CPAP.

Intermediate outcomes, including ESS scores and compliance (hours per night and proportion of nights used) were generally similar for the CPAP and APAP treatment group. Adverse events, reported in 6 studies, were generally similar between groups.

DISCUSSION

Key Findings and Strength of Evidence (Executive Summary Table)

Key Question 1

Case Finding

- No studies assessed the diagnostic accuracy of non-sleep-specialist nurse for case finding and referral.

- One retrospective study reported good agreement between a primary care pulmonologist and a SSP on what sleep test to order for patients referred by their family physician.

Care

- Clinical (*ie*, patient-centered) outcomes were infrequently and inconsistently reported. When reported there was no significant difference in clinical outcomes between OSA treated by primary care/nurses and SSPs. The strength of evidence for quality of life was moderate.

- Intermediate outcomes were more commonly reported. Sleep symptom scores were similar between groups (moderate strength of evidence).

- There was little evidence that treatment compliance differed between patients treated by SSPs and those not, including the proportion of patients with 4 hours or more of CPAP use on 70% or more of nights (moderate strength of evidence).

- Very few studies reported other intermediate outcomes. One reported a significantly lower residual AHI on CPAP in patients referred for PSG by non-sleep specialists and another found that the proportion of patients receiving CPAP within one month of their PSG was significantly higher in patients cared for by a SSP. Strength of evidence for access to care and adverse events was insufficient.

Key Question 3

Titration

- Few studies compared clinical (*ie*, patient-centered) outcomes between in-lab CPAP *titration* and at-home APAP *titration*. In limited reporting, study groups were generally similar on measures of quality of life (moderate strength of evidence) and cognitive symptoms. Some differences were noted for resource utilization and patient preference.

- Intermediate outcomes (*ie*, sleep measures, blood pressure, adverse events, and compliance/adherence) were more commonly reported and generally similar. Strength of evidence for ESS was moderate and strength of evidence for compliance was low. Strength of evidence for adverse events was insufficient.

Treatment

- Twenty-three studies compared *treatment* with CPAP to *treatment* with APAP. The studies enrolled patients with a broad range of baseline AHI values.

- Few studies reported clinical (*ie*, patient-centered) outcomes other than quality of life and patient preference for one *treatment* approach over another. Quality of life, assessed with the SF-36, was generally similar between the CPAP and APAP groups (moderate strength of evidence). Patient preference was also generally similar or favored APAP in studies reporting statistical significance. Strength of evidence was insufficient for access to care.

- Intermediate outcomes including post-treatment ESS scores were frequently reported and generally similar for the 2 *treatment* approaches (moderate strength of evidence). Adverse events were mild and similar for APAP and CPAP (low strength of evidence).

- Compliance, reported as either hours per night or the proportion of nights the device was used, was also similar for the CPAP and APAP *treatment* groups (moderate strength of evidence).

Executive Summary Table. Strength of Evidence

Comparison	Outcome of interest	Strength of evidence[a]	Direction
KQ1: Sleep physician care compared to management by primary care, sleep-specialist nurses, or other non-sleep physicians	Access to care	Insufficient	
	Epworth Sleepiness Score	Moderate	**SIMILAR**
	Quality of life	Moderate	**SIMILAR**
	Compliance, hours per night	Moderate	**SIMILAR**
	Adverse events	Insufficient	
KQ3: Home APAP technology versus standard in-center manual CPAP *titration*	Access to care	Insufficient	
	Epworth Sleepiness Score	Moderate	**SIMILAR**
	Quality of life	Moderate	**SIMILAR**
	Compliance	Low	**SIMILAR**
	Adverse events	Insufficient	
KQ3: APAP versus CPAP *treatment*	Access to care	Insufficient	
	Epworth Sleepiness Score	Moderate	**SIMILAR**
	Quality of life	Moderate	**SIMILAR**
	Compliance	Moderate	**SIMILAR**
	Adverse events	Low	

APAP = Auto-adjusted (autoregulated) continuous positive airway pressure; CPAP = continuous positive airway pressure

[a]Strength of Evidence Definitions (Owens, DK et al, J Clin Epidemiol. 2010;63(5):523-523)

• High: Very confident that estimate of effect lies close to true effect. Few or no deficiencies in body of evidence, findings believed to be stable.

• Moderate: Moderately confident that estimate of effect lies close to true effect. Some deficiencies in body of evidence; findings likely to be stable, but some doubt.

• Low: Limited confidence that estimate of effect lies close to true effect; major or numerous deficiencies in body of evidence. Additional evidence necessary before concluding that findings are stable or that estimate of effect is close to true effect.

• Insufficient: No evidence, unable to estimate an effect, or no confidence in estimate of effect. No evidence is available or the body of evidence precludes judgment.

Limitations

We limited inclusion to English language studies and those performed in the United States, Canada, Western Europe, Australia, or New Zealand. This likely increases applicability of our findings to the VA healthcare system, but may reduce generalizability of our findings to other health settings. We also chose to exclude studies of dentists and anesthesiologists due to our

uncertainty whether or not these persons truly represent "Non-Sleep-Specialist Physicians" given that many dentists and anesthesiologists have substantial practices or background in sleep medicine. We also did not assess the role of surgery or mandibular assist devices for OSA treatment including referral of patients who may be more interested in or better candidates for these options.

A significant limitation is the paucity of high-quality literature regarding key questions 1 and 2. Although we are aware of the existence of clinical care models utilizing sleep respiratory therapists and behavioral sleep medicine providers to provide varying degrees of OSA care, we did not find any studies to address these particular type of practitioners. We also note that, importantly, the providers in many of these "primary care" studies were persons who had substantial experience in sleep medicine. Therefore, the generalizability of these findings to primary care providers with less experience in sleep medicine is not clear.

Applicability and Implementation

Most patients enrolled in studies were obese, middle-aged men, with severe OSA based on both high AHI levels and the presence of excessive daytime sleepiness. Our findings are most applicable to these individuals. We found only one study that was performed in a VA population for Key Question 3. One study was performed at Walter Reed Army Medical Center for key questions 1 and 3. While most studies were not specifically conducted in VA or military populations, because the patients enrolled in these studies were generally older, overweight men with OSA, we believe the findings of our systematic review should be applicable to the population of Veterans served by VA facilities.

Many of the study providers who were not SSPs had prior sleep training and therefore the results may not be fully generalizable to all primary care providers. While data are not conclusive, because our findings indicated similar Epworth Sleepiness Scores, quality of life, and treatment compliance scores among patients evaluated and treated by non-SSPs compared to SSPs it may be reasonable to consider expanded use of non-SSP providers who have received training in sleep medicine, especially where SSPs are in limited supply and demand for OSA services is high. Similarly, greater use of at-home APAP may lessen dependence on backlogged PSG laboratories, as most health outcomes were similar between groups.

Our report focused on methods that might improve the 'supply' side of OSA evaluation and treatment, through use of non-SSPs, electronic consultation, and at-home APAP *titration* and *treatment*. However, healthcare systems struggling to match supply to demand might also consider whether the 'demand' is truly appropriate. We found little to no data in screen-detected patients (*ie*, those found to have abnormal AHI either through direct referral to sleep laboratories or based on results of screening questionnaires such as the Berlin questionnaire but without excessive daytime sleepiness). The evidence to date indicates that the main benefit of OSA detection and treatment is improvement in patient-reported sleepiness symptoms among those with unexplained daytime somnolence. Therefore, VA healthcare providers and decision-makers could potentially achieve the highest value care, including resource use, by targeting case finding approaches and subsequent evaluation and treatment to individuals with unexplained daytime somnolence and who express interest in further evaluation and treatment. Theorectically, this referral approach could be readily be implemented by developing and using electronic medical record templates that describe the evidence-based rationale for the referral recommendations

while requesting that referring providers include information specifically about daytime somnolence and the patient's willingness for further evaluation and treatment in the consult.

Research Gaps/Future Research

Comparative effectiveness trials were lacking for all key questions. Key questions 1 and 2 would particularly benefit from trials to address the outcomes resulting from non-SSP care of sleep apnea patients (Key Question 1) and electronic consultation (Key Question 2). Limited available data suggest that care led by non-SSPs may potentially provide equivalent outcomes to care led by SSPs. Comparative effectiveness trials are needed in order to determine whether such results can be achieved in routine practice, outside of controlled research settings. The available data suggest that with some appropriate training, non-SSPs can potentially provide equivalent outcomes, but the operationalization of such training is unclear. Therefore, future comparative effectiveness trials should describe their training programs. Such trials should also collect clinical outcomes where possible.

As more healthcare systems implement comprehensive electronic medical records (EMRs), we anticipate Key Question 2 will become more feasible to study. In the current climate of increasing numbers of sleep referrals, yet a declining number of SSPs, EMR-based electronic consultation holds significant promise to provide equivalent outcomes in a more cost-effective, time-efficient manner. Although many systems have already implemented electronic consultation systems, evidence supporting this practice is largely lacking and further studies should be conducted. We think comparative effectiveness study designs such as stepped-wedge randomization (where sites are randomly assigned to the time point at which they implement electronic consultation) would allow the creation of good-quality evidence to quantify the risks, benefits, and economic impacts of electronic consultation for patients with known or suspected OSA.

We also note that blinding the intervention is not feasible in many of these randomized trials. However, we recommend that future studies make efforts to assess outcomes by persons blinded to treatment assignment and thereby mitigate the potential for biased assessments.

A large gap in evidence is related to the effectiveness of treatment for individuals without excessive daytime sleepiness (screen detected or case-finding in at-risk asymptomatic individuals) and on outcomes other than daytime sleepiness. For example, a recently published multi-site randomized trial showed that CPAP was not effective for secondary prevention of cardiovascular events in those with established cardiovascular disease and moderate to severe OSA. Additional information gaps include the effectiveness of treatment in those with mild AHI regardless of symptoms. These gaps are supported by recent evidence reports and accompanying clinical practice recommendations by the ACP and US Preventive Services Task Force (USPSTF) and described in greater detail in the introduction to the main report. These are critically important gaps to fill due to the dramatic increase in patients with, or suspected to have, OSA and thus being referred for evaluation and treatment.

Conclusions

Among patients suspected of having OSA, evidence suggests that primary care providers and sleep-specialist nurses might provide similar outcomes to SSPs, although the strength of this evidence was only moderate and many outcomes were inconsistently reported. Likewise, among

patients diagnosed with OSA, evidence suggests that at-home APAP *titration* and *treatment* provides similar outcomes to fixed pressure CPAP titrated in the PSG laboratory, although the strength of evidence was generally low to moderate.

We found no evidence addressing the topic of electronic consultation for the management of known or suspected OSA.

Future studies are needed to determine which patients derive the most benefit from treatment and should be prioritized for testing and treatment, whether newer models of care with less reliance on SSP time (either through utilization of other types of providers or electronic consultation) result in similar outcomes to traditional models, and if effective, how such models should be implemented.

ABBREVIATIONS TABLE

AHI	Apnea-Hypopnea Index
APAP	Auto-adjusting Positive Airway Pressure
BMI	Body Mass Index
CI	Confidence Interval
CPAP	Continuous Positive Airway Pressure
ESS	Epworth Sleepiness Scale
FOSQ	Functional Outcomes of Sleep Questionnaire
HbA1c	Hemoglobin A1c
HST	Home Sleep Testing
LIP	Licensed Independent Practitioner
MeSH	Medical Subject Heading
MID	Minimally Important Difference
ODI	Oxygen Desaturation Index
OSA	Obstructive Sleep Apnea
PSG	Polysomnogram/Polysomnography
PSQI	Pittsburgh Sleep Quality Index
RCT	Randomized Controlled Trial
RDI	Respiratory Disturbance Index
RR	Risk Ratio
SAQLI	Sleep Apnea Quality of Life Index
SASQ	Sleep Apnea Symptom Questionnaire
SF-36	Short Form-36, quality of life scale
SSP	Sleep Specialist Physician
US	United States
USPSTF	United States Preventive Services Task Force
VA	Department of Veterans' Affairs
VAS	Visual Analogue Scale
VINCI	Veterans Administration Informatics and Computing Infrastructure
VSQ-9	Visit-specific Satisfaction Instrument

EVIDENCE REPORT

INTRODUCTION

Obstructive sleep apnea (OSA) is a chronic condition that results from repeated closure of the upper airway during sleep resulting in reduced airflow (hypopnea) or complete airflow cessation (apnea) leading to cyclic sleep disruption. Subsequently, patients with OSA frequently experience excessive daytime sleepiness and decreased quality of life. However, not all individuals have excessive daytime sleepiness and symptoms are not required to make a diagnosis or obtain treatment. OSA has also been associated with a higher risk of myocardial infarction,[1,2] heart failure,[3] stroke,[4,5] and cognitive decline.[6,7] The severity of OSA can be categorized as mild, moderate, or severe based on the number of apnea and hypopnea events per hour (known as the apnea-hypopnea index [AHI]). An AHI of 5/hour to fewer than 15/hour is considered mild, 15/hour to fewer than 30/hour is considered moderate, and 30/hour or greater is considered severe. Continuous positive airway pressure (CPAP) effectively reduces AHI in most patients with OSA, improves blood pressure, and particularly in those with symptoms of excessive daytime sleepiness, improves quality of life and sleep symptoms.[8] CPAP use is also associated with a reduced risk of motor vehicle accidents.[9]

The estimated prevalence of mild to severe OSA based on AHI thresholds (AHI ≥ 5/hour) in the United States (2007-2010 data) among 30- to 70-year-olds regardless of symptom status is 34% for men and 17% for women.[10] The corresponding values for moderate to severe OSA (AHI ≥ 15/hour) are 13% for men and 6% for women. The prevalence of 30-70 year olds with AHI ≥ 5/hour and daytime sleepiness (Epworth Sleepiness Scale score > 10) is 14% for men and 5% for women. Current prevalence of abnormal AHI may be higher due to rising rates of obesity and increased testing for OSA due to heightened awareness of the condition. The proportion of persons with OSA who are asymptomatic or have unrecognized symptoms is unknown but a recent review estimated that 80% of individuals with AHI ≥ 5/hour may be undiagnosed.[11] The cost burden of diagnosing and treating OSA in the US in 2015 was estimated to be $12.4 billion while the cost burden of undiagnosed OSA in the US in 2015 was estimated to be $149.6 billion when costs of comorbidities and mental health, motor vehicle accidents, workplace accidents, and lost productivity were considered.[11] The authors acknowledged the difficulty of determining exact costs.

A prior evidence report and accompanying clinical practice guideline by the American College of Physicians (ACP) recommends that clinicians should target their assessment of OSA to individuals with unexplained daytime sleepiness.[12-14] This is largely because evidence is lacking on the effect of CPAP on improving many other outcomes, including diabetes, coronary heart disease events, and mortality, especially among individuals without daytime sleepiness. Furthermore, the ACP concluded that assessment of OSA in the absence of daytime sleepiness and treatment of persons with low AHI are both low-value care because evidence to date indicates that neither improves clinical outcomes.

A draft evidence report and recommendation statement from the US Preventive Services Task Force (USPSTF) indicated that evidence was insufficient to assess the net benefit of screening for and treatment of asymptomatic OSA.[15]

Specifically, the most evidence was available on CPAP and found that compared to sham intervention, CPAP reduced AHI, ESS score, and blood pressure. Although studies generally showed that treatment with CPAP reduced AHI to near-normal levels, the clinical significance of the small reductions in ESS score and blood pressure is uncertain. Further, given that most of the trials were conducted in referred patients or patients of sleep clinics, the applicability of this evidence to a screen-detected population (*ie*, detection of abnormal AHI in asymptomatic individuals or those without excessive daytime somnolence, including based on findings from "screening" questionnaires) is limited. Despite the consistent observational findings of an association between severe OSA and increased mortality, the USPSTF identified no studies that reported on change in AHI and associated change in mortality. Thus, it is unclear whether treatments that improve AHI would also improve mortality. The USPSTF found inadequate evidence on the link between change in the intermediate outcome (AHI) and reduction in the health outcome (mortality). While the USPSTF found evidence that treatment with CPAP can improve general and sleep-related quality of life in populations referred for treatment, the applicability of this evidence to screen-detected populations is unknown. The USPSTF also found inadequate evidence on whether treatment with CPAP improves other health outcomes (mortality, cognitive impairment, motor vehicle accidents, and cardiovascular or cerebrovascular events). Following release of the USPSTF report, results from a multi-site randomized trial were published and showed that CPAP was not effective for secondary prevention of cardiovascular events in those with established cardiovascular disease and moderate to severe OSA.[16]

Despite the data associating OSA with consequences to health and quality of life and the conclusions from other guideline groups, many persons with OSA remain undiagnosed. In 2005, the National Sleep Foundation administered the Berlin sleep apnea risk questionnaire (a questionnaire commonly used for OSA screening and/or case finding) by phone to a random sample of 1,506 US adults who agreed to complete the questionnaire and found that 31% of men and 21% of women met criteria for high OSA risk.[17] Obesity is a major risk factor for OSA and among those with a body mass index (BMI) of ≥ 30 kg/m^2 (25% of the sample), 57% had high risk for OSA.

In 2010, among 1.8 million US Veterans receiving outpatient care at 136 Veterans Affairs (VA) facilities, 37.4% had a BMI ≥ 30 kg/m^2, suggesting that a substantial portion of US Veterans are at high risk for OSA.[18] A recent analysis of Veterans Administration Informatics and Computing Infrastructure (VINCI) data between 2000 and 2010 showed that among 9.8 million Veterans, the age-adjusted prevalence of diagnosed sleep apnea was 0.4% in 2000 and had increased to 3.0% in 2010 (a relative increase of 650%).[19] As awareness of OSA by patients and providers continues to increase, and as BMI continues to increase in the US and globally,[20] healthcare systems such as the VA need to develop strategies to manage the increasing demand for sleep services. One strategy is to target screening and testing to those most likely to derive benefit from OSA treatment (*eg*, those with significant unexplained daytime sleepiness), as suggested by the ACP. Another strategy is to improve efficiency within healthcare systems by implementing innovative, less resource-intense models of care for OSA.

The traditional model of OSA evaluation and care relies upon primary care providers to refer patients with suspected OSA to a sleep specialist physician (SSP) for consultation. After an initial consultative visit, the SSP can order formal, in-lab polysomnogram (PSG) for diagnostic purposes and for those with confirmed OSA, a PSG for titration of CPAP pressures. The SSP would then typically initiate CPAP at the pressure suggested by the titration PSG, and then the

patient would follow up with the SSP at regular intervals for assessment of treatment compliance and efficacy. Given the rapidly rising requests for OSA diagnostic and treatment services, this traditional model is increasingly viewed as unnecessarily expensive and inefficient for the evaluation and treatment of patients at high risk of OSA. Recent data also indicate a decreasing supply of SSPs to care for patients with known or suspected OSA.[21]

Therefore, new models of OSA care have been proposed and implemented. These new models include home sleep testing (HST) for diagnostic purposes,[22] followed by treatment with an autotitrating CPAP (APAP) device,[23] which has internal algorithms to adjust CPAP pressure to keep the airway open during sleep. These models reduce PSG-associated costs and logistical barriers, yet typically still include consultation and follow up with a SSP. Other proposed models would reduce reliance on SSPs by including non-SSP providers such as nurses or primary care clinics to provide the bulk of OSA diagnosis and treatment.

Although several studies have been conducted to test some of these new models, systematic reviews are lacking. The Minneapolis VA's Evidence-based Synthesis Program (ESP) Center, in partnership with topic nominators and a Technical Expert Panel (TEP), was commissioned to systematically review the evidence regarding the comparative effectiveness, harms, and cost of these new models of OSA evaluation and treatment.

We addressed the following key questions:

Key Question 1. For adults with suspected OSA, what are the effectiveness/harms/resource utilization of case finding and care provided by practitioners who are not sleep physicians (including PCPs, PAs, NPs, technologists, nurses, and respiratory therapists), compared to case finding and care provided by sleep specialist physicians?

KQ1A. Do effectiveness/harms/resource utilization vary by patient characteristics:
- a. Unexplained daytime sleepiness/fatigue
- b. AHI severity
- c. Other risk factors or coexisting conditions associated with OSA (*eg,* obesity, neck circumference, treatment-resistant hypertension, diabetes, age, sex, race)
- d. Symptoms (*eg,* nocturia, loss of libido, snoring, witnessed sleep "apnea")?

Key Question 2. For adults with suspected OSA, what are the effectiveness/harms/resource utilization of electronic consultation versus interactive (*eg*, in-person, telephone) consultation?

KQ2A. Do effectiveness/harms/resource utilization vary by patient characteristics:
- a. Unexplained daytime sleepiness/fatigue
- b. AHI severity
- c. Other risk factors or coexisting conditions associated with OSA (*eg,* obesity, neck circumference, treatment-resistant hypertension, diabetes, age, sex, race)
- d. Symptoms (*eg,* nocturia, loss of libido, snoring, witnessed sleep "apnea")?

Key Question 3. For adults diagnosed with OSA, what are the effectiveness/harms/resource utilization (including cost avoidance) of using in-home autotitrating continuous positive airway

pressure (APAP) technology compared to standard continuous positive airway pressure (CPAP) titrated by in-lab PSG?

KQ3A. Do effectiveness/harms/resource utilization vary by patient characteristics:
a. Unexplained daytime sleepiness/fatigue
b. AHI severity
c. Other risk factors or coexisting conditions associated with OSA (*eg*, obesity, neck circumference, treatment-resistant hypertension, diabetes, age, sex, race)
d. Symptoms (*eg*, nocturia, loss of libido, snoring, witnessed sleep "apnea")?

PICOTS AND ANALYTIC FRAMEWORKS (FIGURES 1A, 1B)

Population:

KQ1, KQ2: Adults with suspected obstructive sleep apnea (OSA) to include:
KQ3: Adults with diagnosed OSA

Intervention

KQ1: supervised practitioner or non-specialist licensed independent practitioner-led care (*ie*, PCP, PA, NP, nurse, technologist, or respiratory therapist) (for case finding *or* treatment)
KQ2: Electronic initial consultation (chart review or algorithm) (for case finding)
KQ3: Home auto-titrating continuous positive airway pressure (APAP)

Comparator

KQ1: Sleep specialist-led care (for case finding *or* treatment)
KQ2: Interactive (*eg*, in-person, telephone) initial consultation (case finding)
KQ3: In-center manual CPAP *titration*

Outcomes

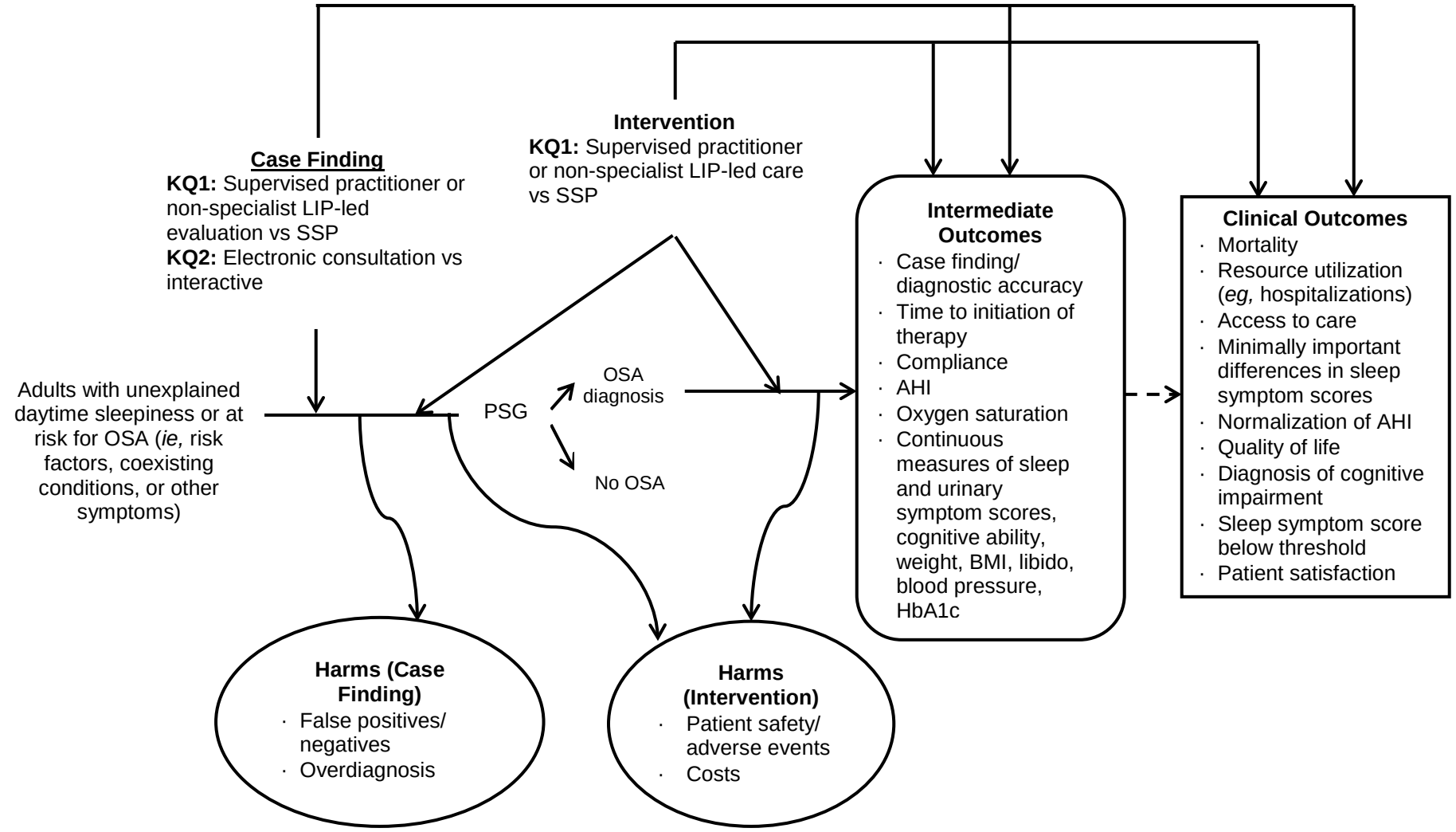
Intermediate Outcomes: time to initiation of therapy; compliance; apnea-hypopnea index (AHI); oxygen saturation; continuous or scale score measures of sleep symptoms, urinary symptoms, cognitive ability, weight, BMI, libido, blood pressure, or HBA1c; harms (false positives/negative, overdiagnosis, patient safety/adverse events); costs

Clinical Outcomes: mortality; resource utilization; access to care; *minimally important differences* in sleep symptom scores, urinary symptom scores, libido, weight change, BMI, blood pressure, or HbA1c; quality of life; diagnosis of cognitive impairment; sleep symptom score below threshold; patient satisfaction

Timing: Any

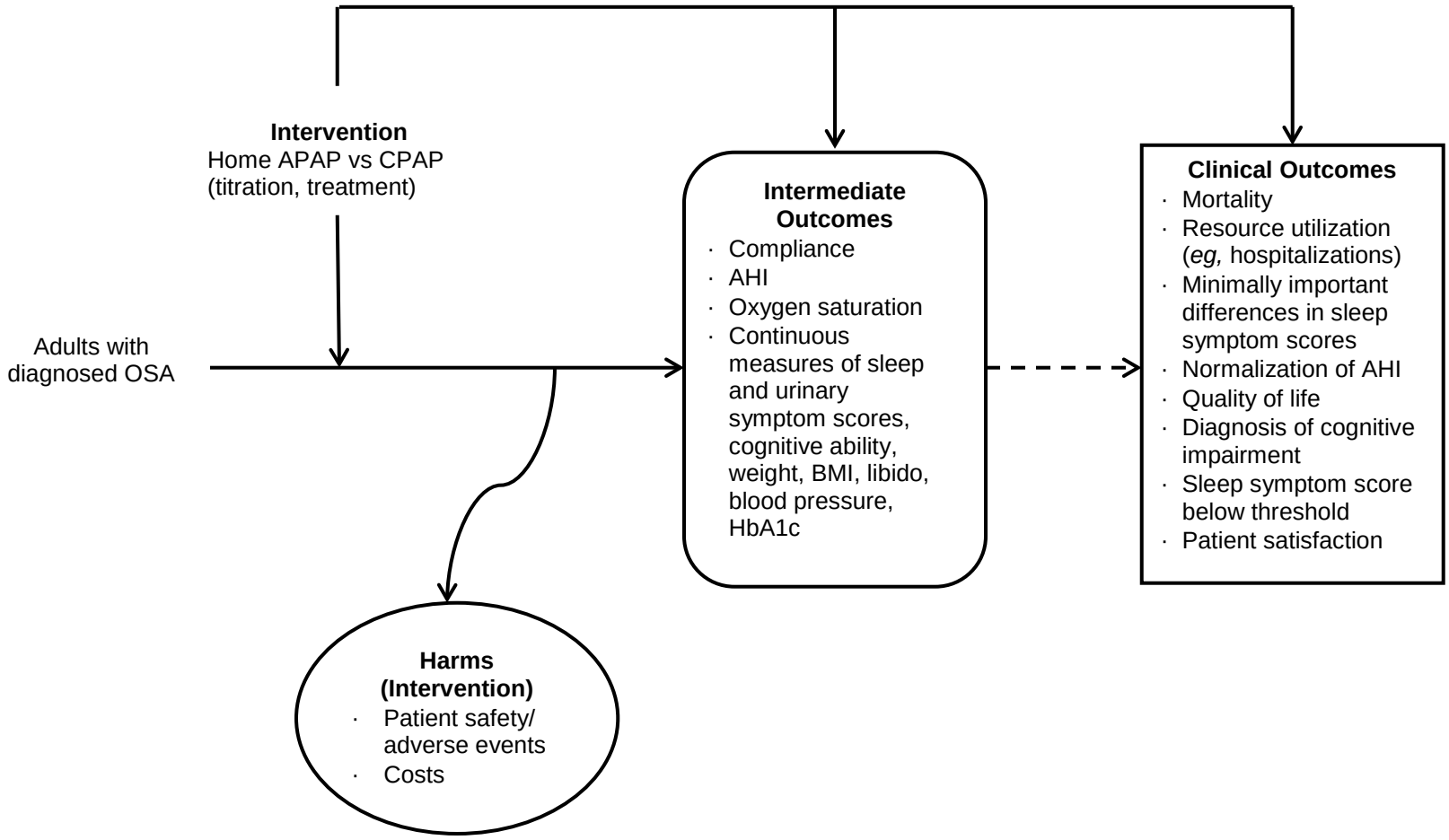
Setting: Study done in North America, Europe, or Australia/New Zealand

Figure 1a. Analytic Framework – Key Questions 1 and 2



Solid Arrow = Linkage; Dotted Arrow = Association; Curved Arrow leads to Harms (Ovals); AHI = Apnea Hypoxia Index; LIP = licensed independent practitioner; OSA = obstructive sleep apnea; PSG = polysomnography; SSP = sleep specialist physician

Figure 1b. Analytic Framework – Key Question 3



Solid Arrow = Linkage; Dotted Arrow = Association; Curved Arrow leads to Harms (Ovals); AHI = Apnea Hypoxia Index; APAP = Auto-adjusting Positive Airway Pressure; CPAP = C Positive Airway Pressure OSA = obstructive sleep apnea

METHODS

TOPIC DEVELOPMENT

This topic was nominated by Kathleen Sarmiento, MD, MPH, Director of Pulmonary Sleep Medicine, VA San Diego and W. Claibe Yarbrough, MD, National Program Director, Pulmonary/Critical Care/Sleep, on behalf of the Specialty Care Services (10P4E) – Pulmonary/Critical Care/Sleep National Program office. The evidence review examines the effectiveness and harms associated with different care models aimed at increasing access to care for Veterans who have obstructive sleep apnea.

SEARCH STRATEGY

We searched MEDLINE (Ovid) and CINAHL for articles published from 2000 through May 2016. Our search was limited to studies enrolling adults and published in the English language. The search for KQs 1 and 2 included the MeSH terms sleep apnea syndromes; sleep apnea, obstructive; health personnel; and remote consultation. The search for KQ3 included the MeSH terms home care services, continuous positive airway pressure, and calibration. The full search strategies are presented in Appendix A. We obtained additional articles by hand searching the reference lists of related systematic reviews and relevant studies.

STUDY SELECTION

All abstracts were independently reviewed by 2 trained investigators and research associates. We included studies of any design that reported results in adults with suspected or diagnosed OSA and were conducted in North America, Europe, Australia, or New Zealand. For KQ1 we excluded studies that did not include a comparison of a supervised practitioner or non-specialist licensed independent practitioner (*eg* primary care physician, physician's assistant, nurse, technologist, or respiratory therapist) to a SSP. We excluded studies evaluating the role of dentists or anesthesiologists. We also excluded studies in which the goal of the intervention was not case finding or care for OSA. For KQ2 we excluded studies that did not compare an electronic initial consultation, without patient contact, to an interactive initial consultation. For KQ3 we excluded studies that did not compare the use of APAP to CPAP for *titration* or *treatment* of OSA. We also excluded studies that used different diagnostic methods in those treated with APAP versus CPAP. In studies of *titration* we only included articles in which the APAP was used for titration at home and CPAP was manually titrated in a lab. We also excluded studies if they did not report any of our outcomes of interest (see PICOTS, above). There was no minimum follow-up duration.

Full-text reports of studies identified as potentially eligible were obtained for further review using the inclusion and exclusion criteria described above. Each article was independently reviewed by 2 investigators or research associates. Reasons for excluding a study at full-text review were noted and disagreements were decided by a third reviewer.

DATA ABSTRACTION

Study characteristics (location, setting, intervention groups, follow-up, aim of study, treatments, inclusion/exclusion criteria, and patient characteristics) as well as intermediate and clinical outcomes (time to initiation of therapy, compliance, AHI, oxygen saturation, sleep symptoms, urinary symptoms, cognitive ability, weight, BMI, libido, blood pressure, HbA1c, harms, overdiagnosis, adverse events, costs, mortality, resource utilization, access to care, quality of life, and patient satisfaction) were extracted onto evidence tables by one investigator or research associate and verified by another.

RISK OF BIAS ASSESSMENT

Trained research methodologists rated the risk of bias of individual studies as low, moderate, or high risk. One methodologist rated risk of bias and the rating was verified by a second investigator trained in risk of bias assessment. Discrepancies were resolved by discussion. For randomized controlled trials (RCTs), risk of bias ratings were based the following criteria: allocation sequence generation, allocation concealment, blinding, incomplete outcome data, and selective outcome reporting – a modification of the Cochrane approach to determining risk of bias.[24]

For observational studies, risk of bias was rated using criteria suggested in the Agency for Healthcare Research and Quality (AHRQ) Methods Guide: selection bias (use of appropriately comparable control group, design/analysis accounted for important confounding and modifying variables); masking of the outcome assessment (outcome assessor); use of intention-to-treat principles (*ie*, inclusion of all comparison group participants in outcomes analyses); attrition bias (if overall or differential dropout/loss to follow-up or exclusions a concern, missing data appropriately handled); and selective reporting of pre-specified outcomes.[25] Observational studies were considered high risk of bias unless all 5 criteria were addressed by the study authors. Studies that addressed all 5 criteria were considered medium or low risk of bias depending on how completely the criteria were addressed.

DATA SYNTHESIS

We created separate evidence tables for each KQ. We described and qualitatively compared the characteristics and findings of included studies. For KQ1 and KQ3, data were analyzed in Comprehensive MetaAnalysis Version 3 (Biostat, Englewood, New Jersey) using DerSimonian and Laird random effects models to calculate weighted mean differences (WMD) for compliance and standardized mean differences (SMD) for changes from baseline for Epworth Sleepiness Scale (ESS) and quality of life (SF-36) scores with corresponding 95% confidence intervals (CI). SMDs can be interpreted by using Cohen's definition of small (0.2), medium (0.5), and large (0.8) effect sizes.[26] Statistical heterogeneity among trials was assessed by using the I^2 test.[27] A score of 75% or greater may indicate considerable heterogeneity.

RATING THE BODY OF EVIDENCE

We assessed strength of evidence using the method described by Owens et al for the following outcomes: access to care, ESS, quality of life, compliance (hours of use per night), and adverse events. Strength of evidence for an outcome was rated as high, moderate, low, or insufficient. This rating was based on precision (degree of certainty in the estimate of effect), consistency

(direction of effect across included studies), directness (whether evidence links intervention directly to health outcomes), and risk of bias of the individual studies (as described above).[28] One methodologist rated strength of evidence and the rating was verified by a second. Discrepancies were resolved by discussion.

PEER REVIEW

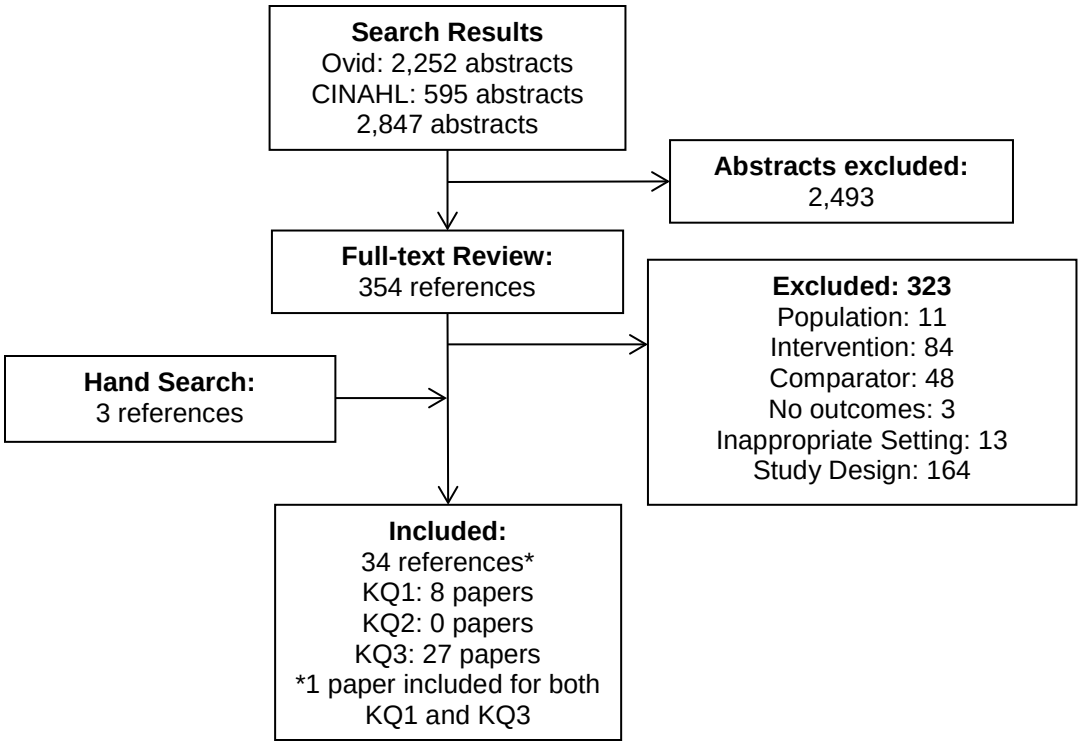
A draft version of this report was reviewed by content experts as well as clinical leadership. Reviewer comments and our responses are presented in Appendix B and the report was modified as needed.

RESULTS

We reviewed 2,847 abstracts, 2,252 from MEDLINE and 595 from CINAHL. We excluded 2,493 abstracts and reviewed the full text of 354 references. During full-text review we excluded 323 articles leaving 31 eligible for inclusion. Hand searching pertinent trials and systematic review identified an additional 3 references. Figure 2 details the process.

LITERATURE FLOW

Figure 2. Literature Flow Chart



KEY QUESTION #1: For adults with suspected OSA, what are the effectiveness/harms/resource utilization of case finding and care provided by practitioners who are not sleep physicians (including PCPs, PAs, NPs, technologists, nurses, and respiratory therapists), compared to case finding and care provided by sleep specialist physicians?

Summary of Findings for Key Question #1

Case Finding

- No studies assessed the diagnostic accuracy of non-sleep-specialist nurse for case finding and referral.
- One retrospective study reported good agreement between a primary care pulmonologist and a SSP on what sleep test to order for patients referred by their family physician.

Care

- Clinical (*ie*, patient-centered) outcomes were infrequently and inconsistently reported. When reported there was no significant difference in clinical outcomes between OSA treated by primary care/nurses and SSPs. The strength of evidence for quality of life was moderate.
- Intermediate outcomes were more commonly reported. Sleep symptom scores were similar between groups (moderate strength of evidence).
- There was little evidence that treatment compliance differed between patients treated by SSPs and those not, including the proportion of patients with 4 hours or more of CPAP use on 70% or more of nights (moderate strength of evidence).
- Very few studies reported other intermediate outcomes. One reported a significantly lower residual AHI on CPAP in patients referred for PSG by non-sleep specialists and another found that the proportion of patients receiving CPAP within one month of their PSG was significantly higher in patients cared for by a SSP. Strength of evidence for access to care and adverse events was insufficient.

Overview of Studies – Table 1

Eight studies (n = 1,401; 4 RCTs) reported results for KQ1.[29-36] Study characteristics are summarized in Table 1 with more details in Appendix C, Table 1. Sleep physician care was compared to management by primary care in 4 studies (n = 564),[31-33,36] sleep-specialist nurses in 3 studies (n = 434),[29,30,34] and other non-sleep physicians in one study (n = 403).[35] The weighted mean age was 52 years and roughly half of the participants (54%) had hypertension. The mean BMI was 34 kg/m^2, ESS was 11.5, and AHI was 32/hour. Three studies (n = 678) reported that their patients had moderate or severe OSA (mean AHI ≥ 15/hour)[29,33,35] and 5 studies (n = 721) reported that their patients had mild OSA (mean ESS 11-14)[30-32] or moderate (mean ESS > 14) sleepiness.[29,33] Three studies (n = 415) required all participants to have symptoms of sleepiness (*ie*, ESS ≥ 8[30,31] or ESS ≥ 12[29]), and participants in these studies had a mean baseline ESS of

13.6 (mild sleepiness). Participants in the remaining 4 studies that reported ESS but did not require sleepiness had a mean baseline ESS of 10.5 (normal sleepiness). Three studies took place in Europe, 3 in North America, and 2 in Australia/New Zealand. Of the 4 RCTs, one was low risk of bias[30] and 3 were medium risk of bias. Of the 4 observational studies, 3 were medium risk of bias and one was high risk of bias.[32]

Table 1. Key Question 1: Study Characteristics

Characteristic	Weighted Mean (range) *Unless otherwise noted*	Number of studies reporting
Total number enrolled/randomized	1,401 (65-403)	8
Randomized controlled trials, total n	589 (65-195)	4
Other, total n	812 (96-403)	4
Primary care, total n	564 (96-210)	4
Sleep-specialist nurse, total n	434 (65-195)	3
Other non-sleep physician provider, total n	403	1
Age of subjects (years)	52.2 (47.7-58.7)	8
Percent Male	67% (47-85.5)	8
Baseline BMI, kg/m^2	34 (30.2-36.6)	7
Baseline ESS	11.5 (8.5-15.7)	7
Baseline AHI (events/hour)	32 (21-43)	3
Percent with hypertension (%)*	54 (38-58.5)	4
Location – North America, total n	716 (103-403)	3
Location – Europe, total n	335 (65-174)	3
Location – Australia/New Zealand, total n	350 (155-195)	2
Required participants to have daytime sleepiness, total n	415 (65-195)	3

*as defined by study

Case Finding

One retrospective study reported on the ability of a primary care pulmonologist to order the proper sleep test.[32] Patients (n = 96) were referred to the primary care pulmonologist by their family physician. An SSP reviewed the sleep tests ordered by the primary care pulmonologist. There was good agreement between the primary care pulmonologist and SSP (kappa = .74) and 93% (89/96) of the referred patients were diagnosed with OSA.

Care

Seven studies reported treatment outcomes in patients being treated by providers other tha SSPs. Three of these studies compared SSP care to primary care (n = 468; 1 RCT), 3 compared SSP care to sleep specialist nurses (n = 434; 3 RCTs), and one compared SSP care to management by a variety of physicians who were not sleep specialists. The nature of care provided by different professionals was varied and often inconsistently reported. Two of the studies (n = 506) were retrospective and gave few details regarding the care given by each practitioner.[35,36] Three studies (n = 560; 2 RCTs and one cohort study) described interventions in which the SSP had much more autonomy than the non-SSP provider, who was generally giving protocol-driven care that followed clinical guidelines.[30,31,33] Two RCTs compared patients who received similar care delivered by different providers and at different locations (home versus a hospital/clinic).[29,34] Many of the articles only described the care given to patients in non-specific terms and often details were only given for one study arm.

Clinical Outcomes – Table 2; Appendix C, Tables 2-3

Clinical outcomes were only reported sporadically. The most commonly reported clinical outcome was resource utilization (k = 5). Two studies reported receipt of treatment finding no significant difference in the proportion of patients using CPAP at follow-up.[31,36] One study found that patients whose follow-up appointments were at the hospital, as opposed to at home, often needed additional help from a specialist nurse for practical problems, although the significance and impact of this finding was not reported.[34] Three studies reported provider contact. One RCT found that patients being managed by a sleep specialist nurse had significantly more scheduled nursing time than patients being managed by a SSP (153 minutes vs 103 minutes, P<.001, effect size 1.25, 95% CI 0.94, 1.56) and patients managed by SSPs had significantly more physician visits than patients receiving nurse-led care (2.4 vs 0.2; P<.001, effect size 2.24, 95% CI 1.88, 2.60).[30] A second RCT compared the time nurse-managed patients spent with a nurse to the time SSP-managed patients spent with a SSP and found groups were similar (effect size 0.32, 95% CI -0.13, 0.66).[34] A third study reported extra visits and calls required by patients but without a measure of statistical significance.[29] We had difficulty determining what would constitute *improved* resource utilization and ultimately, OSA management by a non-SSP had no significant effect on any measure of resource utilization as compared to SSP care.

Three studies reported quality of life and all found that SF-36 scores were similar for patients being managed by primary care (k = 1)[31] or sleep-specialist nurses (k = 2)[30,34] compared to SSPs. The pooled standard mean differences in change from baseline for the mental health and vitality components of the SF-36 are presented in Figure 3.

The 2 studies reporting patient satisfaction both found no significant difference in overall satisfaction between groups.[30,31] One study reported significant differences on several VSQ-9 items but cautioned that the effect sizes were all small and may not be clinically significant.[31] No study reported all-cause mortality, access to care, minimally important differences in symptom scores, cognitive outcomes, libido, or percent of patients achieving physiological targets.

Table 2. Clinical Outcomes Comparing Non-sleep Specialists to Sleep Specialists

Author, Year Enrolled/ Randomized (n) Study Design	All-cause Mortality	Normalization of AHI	MID		Quality of Life		Patient Satisfac-tion		Resource Utilization				Access to Care	Libido	% Achieving Target				Cognitive Symptoms	Other
			ESS	Urinary Symptom Scores	SF-36	SAQLI	VAS	VSQ-9	Hospitalization	Treatment	Provider Contact	Referrals			Weight Loss	BMI	Blood Pressure	HbA1c		
Primary Care *3 studies (n = 468)*																				
Chai-Coetzer 2013 [31] (n = 155); RCT					↔			↕		↔										
Lettieri 2011[33] (n = 210); Obs. cohort																				
Scharf 2004[36] (n = 103); Retro survey/chart review										↔										↔[a]
Nurse *3 studies (n = 434)*																				
Andreu 2012[29] (n = 65); RCT											X									
Antic 2009[30] (n = 195); RCT					↔			↔			↕									
Palmer 2004[34] (n = 174); RCT					↔						↔	X								↔[b]
Not specified *1 study (n = 403)*																				
Pamidi 2012[35] (n = 403); Retro chart review																				
Total	0	0	0	0	3↔	0	0	1↔ 1↕	0	2↔	1↔ 1↕ 1X	1X	0	0	0	0	0	0	0	2↔

[a]subjective symptom improvement, [b]HADS = Hospital Anxiety and Depression Scale

MID = minimally important difference; ESS = Epworth Sleepiness Scale; BMI = body mass index; SF-36 = 36 item short form survey; SAQLI = Calgary Sleep Apnea Quality of Life Index; VAS = visual analogue scale; VSQ-9 = Visit-Specific Satisfaction Instrument; HbA1c = hemoglobin A1c; retro = retrospective; Obs = observational

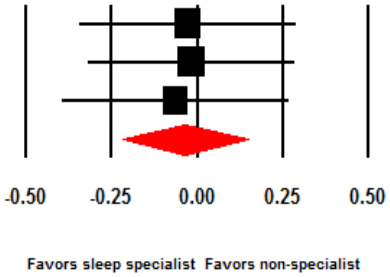
↑ = significantly better with non-SSP than with usual care SSP including less resource utilization, ↔ = non-significant difference between non-SSP and SSP care, ↓ = significantly better with SSP than non-SSP, ↕ = mixed results comparing non-SSP and SSP care, X = between provider significance not reported; number before symbol indicates number of studies

Light gray shading in a cell indicates measure of significance was calculated, not reported by the study.

Figure 3. SF 36 Mental Health and Vitality Scores, Standardized Mean Differences in Mean Change from Baseline, Non-sleep Specialist versus Sleep Specialist[a]

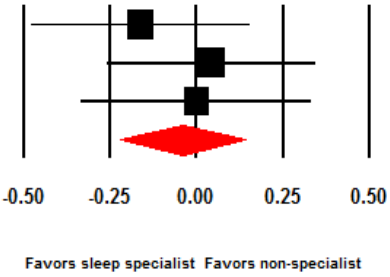
Mental Health

Study name	Non-specialist	Std diff in means	Lower limit	Upper limit	Non-specialist	Sleep specialist
Chai-Coetzer 2013 (31)	PCP	-0.03	-0.34	0.29	81	74
Antic 2009 (30)	Nurse	-0.02	-0.32	0.28	89	81
Palmer 2004 (34)	Nurse	-0.06	-0.40	0.27	68	71
		-0.04	-0.22	0.15		



Vitality

Study name	Comparison	Std diff in means	Lower limit	Upper limit	Non-specialist	Sleep specialist
Chai-Coetzer 2013 (31)	PCP	-0.16	-0.48	0.15	81	74
Antic 2009 (30)	Nurse	0.04	-0.26	0.34	89	81
Palmer 2004 (34)	Nurse	0.00	-0.33	0.33	68	71
		-0.04	-0.22	0.14		



[a]Lower and upper limits represent 95% confidence intervals
PCP = primary care physician

Intermediate Outcomes – Table 3; Appendix C, Tables 4-6

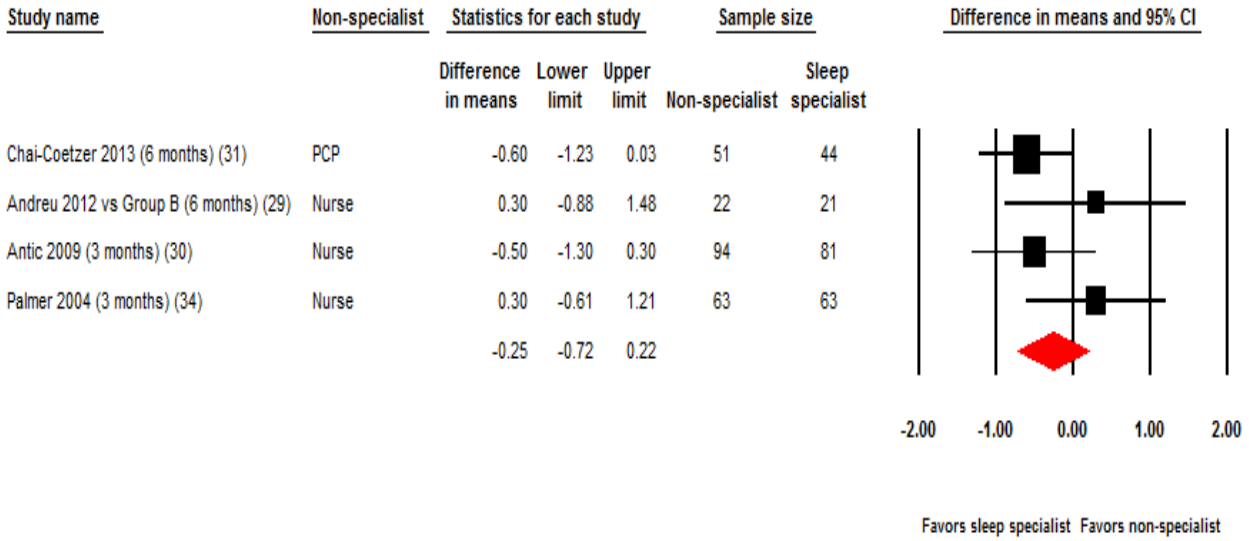
The most frequently reported intermediate outcome was compliance with therapy, reported by 7 studies with some reporting multiple indicators of compliance. The majority of studies reported adherence as hours of use per night (k = 6).[29-31,33-35] Several reported the proportion of patients with regular use (k = 4), defined as ≥ 4 hours of CPAP use on ≥ 70% of nights, a compliance threshold often used to define minimally acceptable compliance for payment reimbursement.[29,33,35,36] One reported the proportion of nights with any use.[33] Six of the 7 studies found no difference in compliance, regardless of measure used, when comparing patients receiving SSP care to those receiving care from non-SSPs. The pooled mean difference from the RCTs was -0.25 (95% CI -0.72, 0.22; I^2 = 21%) (Figure 4). The final study found that patients who were referred for a sleep study by non-SSPs were significantly less compliant, with fewer hours per night and less regular use, than those patients who were referred by SSPs.[35] Cost was reported, in various ways, by 4 of the studies.[29-31,34] Two did not report the significance of differences between SSP and non-SSP care for OSA.[31,34] Two studies, however, found that nurse-led OSA care was associated with significantly lower costs per patient[29] and within-trial costs.[30]

Table 3. Intermediate Outcomes Comparing Non-sleep Specialists to Sleep Specialists

Author, Year Enrolled/ randomized (n) Study design	AHI	Oxygen Saturation: Mean	Oxygen Saturation: Time <90%	Oxygen Saturation: ODI	Oxygen Saturation: Minimum	Sleep Symptom Scores: ESS	Sleep Symptom Scores: FOSQ	Sleep Symptom Scores: SASQ	Sleep Symptom Scores: Arousals	Sleep Symptom Scores: TST	Sleep Symptom Scores: PSQI	Sleep Symptom Scores: Snoring	Sleep Symptom Scores: Other	Weight Loss	BMI	Blood Pressure	HbA1c	Time to Initiation of Therapy	Adverse Events	Cost	Compliance/ Adherence: Hrs/night	Compliance/ Adherence: Proportion of Nights	Compliance/ Adherence: Regular Use[c]	Compliance/ Adherence: Other
Primary Care *3 studies (n = 468)*																								
Chai-Coetzer, 2013 [31] (n = 155); RCT						↔	↔	↔						↔		↔				X	↔			
Lettieri, 2011[33] (n = 210); Obs. cohort						↔													↔		↔	↔	↔	
Scharf, 2004[36] (n = 103); Retro survey/chart review																		↓					↔	
Nurse *3 studies (n = 434)*																								
Andreu, 2012[29] (n = 65); RCT						↔	↕												X	↑[a]	↔		↔	
Antic, 2009[30] (n = 195); RCT						↔	↔						↔[b]					↕		↑	↔			
Palmer, 2004[34] (n = 174); RCT						↔							↔							X	↔			
Not specified *1 study (n = 403)*																								
Pamidi, 2012[35] (n = 403); Retro chart review	↑																				↓		↓	
Total	1↑					5↔	2↔ 1↕	1 ↔					2 ↔	1 ↔		1 ↔		1↓ 1↕	1 ↔ 1 X	2 ↑ 2 X	5 ↔ 1↓	1↔	3 ↔ 1↓	

[a]decreased costs
[b]maintenance of wakefulness
[c] ≥ 4h of use on ≥ 70% of nights
AHI = apnea-hypopnea index; ODI = oxygen desaturation index; ESS = Epworth Sleepiness Scale; FOSQ = Functional Outcomes of Sleep Questionnaire; SASQ = Sleep Apnea Symptom Questionnaire;TST = total sleep time; PSQI = Pittsburgh Sleep Quality Index; BMI = body mass index; retro = retrospective; Obs = observational
↑ = significantly better with non-SSP than with usual care SSP including less resource utilization, ↔ = non-significant difference between non-SSP and SSP care, ↓ = significantly better with SSP than non-SSP, ↕ = mixed results comparing non-SSP and SSP care, X = between provider significance not reported; number before symbol indicates number of studies
Light gray shading in a cell indicates measure of significance was calculated, not reported by the study.

Figure 4. Compliance, Weighted Mean Difference for Hours per Night, Non-sleep Specialist versus Sleep Specialist[a]

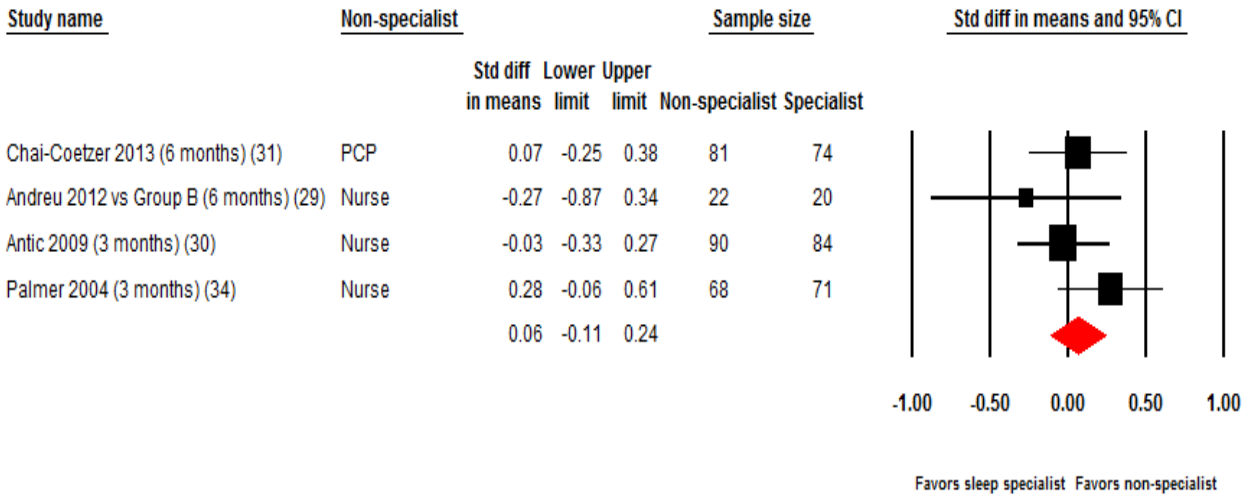


[a]Lower and upper limits represent 95% confidence intervals
PCP = primary care physician

Sleep symptom scores were reported by 5 studies.[29-31,33,34] All 5 reported ESS scores and all found similar results for groups receiving care from different providers. For the RCTs, the SMD for improvement from baseline was 0.06 (95% CI -0.11, 0.24; I^2 = 2%) (Figure 5) and the weighted mean difference (WMD) was 0.30 (95% CI -0.50, 1.10). Three studies also reported FOSQ and one reported SASQ scores; 2 reported no significant between group differences in either score.[30,31] The third study, which used a 3-arm design, found that patients with sleep unit nurse-led follow-up care had lower (*ie*, worse) FOSQ scores than one of the 2 groups of patients with sleep pulmonologist follow-up.[29] The difference was small (one point) but statistically significant. There was no difference between the nurse-followed group and the other pulmonologist group. AHI was reported by one study, which found that residual AHI on CPAP was significantly lower in patients referred for PSG by non-SSPs than in patients referred by SSPs (P<.001).[35]

One RCT reported dryness (54%), nasal congestion (40%), and abrasions (25%) but did not provide separate data for the 2 study groups.[29] An observational study reported no significant differences between groups in the percentages of patients who discontinued therapy although reasons for discontinuing were not reported.[33] In one study, changes in weight loss and blood pressure were similar with primary care management of OSA as compared to SSP care.[31] Time to initiation of therapy was reported in 2 studies. One found that significantly fewer patients in the primary care group received CPAP within one month of PSG when compared to patients in the SSP group (P = .012).[36] The other reported that while groups were similar in satisfaction with time waiting (P = .71), patients receiving nurse-led care were more satisfied with their impression of wait time (P = .004).[30] No studies reported oxygen saturation, HbA1c, or BMI.

Figure 5. Epworth Sleepiness Scores, Standardized Mean Differences for Mean Change from Baseline, Non-sleep Specialist versus Sleep Specialist[a]



[a]Lower and upper limits represent 95% confidence intervals
PCP = primary care physician

KEY QUESTION #1A: Do effectiveness/harms/resource utilization vary by patient characteristics:

a. Unexplained daytime sleepiness/fatigue

b. AHI severity

c. Other risk factors or coexisting conditions associated with OSA (*eg*, obesity, neck circumference, treatment-resistant hypertension, diabetes, age, sex, race)

d. Symptoms (*eg*, nocturia, loss of libido, snoring, witnessed sleep "apnea")?

Few studies provided information to address KQ1A.

A. Unexplained Daytime Sleepiness/Fatigue

Of the studies eligible for KQ1, 2 required patients to have at least mild daytime sleepiness (ESS ≥ 8)[30,31] and one required an ESS ≥ 12.[29] Another study did not report ESS but noted that 68% of patients reported excessive sleepiness at baseline.[36] No study reported results for subgroups of patients based on daytime sleepiness. In addition, with sporadic outcome reporting, it was not possible to determine whether results were different in studies that required a measure of sleepiness for study inclusion versus those that did not.

B. AHI Severity

Four studies reported baseline AHI. Values ranged from 21/hour[33] to 68/hour.[30] In one of the studies, 55% were diagnosed with severe OSA (AHI ≥ 30/hour).[35] Similar to the finding for daytime sleepiness, no study reported results for subgroups of patients based on AHI and it was not possible to determine whether results were different in the studies with higher baseline AHI values.

C. Other Risk Factors or Coexisting Conditions Associated with OSA

Each of the 7 studies with newly evaluated/diagnosed patients reported BMI. Values ranged from 30 kg/m[233] to 36 kg/m^2.[35] One study reported neck circumference (mean of 45.5 cm).[29] No study reported treatment resistant hypertension but 4 reported percentages of study participants with hypertension ranging from 38%[32] to 58%.[35] Three reported percentages of participants with diabetes (14% to 26%).[32,35,36] As noted previously, mean age in the 8 studies was 52 years with little variation across studies (range 48 to 59 years). The mean percentage of male study participants was 67% but values ranged from 47% to 85%. No study reported results for subgroups of patients based on obesity, neck circumference, hypertension, diabetes, age, or gender.

One study from the US (Chicago) reported race with 54% African American and 46% non-African American.[35] In a model adjusted for age, sex, BMI, Medicaid insurance, AHI, ESS, Center for Epidemiologic Studies Depression Scale, and education level, African Americans

used CPAP an average of 56 minutes/day less than non-African Americans (P = .002) It was not reported whether there was an interaction with physician specialty.

D. Symptoms

Four studies addressed snoring. In 2 studies, snoring was an inclusion criterion with "habitual snoring" being a high-risk feature in one study[33] and history of snoring "most" or "every" night being a factor in patient referral in another study.[30] Two studies reported the percentage of patients with snoring at baseline: 100%[29] and 83%.[36] None of the studies reported results for subgroups of patients based on snoring as an inclusion criterion or a baseline factor.

Strength of Evidence for Key Question 1 (Table 4, Appendix D)

There was insufficient evidence for access to care and adverse events. Strength of evidence for quality of life, ESS, and CPAP compliance (hours of use per night) was rated as moderate. Quality of life, ESS, and compliance were similar for patients managed by practitioners who are not sleep physicians compared to patients managed by SSPs.

Table 4. Strength of Evidence for Outcomes, Key Question 1

Comparison	Outcome of Interest	Strength of Evidence	Direction	Findings
KQ1: Sleep physician care compared to management by primary care, sleep-specialist nurses or other non-sleep physicians	Access to care	Insufficient		We found no evidence for this outcome.
	Epworth Sleepiness Scale (ESS)	Moderate	Similar	Based on 4 RCTs (n = 568) with aggregate moderate risk of bias, we found improvement from baseline in ESS scores was similar for patients being managed by primary care/sleep-specialist nurses compared to SSPs (SMD = 0.06 [95% CI -0.11, 0.24]). One observational study also found ESS scores were similar between groups.
	Quality of life	Moderate	Similar	Based on 3 RCTs (n = 524) with aggregate moderate risk of bias, we found quality of life measures were similar for patients being managed by primary care/ sleep-specialist nurses compared to SSPs. SMDs for SF-36 Vitality and Mental Health scores were -0.04 [95% CI -0.22, 0.15]) and -0.04 [95% CI -0.22, 0.14], respectively.
	Compliance, hours per night	Moderate	Similar	Based on 4 RCTs (n = 568) with aggregate moderate risk of bias, we found compliance was similar for patients being managed by primary care/sleep-specialist nurses compared to SSPs (WMD = -0.25 [95% CI -0.72, 0.22]). One observational cohort study also found compliance was similar between groups but one study based on retrospective chart review reported compliance was greater in the SSP group compared to the non-sleep specialist group.
	Adverse events	Insufficient		Based on the findings of one RCT (n = 65) that did not report adverse events by treatment arm, the evidence is insufficient to draw conclusions.

APAP = Auto-adjusted (autoregulated) continuous positive airway pressure; CPAP = continuous positive airway pressure; RCT = randomized controlled trial; SAQLI = Sleep Apnea Quality of Life Index; SF-36 = Short Form-36; SMD = standardized mean difference; WMD = weighted mean difference.

[a]Strength of Evidence Definitions;[28] See Appendix D for more details

- High: Very confident that estimate of effect lies close to true effect. Few or no deficiencies in body of evidence, findings believed to be stable.
- Moderate: Moderately confident that estimate of effect lies close to true effect. Some deficiencies in body of evidence; findings likely to be stable, but some doubt.
- Low: Limited confidence that estimate of effect lies close to true effect; major or numerous deficiencies in body of evidence. Additional evidence necessary before concluding that findings are stable or that estimate of effect is close to true effect.
- Insufficient: No evidence, unable to estimate an effect, or no confidence in estimate of effect. No evidence is available or the body of evidence precludes judgment.

KEY QUESTION #2: For adults with suspected OSA, what are the effectiveness/harms/resource utilization of electronic consultation versus interactive (*eg*, in-person, telephone) consultation?

Key Question #2A: Do effectiveness/harms/resource utilization vary by patient characteristics:

a. Unexplained daytime sleepiness/fatigue

b. AHI severity

c. Other risk factors or coexisting conditions associated with OSA (*eg*, obesity, neck circumference, treatment-resistant hypertension, diabetes, age, sex, race)

d. Symptoms (*eg*, nocturia, loss of libido, snoring, witnessed sleep "apnea")?

We found no studies that addressed KQ2 or KQ2a.

KEY QUESTION #3: For adults diagnosed with OSA, what are the effectiveness/harms/resource utilization (including cost avoidance) of using in-home autotitrating continuous positive airway pressure (APAP) technology compared to standard continuous positive airway pressure (CPAP) titrated by in-lab PSG?

Summary of Findings for Key Question 3

Titration

- Few studies compared clinical (ie, patient-centered) outcomes between in-lab CPAP titration and at-home APAP titration. In limited reporting, study groups were generally similar on measures of quality of life (moderate strength of evidence) and cognitive symptoms. Some differences were noted for resource utilization and patient preference.

- Intermediate outcomes (ie, sleep measures, blood pressure, adverse events, and compliance/adherence) were more commonly reported and generally similar. Strength of evidence for ESS was moderate and strength of evidence for compliance was low. Strength of evidence for adverse events was insufficient.

Treatment

- Twenty-three studies compared treatment with CPAP to treatment with APAP. The studies enrolled patients with a broad range of baseline AHI values.

- Few studies reported clinical (ie, patient-centered) outcomes other than quality of life and patient preference for one treatment approach over another. Quality of life, assessed with the SF-36, was generally similar between the CPAP and APAP groups (moderate strength of evidence). Patient preference was also generally similar or favored APAP in studies reporting statistical significance. Strength of evidence was insufficient for access to care.

- Intermediate outcomes including post-treatment ESS scores were frequently reported and generally similar for the 2 treatment approaches (moderate strength of evidence). Adverse events were mild and similar for APAP and CPAP (low strength of evidence).

- Compliance, reported as either hours per night or the proportion of nights the device was used, was also similar for the CPAP and APAP treatment groups (moderate strength of evidence).

Titration

Overview of Studies – Table 5; Appendix C, Table 7

Four studies compared *titration* of positive airway pressure using in-lab CPAP versus at-home APAP to determine a final long-term CPAP pressure setting.[33,37-39] Three were RCTs conducted in Canada,[39] Australia,[38] and Spain.[37] The fourth study, a cohort study, was done in the US.[33] Follow-up periods ranged from 4 to 12 weeks. Sample sizes ranged from 68[39] to 245[37] at baseline but at least 10% of the sample was not included in the final analysis of each study. Each of the studies was rated medium risk of bias.

Table 5. Key Question 3, Titration: Study Characteristics

Characteristic	Weighted Mean (range) *Unless otherwise noted*	Number of studies reporting
Total number enrolled/randomized	622 (68-245)	4
Randomized controlled trials, total n	482 (68-245)	3
Other, total n	140	1
Age of subjects (years)	50 (46-54)	4
Percent Male	76 (67-88)	4
Baseline BMI (kg/m^2)	33 (29.3-38.5)	4
Baseline ESS	14.7 (14.0-15.6)	3
Baseline AHI (events/hour)	43 (21.2-62.3)	3
Percent with hypertension (%)*	45 (32-56)	2
Location - North America, total n	208 (68-140)	2
Location - Europe, total n	213	1
Location - Australia/New Zealand, total n	169	1

*as defined by study

In 3 of the studies, all of the study participants had a diagnostic PSG.[33,37,38] In the cohort study, those diagnosed with OSA (defined as an AHI > 5/hour with compatible symptoms) and eligible for the home sleep study program (having at least 2 high-risk features such as habitual snoring and daytime fatigue) were included.[33] The RCT from Australia required patients to have symptomatic OSA, defined as ESS ≥ 8 and AHI ≥ 15/hour.[38] The RCT from Spain also required patients to have symptomatic OSA, defined as ESS ≥ 12 and AHI ≥ 30/hour.[37]

In the remaining RCT, patients with clinical suspicion of moderate to severe OSA were evaluated using a diagnostic algorithm. Patients with ESS ≥ 10, Sleep Apnea Clinical Score (SACS) ≥ 15, and RDI ≥ 15/hour (respiratory disturbance index, a measure similar to AHI) were eligible for the study.[39] The group assigned to in-lab CPAP titration underwent PSG to evaluate the performance of the diagnostic algorithm and 34 of 36 patients had an AHI ≥ 15/hour (probability of moderate to severe OSA 0.94 (95% CI 0.81, 0.99)).

Baseline ESS scores were similar in all 4 studies, ranging from 14.0 to 15.5 with the highest score in the study requiring an AHI ≥ 30/hour.[37] Baseline AHI values ranged from 21/hour[33] to 62/hour.[37]

Clinical Outcomes - Table 6, Appendix C Tables 8-9

None of the studies reported all-cause mortality; minimally important differences in AHI, ESS, or urinary symptom scores; patient satisfaction; hospitalizations; access to care; libido; or percent achieving targets for weight loss, BMI, blood pressure, or HbA1c. Quality of life measures were most commonly reported including SF-36 (2 RCTs) and the Sleep Apnea Quality of Life (SAQLI) (1 RCT).

In the study from Spain, SF-36 scores were reported as change from baseline. There was a significant (P<.01) improvement from baseline for both groups in the physical score while the

mental score improved from baseline only for the in-lab CPAP titration group.[37] The Australian study found CPAP and APAP titration resulted in similar scores at the 4-week follow-up for both the physical and mental components of the SF-36.[38] Standard mean difference for the change from baseline in the mental and physical health components of the SF-36 are depicted in Figure 6.

One RCT reported several resource utilization outcomes.[38] During the 4-week study period, there were more humidifiers issued to the home APAP group but fewer chin straps. Technologist staff time for education was similar in the 2 groups as was the time required for follow-up clinics and telephone calls. Technologist time was significantly higher for the home APAP group compared to the in-lab titration group on titration morning. Technologist time for the home APAP group included time needed to download data from the home device. Physician time for titration study reporting was significantly higher in the in-lab titration group compared to the home APAP group as the physician analyzed the manual titration results to determine the fixed pressure for CPAP therapy. Physician time for follow-up clinics was similar.

The same study reported cognitive outcomes.[38] Baseline and 4-week scores on the Trails A and Trails B cognitive function tests were similar in the in-lab CPAP and home APAP groups.

One study reported patient preference.[39] Sixty-two percent in the in-lab study group would have preferred home management while 6% of the home groups would have preferred in-lab management.

Table 6. Clinical Outcomes Comparing APAP to CPAP, Titration Studies

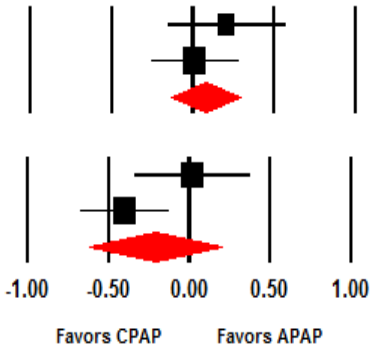
Author, Year Enrolled/ Randomized (n) Study Design	All-cause Mortality	Normalization of AHI	MID		Quality of Life		Patient Satisfac-tion		Resource Utilization				Access to Care	Libido	% Achieving Target				Cognitive Symptoms	Patient Preference	Other
			ESS	Urinary Symptom Scores	SF-36	SAQLI	VAS	VSQ-9	Hospitalization	Treatment	Provider Contact	Referrals			Weight Loss	BMI	Blood Pressure	HbA1c			
Lettieri 2011[33] (n = 140) Obs. Cohort																					
McArdle 2010[38] (n = 169) RCT					↔					↕	↓								↔		
Mulgrew 2007[39] (n = 68) RCT						↔														X	
Masa 2004[37] (n = 245) RCT					↕																
Total *4 studies* *(n = 622)*	0	0	0	0	1↔ 1↕	1↔	0	0	0	1↕	1↓	0	0	0	0	0	0	0	1↔	1X	0

MID = minimally important difference; AHI = apnea-hypopnea index; ESS = Epworth Sleepiness Scale; BMI = body mass index; SF-36 = 36 item short form survey; SAQLI = Sleep Apnea Quality of Life Index; VAS = visual analogue scale; VSQ-9 = Visit-Specific Satisfaction Instrument; HbA1c = hemoglobin A1c; Obs = observational

↑ = significantly better with APAP than CPAP, ↔ = non-significant difference between APAP and CPAP, ↓ = significantly better with CPAP than APAP, ↕ = mixed results comparing APAP and CPAP, X = between group significance not reported; number before symbol indicates number of studies

Figure 6. SF 36 Scores, Standardized Mean Differences in Mean Change from Baseline, Titration Studies[a]

Study name	Comparison	Std diff in means	Lower limit	Upper limit	Sample size APAP	Sample size CPAP
McArdle 2010 (38)	Mental	0.21	-0.16	0.57	58	60
Masa 2004 (37)	Mental	0.01	-0.26	0.28	106	107
		0.08	-0.14	0.29		
McArdle 2010 (38)	Physical	0.02	-0.34	0.38	58	60
Masa 2004 (37)	Physical	-0.40	-0.67	-0.13	106	107
		-0.21	-0.61	0.20		



[a]Lower and upper limits represent 95% confidence intervals

Intermediate Outcomes – Table 7; Appendix C, Tables 10-12

Changes in sleep measures were commonly reported with the in-lab CPAP titration and home APAP titration groups, and found to be similar at follow-up. Three RCTs reported AHI values on PAP were similar at 4 weeks[38] and 12 weeks.[37,39] Two RCTs reported oxygen saturation outcomes finding the groups to be similar at 4 weeks[38] and 12 weeks.[37] All 4 studies reported ESS scores and found the titration groups were similar at follow-up. For the 2 RCTs reporting mean ESS scores, the SMD in change in ESS scores from baseline was 0.00 (95% CI -0.22, 0.21; I^2 = 0%) (Figure 7) and the WMD was -0.01 (95% CI -1.11, 1.10).[37,38]

Other sleep symptom measures included the FOSQ,[37] Arousal Index,[37,38] and total sleep time;[38] these were also similar at follow-up for the in-lab CPAP and home APAP groups.

One study reported blood pressure.[38] At 4 weeks follow-up, the in-lab CPAP and home APAP groups had similar systolic and diastolic blood pressures. No studies reported on weight loss, BMI, or HbA1c.

The cohort study reported on discontinuation of therapy during the 4-6 week follow-up and found similar discontinuation rates – 8.6% of the in-lab CPAP group and 10% of the home APAP groups discontinued treatment.[33] The RCT from Spain reported "secondary effects" (*eg*, oral dryness, mask intolerance, noise, headache, claustrophobia, smothering sensations, bed partner intolerance) noting that there were no important differences between the in-lab CPAP and home APAP groups.[37]

One study reported staff costs per patient and capital equipment and consumable costs per patient. Both were higher for the in-lab CPAP group but the statistical significance of the findings was not reported.[38]

Each of the studies reported on PAP use (*ie*, hours per night). Three studies reported that the in-lab CPAP and home APAP groups were similar[33,37,38] while one reported significantly (P = .02) more hours per night (median 6.0 vs 5.4) in the home APAP group.[39] Mean differences from the RCTs are presented in Figure 8. The pooled mean difference was 0.02 (95% CI -0.41, 0.45).

Other reported adherence/compliance measures were similar for the 2 groups including percentage of nights used,[33] use of device for more than 4 hours per night for more than 70% of nights,[33] and percentage of patients continuing to use device at 4 week follow-up.[38]

One study reclassified participants into "adherent" and "non-adherent" groups based on use of the device for more than 4 hours per night for more than 70% of nights.[33] Age, gender, BMI, AHI, and baseline sleepiness did not influence use of the PAP device.

Table 7. Intermediate Outcomes Comparing APAP to CPAP, Titration Studies

Author, Year Enrolled/Randomized (n) Study Design	AHI	Oxygen Saturation				Sleep Symptom Scores							Weight Loss	BMI	Blood Pressure	HbA1c	Adverse Events	Costs	Compliance/ Adherence			
		Mean	Time <90%	ODI	Minimum	ESS	FOSQ	SASQ	Arousals	Total Sleep Time	PSQI	Snoring	Other							Hrs/night	Proportion of Nights	Regular Use[a]	Other
Lettieri, 2011[33] (n = 140) Obs. Cohort						↔											↔		↔	↔	↔	
McArdle, 2010[38] (n = 169) RCT	↔	↔	↔			↔			↔	↔					↔			X	↔			↔[b]
Mulgrew, 2007[39] (n = 68) RCT	↔					↔													↑			
Masa, 2004[37] (n = 245) RCT	↔		↔			↔	↔		↔								↔		↔			
Total *4 studies (n = 622)*	3↔	1↔	2↔	0	0	4↔	1↔	0	2↔	1↔	0	0	0	0	0	1	0	2↔	2X	1↑ 3↔	1↔	1↔	1↔

[a]as defined by study

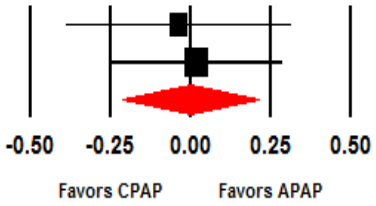
[b]percentage of patients continuing to use CPAP

AHI = apnea-hypopnea index; ODI = oxygen desaturation index; ESS = Epworth sleepiness scale; FOSQ = functional outcomes of sleep questionnaire; PSQI = Pittsburgh sleep quality index; BMI = body mass index

↑ = significantly better with APAP than CPAP, ↔ = non-significant difference between APAP and CPAP, ↓ = significantly better with CPAP than APAP, ↕ = mixed results comparing APAP and CPAP, X = between group significance not reported; number before symbol indicates number of studies

Figure 7. Epworth Sleep Scores, Standardized Mean Difference for Mean Change from Baseline, Titration Studies[a]

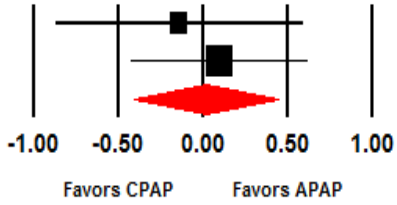
Study name	Std diff in means	Lower limit	Upper limit	APAP	CPAP
McArdle 2010 (4 weeks) (38)	-0.04	-0.39	0.32	62	61
Masa 2004 (12 weeks) (37)	0.02	-0.25	0.29	106	107
	-0.00	-0.22	0.21		



[a]Lower and upper limits represent 95% confidence intervals

Figure 8. Compliance, Weighted Mean Difference for Hours per Night, Titration Studies[a]

Study name	Difference in means	Lower limit	Upper limit	APAP	CPAP
McArdle 2010 (4 weeks) (38)	-0.14	-0.87	0.59	68	70
Masa 2004 (12 weeks) (37)	0.10	-0.42	0.62	106	107
	0.02	-0.41	0.45		



[a]Lower and upper limits represent 95% confidence intervals

Treatment

Overview of Studies –Table 8; Appendix C, Table 7

We included 23 studies (with 1,260 patients) comparing CPAP to APAP for *treatment* of OSA.[40-62] Most of the studies included an in-lab CPAP titration study, but 4 of the 23 studies put patients on APAP without a titration study.[43,56,59,62] In 2 of those 4 studies, the CPAP fixed pressure was based on a 1-week[62] or 2-week[59] adaptation period of APAP (rather than an in-lab titration study). Three studies were rated low risk of bias[52,54,61] and the remaining studies medium risk of bias.

Table 8. Key Question 3, Treatment: Study Characteristics

Characteristic	Weighted Mean (range) *Unless otherwise noted*	Number of studies reporting
Total number enrolled/randomized	1260 (10-200)	23
Randomized controlled trials, total n	486 (21-109)	7
Crossover studies, total n	600 (10-200)	15
Other, total n	174	1
Age of subjects (years)	53 (45-57)	23
Percent Male	87 (75-100)	22
Baseline BMI	33 (29.3-49.9)	22
Baseline ESS	11.8 (6.4-17.4)	20
Baseline AHI (events/hour)	44 (14.7-75.8)	19
Percent with hypertension (%)*	28 (20-55)	2
Location - North America, total n	119 (10-109)	2
Location - Europe, total n	1,018 (20-200)	17
Location - Australia/New Zealand, total n	77 (12-55)	3
Location - Multi-national, total n	46	1

*as defined by study

There were 22 RCTs (15 of which used a crossover design) and one retrospective cohort study.[58]. One study was conducted in the US with a VA population.[43] Of the remaining studies, one was conducted in Canada, 3 in Australia or New Zealand, and 17 in Europe.

Individual study sample sizes ranged from 10[47] to 200.[61] Among the non-crossover RCTs, treatment periods ranged from one[43] to 6[50,62] months. In the crossover studies, treatment phases ranged from 6 nights[40] to 2 months/8 weeks.[42,45,46,51,53,60] Several of the crossover studies reported either a washout period between study arms or exclusion of data collected during an initial period of use following a change to a different protocol.

Patient inclusion criteria varied in the 23 studies. All but 3 studies enrolled patients based on AHI. Minimum AHIs for enrollment were 5/hour,[46,51,58] 10/hour,[42,45,54,57,59] 15/hour,[40,41,47,49,61] 20/hour,[53,55,60] and 30/hour.[48,50,56] One study included patients with a home sleep test AHI above 10/hour or a laboratory PSG AHI above 20/hour.[44] Three studies did not specify a minimum AHI. In one, patients were referred for PSG based on the Berlin Questionnaire.[43] Another included patients based on excessive daytime sleepiness and an ESS above 9.[62] The third study enrolled patients who were already being treated with CPAP.[52] Of the studies specifying an AHI, 11 also required clinical symptoms: sleepiness in 4[44,54,59,61] and daytime symptoms in 2.[41,58]

Baseline values for AHI ranged from 15/hour[51] (a study that required participants to have AHI values between 5/hour to 30/hour) to 76/hour[40] (a study that required participants to have an AHI of 15/hour or higher with morbid obesity). The baseline value was below 20/hour in one study, between 31 and 40/hour in 3 studies, between 41 and 50/hour in 8 studies, and 51/hour or higher in 7 studies. No studies had baseline values between 21 and 30/hour; 4 did not report baseline AHI.

Baseline ESS scores ranged from 6.4[58]to 17.4[40] in 21 studies. Two studies did not report baseline ESS scores.

Clinical Outcomes – Table 9; Appendix C, Table 8-9

The most frequently reported clinical outcomes were patient preference (12 studies) and quality of life assessed with the SF-36 (9 studies). In 7 of the 12 studies reporting preference, patients preferred APAP over CPAP. The difference was reported to be statistically significant in 3 studies.[45,54,57] Four studies did not report statistical significance.[40,42,48,53] The percentage of patients expressing a preference for APAP or CPAP was similar in 5 studies with 2 reporting the statistical significance for the comparison.[47,61] No study reported a significantly higher patient preference for CPAP. In 4 studies reporting, between 10% and 72% of patients did not express a preference.[48,51,59,61]

Seven of the 9 studies reporting SF-36 quality of life found the APAP and CPAP groups to be similar post-intervention.[44,46,52,54,59,61,62] One study reported a significantly higher mental health composite score in the APAP group.[49] The remaining study did not report statistical significance for between group comparisons.[50] One study also reported groups were similar for the SAQLI.[62]

All-cause mortality was reported in 3 studies with no or few events.[43,58,61] One study reported that a minimally important difference in ESS scores was achieved in both study groups.[59] The minimally important difference was 2 points. Patient satisfaction (2 studies) was similar for the APAP and CPAP groups.[46,58] Four studies reported measures of resource utilization including hospitalization for chest pain,[43] seeking help from the sleep center,[61] extra calls or visits with sleep nurses,[62] or unplanned contacts and duration of unplanned contacts.[46] finding the groups were similar or not reporting statistical significance.

No study reported minimally important differences in urinary symptom scores, access to care, or diagnosis of cognitive impairment.

Table 9. Clinical Outcomes Comparing APAP to CPAP, Treatment Studies

Author, Year Enrolled/Randomized (n) Study Design	All-cause Mortality	Normalization of AHI	MID		Quality of Life		Patient Satisfaction		Resource Utilization		Access to Care	Diagnosis of Cognitive Impairment	Patient Preference	Other
			ESS	Urinary Symptom Score	SF-36	SAQLI	VAS	VSQ-9	Hospital i-zation	Provider Contact				
Bakker 2011[40] (n = 12) crossover													X	
Drummond 2010[43] (n = 109) RCT	X								X					
Vennelle 2010[61] (n = 200) crossover	X				↔					↔			↔	
Damjanovic 2009[41] (n = 100) RCT														
Galetke 2008[45] (n = 20) crossover													↑	
Fietze 2007[44] (n = 21) RCT					↔									
Meurice 2007[50] (n = 83) RCT					X									
Nolan 2007[51] (n = 29) crossover													X	
Patruno 2007[55] (n = 40) RCT														
Richard 2007[58] (n = 174) retro cohort	X						↔							
Nolan 2006[52] (n = 27) crossover					↔								X	
Nussbaumer 2006[54] (n = 34) crossover					↔								↑	
West 2006[62] (n = 98) RCT					↔	↔				X				
Hukins 2004[46] (n = 55) crossover					↔		↔			↔				
Hussain 2004[47] (n = 10) crossover													↔	
Marrone 2004[48] (n = 22) crossover													X	
Noseda 2004[53] (n = 27) crossover													X	
Massie 2003[49] (n = 46) crossover					↑									
Planes 2003[56] (n = 35) RCT														↔[a]
Senn 2003[59] (n = 31) crossover			↔		↔								X	

Randerath 2001[57] (n = 52) crossover													↑	
D'Ortho 2000[42] (n = 25) crossover													X	
Teschler 2000[60] (n = 10) crossover														
Total *23 studies* *(n = 1,260)*	3X	0	1↔	0	1↑ 7↔ 1X	1↔	2↔	0	1X	2↔ 1X	0	0	3↑ 2↔ 7X	1↔

[a]tolerance

↑ = significantly better with APAP than CPAP, ↔ = non-significant difference between APAP and CPAP, ↓ = significantly better with CPAP than APAP, ↕ = mixed results comparing APAP and CPAP, X = between group significance not reported

Intermediate Outcomes – Table 10; Appendix C, Tables 10-12

Intermediate outcomes were more frequently reported. Seventeen studies reported AHI after the treatment periods with 12 finding APAP and CPAP resulted in similar AHI values, one reporting significantly better results with CPAP compared to APAP, and 4 not reporting the significance of the findings. Oxygen saturation measures were reported in 11 studies with most finding APAP and CPAP to be similar.

Sleep symptom scores were also generally similar. The most frequently reported score was the ESS (reported in 20 of the 23 studies). Sixteen studies found the CPAP and APAP groups to be similar, 2 reported the scores were statistically significantly better for the APAP groups, and 2 did not report significance. For the 4 pooled parallel group RCTs, the SMD in improvement from baseline was 0.18 (95% CI -0.06, 0.43; $I^2 = 0\%$) **(**Figure 9) and the WMD was 0.85 (95% CI -0.28, 1.99)**.** We included data from the longest follow-up time reported. Findings for other sleep measures including the FOSQ, arousals, total sleep time, PSQI, and snoring were similar between the CPAP and APAP groups.

Few studies measured changes in weight, BMI, or blood pressure with none reporting a significant difference between groups.[41,51,55,56,62] One study reported statistically significant decreases in systolic and diastolic blood pressure in the CPAP group but not the APAP group but did not report the significance of the difference between study groups.[55] No study reported changes in HbA1c.

One study reported time to initiation of therapy follow diagnosis.[56] Mean time in the APAP group was 12 days compared to 27 days in the CPAP group (P <.01). The difference was attributed to wait time for PSG.

Adverse events were reported in 6 studies. Three found similar frequency of events (dry mouth, blocked/runny nose, skin irritation, nasal irritation, pressure feeling too high, claustrophobia) in the CPAP and APAP groups.[51,57,59] One found incidence of nasal and throat/mouth symptoms was similar for different APAP devices but did not compare findings to CPAP.[52] One study reported mixed results (*ie,* no difference in nasal irritation/obstruction, pressure intolerance, or partner dislike but fewer total side effects [P = .02] in the APAP group).[46] One reported no side effects requiring alteration of treatment.[54] In this study, the incidence of nasal and mouth/throat

side effects was similar for the CPAP and APAP groups while discomfort with air pressure was significantly higher in the CPAP group.

One study reported costs during a 2-month treatment period.[56] Hospital costs were significantly lower in the APAP group while telecommunication costs (to transmit data from home to sleep laboratory) were significantly higher (both P <.001). Costs for equipment and home nurse visits were similar for the APAP and CPAP groups. Total cost per patient of the 2-month treatment was lower in the APAP group (€1,264 vs €1,720; P <.01).

Table 10. Intermediate Outcomes Comparing APAP to CPAP, Treatment Studies

Author, Year Enrolled/Randomized (n) Study Design	AHI	Oxygen Saturation				Sleep Symptom Scores							Weight Loss	BMI	Blood Pressure	HbA1c	Time to Initiation of Therapy	Adverse Events	Costs	Compliance/ Adherence			
		Mean	Time <90%	ODI	Minimum	ESS	FOSQ	SASQ	Arousals	Total Sleep Time	PSQI	Snoring	Other								Hrs/night	Proportion of	Regular Use	Other
Bakker 2011[40] (n = 12) crossover	↔	↓	↔	↔		↔				↔											↔			
Drummond 2010[43] (n = 109) RCT						↔	↔																	
Vennelle 2010[61] (n = 200) crossover	↔					↑															↑			
Damjanovic 2009[41] (n = 100) RCT	↔			↔		↔			↔						X						↔	↔		↔[a]
Galetke 2008[45] (n = 20) crossover	↔				↔	↔			↔			↔									↔			
Fietze 2007[44] (n = 21) RCT	↔					↔															↔	↔		
Meurice 2007[50] (n = 83) RCT	X	↔	↔			↔				X											↔			
Nolan 2007[51] (n = 29) crossover	↔	↔				↔			↑	↔		↔		X		X			X		↔	↔		
Patruno 2007[55] (n = 40) RCT	↓	↔		↓										X	X						↔			
Richard 2007[58] (n = 174) retro cohort																					↔	↔	X	
Nolan 2006[52] (n = 27) crossover						↔												X			↕	↕		
Nussbaumer 2006[54] (n = 34) crossover	↔			↔		↔												↕			↔		↔	
West 2006[62] (n = 98) RCT	↔					↔							↔[b]		↔						↔	↔		
Hukins 2004[46] (n = 55) crossover						↔												↑			↔	↔		
Hussain 2004[47] (n = 10) crossover	↔			↔	↓	↔			↔	↔											↔			

Marrone 2004[48] (n = 22) crossover						↔														↔	↔		
Noseda 2004[53] (n = 27) crossover						↑			↔		↔									↔	↔		
Massie 2003[49] (n = 46) crossover	↔					↔														↑	↔		
Planes 2003[56] (n = 35) RCT	↔		↔			↔								X			↑		↑	↔		↔	
Senn 2003[59] (n = 31) crossover	↔		↔			↔												↔		↔		↔	
Randerath 2001[57] (n = 52) crossover	X					X			X	X		X						↔		X	X		
D'Ortho 2000[42] (n = 25) crossover	X	X				X														↔		X	
Teschler 2000[60] (n = 10) crossover	X																			↔	↔		
Total *23 studies* *(n = 1,260)*	12 ↔ 1↓ 4X	3↔ 1↓ 1X	4↔	4↔ 1↓	1↓ 1←	2↑ 16↔ 2X	1↔	0	1↑ 4↔ 1X	3↔ 2X	1↔	2↔ 1X	1↔	1X	3X	1↔ 2X	0	1↑	1↑ 2↔ 1↕ 2X	1↑				

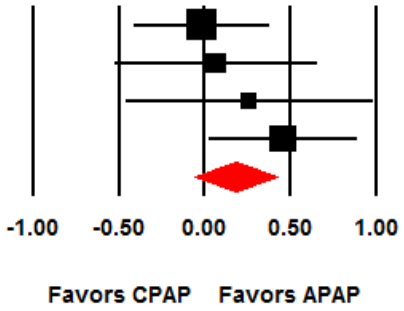
[a] hours use/sleep time
[b] maintenance of wakefulness
↑ = significantly better with APAP than CPAP, ↔ = non-significant difference between APAP and CPAP, ↓ = significantly better with CPAP than APAP, ↕ = mixed results comparing APAP and CPAP, X = between group significance not reported

Twenty-two of the 23 included studies reported hours per night of use with 19 of 22 reporting the CPAP and APAP groups were similar. Of the remaining 3 studies, 2 noted significantly higher use with APAP[49,61] and one reported mixed results for different APAP devices.[52] The differences in means for the 5 non-crossover RCTs are presented in Figure 10. Using data from the longest reported follow-up time, the pooled difference was -0.08 (95% CI -0.55, 0.38; I^2 = 14%).

The results for proportion of nights used were similar; 10 of 12 studies found the CPAP and APAP groups to be similar, one reported mixed results,[52] and one did not report statistical significance.[57] "Regular use," as defined by the study, was similar in 3 of the 5 studies reporting (with 2 not reporting significance).

Figure 9. Epworth Sleepiness Scores, Standardized Mean Difference for Mean Change from Baseline from Parallel Group RCTs, Treatment Studies[a]

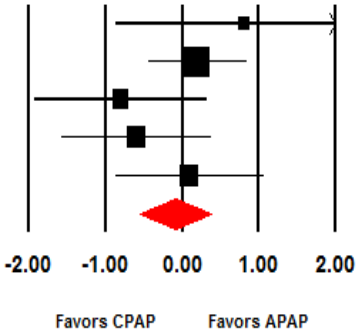
Study name	Std diff in means	Lower limit	Upper limit	Sample size APAP	Sample size CPAP
Damjanovic 2009 (9 months) (41)	-0.02	-0.41	0.37	50	50
Meurice 2007 (6 months) (50)	0.07	-0.53	0.66	51	14
Planes 2003 (2 months) (56)	0.26	-0.46	0.98	16	14
Drummond 2010 (final, 2-6 months) (43)	0.46	0.03	0.88	42	44
	0.18	-0.06	0.43		



[a]Lower and upper limits represent 95% confidence intervals

Figure 10. Compliance, Weighted Mean Difference for Hours per Night from Parallel Group RCTs, Treatment Studies[a]

Study name	Difference in means	Lower limit	Upper limit	Sample size APAP	Sample size CPAP
Fietze 2007 (6 weeks) (44)	0.80	-0.86	2.46	10	11
Patruno 2007 (3 months) (55)	0.20	-0.44	0.84	15	16
Planes 2003 (2 months) (56)	-0.80	-1.92	0.32	16	14
Meurice 2007 (6 months) (50)	-0.60	-1.57	0.37	51	14
Damjanovic 2009 (9 months) (41)	0.10	-0.87	1.07	50	50
	-0.08	-0.55	0.38		



[a]Lower and upper limits represent 95% confidence intervals

KEY QUESTION #3A: Do effectiveness/harms/resource utilization vary by patient characteristics:

a. Unexplained daytime sleepiness/fatigue

b. AHI severity

c. Other risk factors or coexisting conditions associated with OSA (*eg*, obesity, neck circumference, treatment-resistant hypertension, diabetes, age, sex, race)

d. Symptoms (*eg*, nocturia, loss of libido, snoring, witnessed sleep "apnea")?

Few studies reported results by patient characteristics.

A. Unexplained Daytime Sleepiness/Fatigue

No study reported results for subgroups of patients based on unexplained daytime sleepiness or fatigue.

B. AHI Severity

One treatment study grouped patients by baseline AHI (< 60/hour or ≥ 60/hour).[42] Duration of use (hours/night) of CPAP and APAP was similar in the overall study population. There was significantly longer use of both CPAP and APAP in patients with baseline AHI ≥ 60/hour compared to those with baseline AHI <60/hour. A second study compared baseline AHI in compliant and non-compliant patients (with compliance defined as at least 4 hours per night for at least 5 days per week). AHI levels did not differ (51/hour for the compliant group, 47/hour for the non-compliant group, P = .40); results were not reported by treatment group.[58] A third study reported that compliance was similar between CPAP and APAP treatment groups in post-hoc subgroup analyses of patients with differing degrees of OSA severity.[46]

Two treatment studies compared baseline AHI in patients who expressed a preference for CPAP or APAP. One found that baseline AHI was a significant predictor of preference for CPAP or APAP.[48] Patients who preferred APAP had a higher baseline AHI than patients who preferred CPAP (73.1 vs 60.0/hour, P<.02). The other found AHI values were similar whether patients preferred APAP or CPAP (16.3 vs 14.2/hour, P = .49).[51]

C. Other Risk Factors or Coexisting Conditions Associated with OSA

One treatment study included only morbidly obese (BMI ≥ 40 kg/m^2) patients.[40] Neck circumferences was 46.5 (4.0) cm. CPAP and APAP generally produced similar results; patients were not stratified by BMI.

Two treatment studies evaluated age and compliance. In both studies, older patients tended to be more compliant with therapy but the findings were not statistically significant.[45,58] One study also reported that baseline BMI and ESS were not associated with compliance.[58] Neither study reported results by treatment group (CPAP or APAP).

Two studies reported other factors associated with preference for CPAP or APAP. One found that neither age or ESS score at baseline were significant predictors of preference.[48] Another reported that neither age, BMI, neck circumference, or ESS score were significant predictors.[51]

D. Symptoms

One titration study reported baseline percentages of patients with habitual snoring (88%), observed apneas (61%), and nocturia (27%).[37] The study did not report results for subgroups of patients based on these characteristics.

Strength of Evidence for Key Question 3 (Table 4, Appendix D)

Titration

For studies comparing in-lab CPAP *titration* to at-home APAP *titration*, there was insufficient evidence for the outcomes of access to care and adverse events. Strength of evidence was moderate for quality of life and low for compliance (hours/night) and ESS with similar findings for the 2 study groups.

Treatment

Among studies comparing CPAP to APAP for *treatment* of OSA, there was insufficient evidence for access to care. Strength of evidence was moderate for quality of life, ESS, and compliance (hours/night) and low for adverse events. These outcomes were similar in the 2 treatment groups.

Table 11. Strength of Evidence for Outcomes, Key Question 3

Comparison	Outcome of Interest	Strength of Evidence	Direction	Findings
KQ3: Home APAP technology versus standard in-center manual CPAP *titration*	Access to care	Insufficient		We found no evidence for this outcome.
	Epworth Sleepiness Scale (ESS)	Moderate	Similar	Based on 2 RCTs (n = 414) with moderate risk of bias, we found improvement from baseline in ESS scores was similar for patients allocated to home APAP titration compared to patients allocated to in-center CPAP titration (SMD = 0.0 [95% CI -0.22, 0.21]). One moderate risk of bias RCT (n = 68) found median change in EES scores from baseline was also similar between groups (MD 1 [95% CI -1, 4]). One observational cohort study also found ESS scores were similar between groups.
	Quality of life	Moderate	Similar	Based on 2 RCTs (n = 414) with moderate risk of bias, we found quality of life measures were similar for patients allocated to home APAP titration compared to patients allocated to in-center CPAP titration. The SMDs for SF-36 Mental Health and Physical Health scores were 0.08 [95% CI -0.14, 0.29] and -0.21 [95% CI -0.61, 0.20], respectively. Results for the Physical Health scores were imprecise. One moderate risk of bias RCT (n = 68) found median improvement from baseline in the SAQLI was similar between groups (median difference = 0.17 [95% CI -0.6, 0.9])
	Compliance, hours per night	Low	Similar	Based on 2 RCTs (n = 414) with moderate risk of bias, we found compliance was similar for patients allocated to home APAP titration compared to patients allocated to in-center CPAP titration (WMD = 0.02 [95% CI -0.41, 0.45]). One moderate risk of bias RCT (n = 68) found median compliance was better in the APAP group versus the CPAP group (MD -1.1 [95% CI -2.0, -0.2]). One observational cohort study found compliance was similar between groups.
	Adverse events	Insufficient		Based on the findings of one RCT (n = 245) that reported no "important differences" in adverse events between the home APAP and in-lab CPAP and groups, the evidence is insufficient to draw conclusions.
KQ3: APAP versus CPAP *treatment*	Access to care	Insufficient		We found no evidence for this outcome.
	Epworth Sleepiness Scale (ESS)	Moderate	Similar	Based on 4 parallel group RCTs (n = 327) with aggregate moderate risk of bias, we found improvement from baseline in ESS scores was similar for patients allocated to APAP treatment compared to patients allocated to CPAP treatment (SMD = 0.18 [95% CI -0.06, 0.43]). Two parallel group trials not pooled (reported as a median or data not shown) also found improvement from baseline in ESS scores similar between groups. Ten crossover RCTs (n = 269) reported similar improvements between groups and 2 (N = 227) reported greater improvement with APAP.

Comparison	Outcome of Interest	Strength of Evidence	Direction	Findings
	Quality of life	Moderate	Similar	Based on 3 parallel group RCTs (n = 202) with aggregate moderate risk of bias, we found quality of life measures (SF-36, SAQLI) were similar for patients allocated to APAP treatment compared to patients allocated to CPAP treatment (data were not pooled due to variation in reporting of results, *ie,* reported as medians). Six crossover RCTs (n = 393) also reported no differences in most of the quality of life measures between the treatment groups.
	Compliance, hours per night	Moderate	Similar	Based on 5 parallel group RCTs (n = 279) with aggregate moderate risk of bias, we found compliance was similar for patients allocated to APAP treatment compared to patients allocated to CPAP treatment (WMD = -0.08 [95% CI -0.55, 0.38]). One parallel group RCT reporting median compliance, most of the remaining crossover RCTs, and one observational study also found compliance was similar between groups.
	Adverse events	Low	Similar	Adverse events were infrequently reported. One parallel group RCT (n = 109) reported adverse events, chest pain in 12% and 9% of APAP and CPAP patients, respectively. Five crossover trials reported adverse events for both APAP and CPAP treatments. One trial (n = 55) reported a higher frequency of total number of events and another trial (n = 34) reported a higher incidence of pressure discomfort with CPAP therapy arm compared with the APAP treatment. There trials (n = 112) reported no differences in adverse events between the treatment groups.

APAP = Auto-adjusted (autoregulated) continuous positive airway pressure; CPAP = continuous positive airway pressure; RCT = randomized controlled trial; SAQLI = Sleep Apnea Quality of Life Index; SF-36 = Short Form-36; SMD = standardized mean difference; WMD = weighted mean difference.

[a]Strength of Evidence Definitions;[28] See Appendix D for more details

• High: Very confident that estimate of effect lies close to true effect. Few or no deficiencies in body of evidence, findings believed to be stable.

• Moderate: Moderately confident that estimate of effect lies close to true effect. Some deficiencies in body of evidence; findings likely to be stable, but some doubt.

• Low: Limited confidence that estimate of effect lies close to true effect; major or numerous deficiencies in body of evidence. Additional evidence necessary before concluding that findings are stable or that estimate of effect is close to true effect.

• Insufficient: No evidence, unable to estimate an effect, or no confidence in estimate of effect. No evidence is available or the body of evidence precludes judgment.

SUMMARY AND DISCUSSION

Our systematic review compared outcomes associated with 3 key questions: 1) use of non-SSP providers compared to SSPs; 2) electronic consultation compared to interactive consultation; and 3) in-home APAP compared to in-laboratory CPAP. Overall the evidence was often lacking to fully address the key questions and we found no data for Key Question 2.

For Key Question 1, we found 8 studies, of which 4 were randomized trials. Access to care was not reported in any of the studies, so we were unable to determine if use of non-SSP providers improves access to OSA diagnosis and treatment. Likewise, adverse events including overdiagnosis/overtreatment and underdiagnosis/undertreatment were not reported. We found moderate-strength evidence that care provided by non-SSP providers and SSPs resulted in similar Epworth Sleepiness Scores, quality of life, and treatment compliance.

For Key Question 3, we found 4 studies comparing at-home APAP to in-laboratory CPAP for *titration* of CPAP pressures and 23 for *treatment*. For both *titration* and *treatment*, access to care was not reported in any of the studies. Adverse events were rarely reported and therefore strength of evidence was insufficient to low to determine whether at-home APAP and in-laboratory CPAP differ for adverse events. We found moderate-strength evidence in both *titration* and *treatment* studies that at-home APAP and in-lab CPAP were similar in regards to Epworth Sleepiness Scores and quality of life outcomes. For PAP *titration*, at-home APAP and in-laboratory CPAP resulted in similar treatment compliance, though strength of evidence was low. For PAP *treatment*, at-home APAP and in-laboratory CPAP resulted in similar treatment compliance (moderate strength of evidence).

Our systematic review was not intended to determine the best OSA screening tools nor in whom OSA treatments are most effective. However, most of the included studies in our report enrolled patients who were generally older, obese, and sleepy, as determined by Epworth Sleepiness Scores. Furthermore, 2 recent evidence reports and clinical practice recommendations have addressed OSA screening and treatment.[12-15] This additional evidence may assist clinicians and policymakers when determining referral recommendations and evaluation pathways to target individuals with the greatest likelihood of benefiting while minimizing harms, as well as financial and opportunity costs associated with diagnosis and treatment of OSA. Specifically, a prior evidence report and accompanying clinical practice guideline by the ACP recommends that clinicians target their assessment of OSA to individuals with unexplained daytime sleepiness. The ACP concluded that assessment of OSA in the absence of daytime sleepiness or treatment of persons with low AHI scores is low-value care because evidence to date indicates that neither improves clinical outcomes.[14] Additionally, a draft evidence report and recommendation statement from the U.S. Preventive Services Task Force indicates that evidence is insufficient to assess the net benefit of screening for and treatment of asymptomatic OSA.[15] While good evidence has established that people with severe OSA are at increased mortality risk compared to controls, trials of CPAP and other treatments have not established whether treatment reduces mortality or improves most other health outcomes, except possibly for sleep-related quality of life. In addition to the findings included in the ACP and USPSTF reports, a recently published multi-site randomized trial showed that CPAP was not effective for secondary prevention of cardiovascular events in those with established cardiovascular disease and moderate to severe

OSA.[16] Furthermore, the USPSTF evidence report also found uncertainty about the clinical utility of screening tools and that most screening questionnaires had poor diagnostic accuracy.

When redesigning OSA care models we recommend that clinicians and policymakers consider the limited data to support some of these newer models (*eg,* hiring non-SSPs to provide sleep care) and the existing evidence used to inform clinical practice recommendations by the ACP and USPSTF. While the existing data do not suggest significant harm from these newer care models, current data are insufficient to draw firm conclusions.

LIMITATIONS

Inherent to the process of systematic review there are certain general limitations, some related to our inclusion criteria, others inherent in the available research. For example, we limited inclusion to English language studies and those performed in the United States, Canada, Western Europe, Australia, or New Zealand. We did this to target research most likely to be applicable to clinical practice in the United States and Veterans Affairs facilities. However, these search criteria also mean that our findings may not apply to the evaluation and care of OSA patients in other healthcare settings.

We also chose to exclude studies of dentists and anesthesiologists due to our uncertainty whether or not these persons truly represent "Non-Sleep-Specialist Physicians" given that many dentists and anesthesiologists have substantial practices or background in sleep medicine. Therefore, our systematic review does not address the role of these practitioners on sleep apnea care. We also did not assess the role of surgery or mandibular assist devices for OSA treatment including referral of patients who may be more interested in or better candidates for these options.

A significant limitation is the paucity of high-quality literature regarding key questions 1 and 2. For initial evaluation and treatment of OSA (Key Question 1), we identified 4 randomized trials and 4 observational studies. The non-SSP providers were primary care providers in 4 studies, nurses in 3 studies, and a pulmonologist in one study. Therefore, although we are aware of the existence of clinical care models utilizing sleep respiratory therapists to provide varying degrees of OSA care, we did not find any studies to address this particular type of practitioner. Due to small sample sizes, we were also unable to directly compare primary care-based models of OSA evaluation and treatment to sleep nurse-based models.

We also note that, importantly, the providers in many of these "primary care" studies were persons who had substantial experience in sleep medicine. For example, in the randomized trial of Chai-Coetzer and colleagues,[31] one of the 4 primary care nurses had 15 years of experience at a tertiary care sleep medicine center. In the study by Chamorro and colleagues,[32] the non-SSP provider was a 'primary care pulmonologist' – the degree of sleep training and experience of this person was unclear. Therefore, the generalizability of these findings to primary care providers with less experience in sleep medicine is not clear.

Due to the lack of studies regarding electronic consultation for persons suspected of having OSA (Key Question 2), we are unable to determine the effectiveness/harms/resource utilization of this evolving practice.

We found the most data for Key Question 3, which compared at-home autotitrating continuous positive airway pressure (APAP) to the more traditional CPAP titrated in the PSG laboratory. Several studies that compared at-home APAP to in-lab CPAP[63-65] were excluded as these studies also used different methods of diagnostic testing (*eg*, HST preceded at-home APAP, while PSG preceded in-lab CPAP) and we were concerned that the differences in diagnostic testing might confound our analysis of outcomes between APAP and CPAP. Although these studies were not included in our formal analysis, they reassuringly also further support the notion that APAP and CPAP (combined with at-home vs in-laboratory testing, respectively) result in similar outcomes in regards to: 1) 6-week ESS and PAP adherence among 106 Veterans in Florida using 2-3 nights of APAP for *titration*,[63] 3-month ESS, SF-12, and adherence among 182 Veterans in Pennsylvania using 4-5 nights of APAP for *titration*,[64] and 3-month ESS and SF-36 among 142 patients at 7 academic sleep centers in the U.S. using 5-7 nights of APAP for *titration*;[65] this last study also found higher 3-month PAP adherence in those assigned to the home study plus APAP titration arm.[65]

APPLICABILITY AND IMPLEMENTATION OF FINDINGS

Most patients enrolled in studies were obese, middle-aged men, with severe OSA based on both high AHI levels and the presence of excessive daytime sleepiness. Our findings are most applicable to these individuals. We found only one study that was performed in a VA population for Key Question 3.[43] One study was performed at Walter Reed Army Medical Center for key questions 1 and 3.[33] While most studies were not specifically conducted in VA or military populations, because the patients enrolled in these studies were generally older, overweight men with OSA, we believe the findings of our systematic review should be applicable to the population of Veterans served by VA facilities.

Many of the study providers who were not SSPs had prior sleep training and therefore the results may not be fully generalizable to all primary care providers.While data are not conclusive, because our findings indicated similar Epworth Sleepiness Scores, quality of life, and treatment compliance scores among patients evaluated and treated by non-SSPs compared to SSPs, it may be reasonable to consider expanded use of non-SSP providers who have received training in sleep medicine, especially where SSPs are in limited supply and demand for OSA services is high. Similarly, greater use of at-home APAP may lessen dependence on backlogged PSG laboratories, as most health outcomes were similar between groups.

Our report focused on methods that might improve the ‘supply’ side of OSA evaluation and treatment, through use of non-SSPs, electronic consultation, and at-home APAP titration and treatment. However, healthcare systems struggling to match supply to demand might also consider whether the ‘demand’ is truly appropriate. We found little to no data in screen-detected patients (*ie*, those found to have abnormal AHI either through direct referral to sleep laboratories or based on results of screening questionnaires such as the Berlin questionnaire but without excessive daytime sleepiness). The evidence to date indicates that the main benefit of OSA detection and treatment is improvement in patient-reported sleepiness symptoms among those with unexplained daytime somnolence. Therefore, VA healthcare providers and decision-makers could potentially achieve the highest value care, including resource use, by targeting case finding approaches and subsequent evaluation and treatment to individuals with unexplained daytime somnolence and who express interest in further evaluation and treatment. Theorectically, this referral approach could be readily be implemented by developing and using electronic medical

record templates that describe the evidence-based rationale for the referral recommendations while requesting that referring providers include information specifically about daytime somnolence and the patient's willingness for further evaluation and treatment in the consult.

RESEARCH GAPS/FUTURE RESEARCH

Comparative effectiveness trials were lacking for all key questions. Key questions 1 and 2 would particularly benefit from trials to address the outcomes resulting from non-SSP case finding and care of sleep apnea patients (Key Question 1) and outcomes resulting from electronic consultation (Key Question 2). Limited available data suggest that care led by non-SSPs may potentially provide equivalent outcomes to care led by SSPs. Comparative effectiveness trials are needed in order to determine whether such results can be achieved in routine practice, outside of controlled research settings. The available data suggest that with appropriate training, non-SSPs can potentially provide equivalent outcomes, but the operationalization of such training is unclear. Therefore, future comparative effectiveness trials should describe their training programs in detail. Such trials should also collect clinical outcomes where possible.

In regards to case finding, future studies should compare outcomes among differing strategies such as offering sleep studies and treatment to a broad set of patients (*eg,* with OSA risk factors, regardless of number of risk factors or symptoms) versus a narrow set of criteria (*eg,* high OSA risk or very symptomatic). Outcomes of interest in such studies would include consultation rates, sleep study rates, percent of sleep studies confirming OSA, severity of OSA, changes in sleepiness, and adherence to CPAP. Such studies might also compare case finding led by sleep specialists, primary care providers, nurses, or even automated through electronic medical record (EMR) systems.

As more healthcare systems implement comprehensive EMRs, we anticipate Key Question 2 will become more feasible to study. In the current climate of increasing numbers of sleep referrals, yet a declining number of SSPs, EMR-based electronic consultation holds significant promise to provide equivalent outcomes in a more cost-effective, time-efficient manner. Although many systems have already implemented electronic consultation systems,[66] evidence supporting this practice is largely lacking and further studies should be conducted. We think comparative effectiveness study designs such as stepped-wedge randomization (where sites are randomly assigned to the time point at which they implement electronic consultation) would allow the creation of good-quality evidence to quantify the risks, benefits, and economic impacts of electronic consultation for patients with known or suspected OSA.

Many studies have addressed Key Question 3 and while we found similar sleepiness, quality of life, and adherence between APAP and CPAP, future studies should include longer-term follow-up to allow better determination of long-term clinical ouctomes. Future studies should also include more rigorous collection and reporting of adverse event data.

A large gap in evidence is related to the effectiveness of treatment for individuals without excessive daytime sleepiness (screen detected or case-finding in at-risk asymptomatic individuals) and on outcomes other than daytime sleepiness. For example, a recently published multi-site randomized trial showed that CPAP was not effective for secondary prevention of cardiovascular events in those with established cardiovascular disease and moderate to severe OSA.[16] Additional information gaps include the effectiveness of treatment in those with mild

AHI regardless of symptoms. These gaps are supported by recent evidence reports and accompanying clinical practice recommendations by the ACP and USPSTF and described in greater detail in the introduction and above. There are also critically important gaps to fill due to the dramatic increase in patients with, or suspected to have, OSA and thus being referred for evaluation and treatment.

CONCLUSIONS

Among patients suspected of having OSA, evidence suggests that primary care providers and sleep-specialist nurses might provide similar outcomes to SSPs, although the strength of this evidence was only moderate and many outcomes were inconsistently reported. Likewise, among patients diagnosed with OSA, evidence suggests that at-home APAP *titration* and *treatment* provides similar outcomes to fixed pressure CPAP titrated in the PSG laboratory, although the strength of evidence was generally low to moderate.

We found no evidence addressing the topic of electronic consultation for the management of known or suspected OSA.

Future studies are needed to determine which patients derive the most benefit from treatment and should be prioritized for testing and treatment, whether newer models of care with less reliance on SSP time (either through utilization of other types of providers or electronic consultation) result in similar outcomes to traditional models, and if effective, how such models should be implemented.

REFERENCES

1. Marin JM, Carrizo SJ, Vicente E, Agusti AG. Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: an observational study. *Lancet.* 2005;365(9464):1046-1053.
2. Peker Y, Hedner J, Norum J, Kraiczi H, Carlson J. Increased incidence of cardiovascular disease in middle-aged men with obstructive sleep apnea: a 7-year follow-up. *Am J Respir Crit Care Med.* 2002;166(2):159-165.
3. Gottlieb DJ, Yenokyan G, Newman AB, et al. Prospective study of obstructive sleep apnea and incident coronary heart disease and heart failure: the sleep heart health study. *Circulation.* 2010;122(4):352-360.
4. Redline S, Yenokyan G, Gottlieb DJ, et al. Obstructive sleep apnea-hypopnea and incident stroke: the sleep heart health study. *Am J Respir Crit Care Med.* 2010;182(2):269-277.
5. Yaggi HK, Concato J, Kernan WN, Lichtman JH, Brass LM, Mohsenin V. Obstructive sleep apnea as a risk factor for stroke and death. *N Engl J Med.* 2005;353(19):2034-2041.
6. Blackwell T, Yaffe K, Laffan A, et al. Associations of objectively and subjectively measured sleep quality with subsequent cognitive decline in older community-dwelling men: the MrOS sleep study. *Sleep.* 2014;37(4):655-663.
7. Yaffe K, Laffan AM, Harrison SL, et al. Sleep-disordered breathing, hypoxia, and risk of mild cognitive impairment and dementia in older women. *JAMA.* 2011;306(6):613-619.
8. Giles TL, Lasserson TJ, Smith BH, White J, Wright J, Cates CJ. Continuous positive airways pressure for obstructive sleep apnoea in adults. *Cochrane Database Syst Rev.* 2006;(3):CD001106.
9. Tregear S, Reston J, Schoelles K, Phillips B. Continuous positive airway pressure reduces risk of motor vehicle crash among drivers with obstructive sleep apnea: systematic review and meta-analysis. *Sleep.* Oct 2010;33(10):1373-1380.
10. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM. Increased prvalence of sleep-diordered breathing in adults. *Am J Epidemol.* 2013;177(9):1006-1014.
11. Frost & Sullivan. *Hidden Health Crisis Costing America Billions: Underdiagnosing and Undertreating Obstructive Sleep Apnea Draining Healthcare System.* American Academy of Sleep Medicine, 2016, http://www.aasmnet.org/Resources/pdf/sleep-apnea-economic-crisis.pdf. Accessed 22 Sept 2016
12. Balk EM, Moorthy D, Obadan NO, et al. *Diagnosis and Treatment of Obstructive Sleep Apnea in Adults. Comparative Effectiveness Review No. 32.* Rockville, MD: AHRQ Publication No. 11-EHC052-EF, 2011.
13. Qaseem A, Dallas P, Owens DK, et al. Diagnosis of obstructive sleep apnea in adults: a Clinical Practice Guideline from the American College of Physicians. *Ann Intern Med.* 2014;161:210-220.

14. Qaseem A, Holty J-E, Owens DK, Dallas P, Starkey M, Shekelle P. Management of obstructive sleep apnea in adults: a Clinical Practice Guideline from the American College of Physicians. *Ann Intern Med.* 2013;159(7):471-483.

15. U.S. Preventive Services Task Force. *Draft Recommendation Statement: Obstructive Sleep Apnea in Adults: Screening.* 2016. Available from: https://www.uspreventiveservicestaskforce.org/Page/Document/draft-recommendation-statement166/obstructive-sleep-apnea-in-adults-screening. Accessed 21 Sept 2016

16. McEvoy RD, Antic NA, Heeley E, et al. CPAP for prevention of cardiovascular events in obstructive sleep apnea. *New Engl J Med.* 2016;375(10):919-931.

17. Hiestand DM, Britz P, Goldman M, Phillips B. Prevalence of symptoms and risk of sleep apnea in the US population: Results from the national sleep foundation sleep in America 2005 poll. *Chest.* Sep 2006;130(3):780-786.

18. Das SR, Kinsinger LS, Yancy WS, Jr., et al. Obesity prevalence among veterans at Veterans Affairs medical facilities. *Am J Prev Med.* Apr 2005;28(3):291-294.

19. Alexander M, Ray MA, Hebert JR, et al. The National Veteran Sleep Disorder Study: descriptive epidemiology and secular trends, 2000-2010. *Sleep.* 2016;39(7);1399-1410.

20. NCD Risk Factor Collaboration. Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *Lancet.* 2016;387:1377-1396.

21. Phillips B, Gozal D, Malhotra A. What is the future of sleep medicine in the United States? *Am J Respir Crit Care Med.* 2015;192(8):915-917.

22. Collop NA, Anderson WM, Boehlecke B, et al. Clinical guidelines for the use of unattended portable monitors in the diagnosis of obstructive sleep apnea in adult patients. Portable Monitoring Task Force of the American Academy of Sleep Medicine. *J Clin Sleep Med.* 2007;3(7):737-747.

23. Morgenthaler TI, Aurora RN, Brown T, et al. Practice parameters for the use of autotitrating continuous positive airway pressure devices for titrating pressures and treating adult patients with obstructive sleep apnea syndrome: an update for 2007. An American Academy of Sleep Medicine report. *Sleep.* Jan 2008;31(1):141-147.

24. Higgins H, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. *The Cochrane Collaboration*, 2011.

25. Viswanathan M, Ansarai MT, Berkman ND, et al. *Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions.* Rockville, MD: Agency for Healthcare Research and Quality, 2008.

26. Cohen J. *Statistical power analysis for the behavioral sciences*, 2nd ed. Hillsdale, NJ: Lawrence Earlbaum Associates; 1988.

27. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327(7414):557-560.

28. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions--agency for healthcare research and quality and the effective health-care program. *J Clin Epidemiol.* 2010;63(5):513-523.

29. Andreu AL, Chiner E, Sancho-Chust JN, et al. Effect of an ambulatory diagnostic and treatment programme in patients with sleep apnoea. *Eur Respir J.* 2012;39(2):305-312.

30. Antic NA, Buchan C, Esterman A, et al. A randomized controlled trial of nurse-led care for symptomatic moderate-severe obstructive sleep apnea. *Am J Respir Crit Care Med.* 2009;179(6):501-508.

31. Chai-Coetzer CL, Antic NA, Rowland LS, et al. Primary care vs specialist sleep center management of obstructive sleep apnea and daytime sleepiness and quality of life: a randomized trial. *JAMA.* 2013;309(10):997-1004.

32. Chamorro N, Sellares J, Millan G, et al. An integrated model involving sleep units and primary care for the diagnosis of sleep apnoea. *Eur Respir J.* 2013;42(4):1151-1154.

33. Lettieri CF, Lettieri CJ, Carter K. Does home sleep testing impair continuous positive airway pressure adherence in patients with obstructive sleep apnea? *Chest.* 2011;139(4):849-854.

34. Palmer S, Selvaraj S, Dunn C, et al. Annual review of patients with sleep apnea/hypopnea syndrome: a pragmatic randomised trial of nurse home visit versus consultant clinic review. *Sleep Med.* 2004;5(1):61-65.

35. Pamidi S, Knutson KL, Ghods F, Mokhlesi B. The impact of sleep consultation prior to a diagnostic polysomnogram on continuous positive airway pressure adherence. *Chest.* 2012;141(1):51-57.

36. Scharf SM, DeMore J, Landau T, Smale P. Comparison of primary-care practitioners and sleep specialists in the treatment of obstructive sleep apnea. *Sleep Breath.* 2004;8(3):111-124.

37. Masa JF, Jimenez A, Duran J, et al. Alternative methods of titrating continuous positive airway pressure: a large multicenter study. *Am J Respir Crit Care Med.* 2004;170(11):1218-1224.

38. McArdle N, Singh B, Murphy M, et al. Continuous positive airway pressure titration for obstructive sleep apnoea: automatic versus manual titration. *Thorax.* 2010;65(7):606-611.

39. Mulgrew AT, Fox N, Ayas NT, Ryan CF. Diagnosis and initial management of obstructive sleep apnea without polysomnography: a randomized validation study. *Ann Intern Med.* 2007;146(3):157-166.

40. 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;20(1 Pt 2):233-240.

41. Damjanovic D, Fluck A, Bremer H, Muller-Quernheim J, Idzko M, Sorichter S. Compliance in sleep apnoea therapy: influence of home care support and pressure mode. *Eur Respir J.* 2009;33(4):804-811.

42. d'Ortho M, Grillier-Lanoir V, Levy P, et al. Constant vs automatic continuous positive airway pressure therapy: home evaluation. *Chest.* 2000;118(4):1010-1017.

43. Drummond F, Doelken P, Ahmed QA, et al. Empiric auto-titrating CPAP in people with suspected obstructive sleep apnea. *J Clin Sleep Med.* 2010;6(2):140-145.

44. Fietze I, Glos M, Moebus I, Witt C, Penzel T, Baumann G. Automatic pressure titration with APAP is as effective as manual titration with CPAP in patients with obstructive sleep apnea. *Respiration.* 2007;74(3):279-286.

45. Galetke W, Anduleit N, Richter K, Stieglitz S, Randerath WJ. Comparison of automatic and continuous positive airway pressure in a night-by-night analysis: a randomized, crossover study. *Respiration.* 2008;75(2):163-169.

46. Hukins C. Comparative study of autotitrating and fixed-pressure CPAP in the home: a randomized, single-blind crossover trial. *Sleep.* 2004;27(8):1512-1517.

47. Hussain SF, Love L, Burt H, Fleetham JA. A randomized trial of auto-titrating CPAP and fixed CPAP in the treatment of obstructive sleep apnea-hypopnea. *Respir Med* 2004;98(4):330-333.

48. Marrone O, Resta O, Salvaggio A, Giliberti T, Stefano A, Insalaco G. Preference for fixed or automatic CPAP in patients with obstructive sleep apnea syndrome. *Sleep Med.* 2004;5(3):247-251.

49. Massie CA, McArdle N, Hart RW, et al. Comparison between automatic and fixed positive airway pressure therapy in the home. *Am J Respir Crit Care Med.* 2003;167(1):20-23.

50. Meurice JC, Cornette A, Philip-Joet F, et al. Evaluation of autoCPAP devices in home treatment of sleep apnea/hypopnea syndrome. *Sleep Med.* 2007;8(7-8):695-703.

51. Nolan GM, Doherty LS, Mc Nicholas WT. Auto-adjusting versus fixed positive pressure therapy in mild to moderate obstructive sleep apnoea. *Sleep.* Feb 2007;30(2):189-194.

52. Nolan GM, Ryan S, O'Connor T M, McNicholas WT. Comparison of three auto-adjusting positive pressure devices in patients with sleep apnoea. *Eur Respir J.* 2006;28(1):159-164.

53. Noseda A, Kempenaers C, Kerkhofs M, Braun S, Linkowski P, Jann E. Constant vs auto-continuous positive airway pressure in patients with sleep apnea hypopnea syndrome and a high variability in pressure requirement. *Chest.* 2004;126(1):31-37.

54. Nussbaumer Y, Bloch KE, Genser T, Thurnheer R. Equivalence of autoadjusted and constant continuous positive airway pressure in home treatment of sleep apnea. *Chest.* 2006;129(3):638-643.

55. Patruno V, Aiolfi S, Costantino G, et al. Fixed and autoadjusting continuous positive airway pressure treatments are not similar in reducing cardiovascular risk factors in patients with obstructive sleep apnea. *Chest.* 2007;131(5):1393-1399.

56. Planes C, D'Ortho MP, Foucher A, et al. Efficacy and cost of home-initiated auto-nCPAP versus conventional nCPAP. *Sleep.* 2003;26(2):156-160.

57. Randerath WJ, Schraeder O, Galetke W, Feldmeyer F, Ruhle KH. Autoadjusting CPAP therapy based on impedance efficacy, compliance and acceptance. *Am J Respir Crit Care Med.* Mar 2001;163(3 Pt 1):652-657.

58. Richard W, Venker J, den Herder C, et al. Acceptance and long-term compliance of nCPAP in obstructive sleep apnea. *Eur Arch Otorhinolaryngol.* 2007;264(9):1081-1086.

59. Senn O, Brack T, Matthews F, Russi EW, Bloch KE. Randomized short-term trial of two autoCPAP devices versus fixed continuous positive airway pressure for the treatment of sleep apnea. *Am J Respir Crit Care Med.* 2003;168(12):1506-1511.

60. Teschler H, Wessendorf TE, Farhat AA, Konietzko N, Berthon-Jones M. Two months auto-adjusting versus conventional nCPAP for obstructive sleep apnoea syndrome. *Eur Respir J.* 2000;15(6):990-995.

61. Vennelle M, White S, Riha RL, Mackay TW, Engleman HM, Douglas NJ. Randomized controlled trial of variable-pressure versus fixed-pressure continuous positive airway pressure (CPAP) treatment for patients with obstructive sleep apnea/hypopnea syndrome (OSAHS). *Sleep.* 2010;33(2):267-271.

62. West SD, Jones DR, Stradling JR. Comparison of three ways to determine and deliver pressure during nasal CPAP therapy for obstructive sleep apnoea. *Thorax.* 2006;61(3):226-231.

63. Berry RB, Hill G, Thompson L, McLaurin V. Portable monitoring and autotitration versus polysomnography for the diagnosis and treatment of sleep apnea. *Sleep.* 2008;31(10):1423-1431.

64. Kuna ST, Gurubhagavatula I, Maislin G, et al. Noninferiority of functional outcome in ambulatory management of obstructive sleep apnea. *Am J Respir Crit Care Med.* 2011;183(9):1238-1244.

65. Rosen CL, Auckley D, Benca R, et al. A multisite randomized trial of portable sleep studies and positive airway pressure autotitration versus laboratory-based polysomnography for the diagnosis and treatment of obstructive sleep apnea: the HomePAP study. *Sleep.* 2012;35(6):757-767.

66. Chen LM, Kennedy EH, Sales A, Hofer TP. Use of health IT for higher-value critical care. *N Engl J Med.* 2013;368(7):594-597.

APPENDIX A. SEARCH STRATEGIES

Ovid (Medline) KQ1 and KQ2:

1. Sleep Apnea Syndromes/di,th [Diagnosis, Therapy]
2. Sleep Apnea, Obstructive/di,th [Diagnosis, Therapy]
3. (protocol: or algorithm:).mp.
4. Patient care team/ or nurse's practice patterns/ or health personnel/ or allied health personnel/
5. Sleep apnea syndromes/nu or sleep apnea, obstructive/nu
6. (nurse led or nurse-led).ti,ab.
7. (nurse: or nursing or technician:).mp.
8. Primary health care/ or physicians/ or (nurse* or technician or special* or primary care or physician).ti,ab.
9. "referral and consultation"/ or (electronic adj consult).mp. or consult*.mp. or telemedicine/ or remote consultation/
10. Mass screening/
11. Continuous positive airway pressure/mt, nu [methods, nursing]
12. Polysomnography/nu [nursing]
13. Chart review.mp. or risk assessment/
14. 1 or 2
15. 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13
16. 14 and 15
17. Limit 16 to (English language and yr = "2000-Current")
18. Limit 17 to "all child (0 to 18 years)"
19. Limit 17 to "all adult (19 plus years)"
20. 18 not 19
21. 17 not 20

Ovid (Medline) KQ3:

1. Sleep Apnea Syndromes/th [therapy] or Sleep Apnea, Obstructive/th [therapy]
2. (titrat* and (manual or conventional or standard or fixed or (auto* or APAP))).mp.
3. Home care services/
4. (positive-pressure respiration/ or continuous positive airway pressure/ or intermittent positive-pressure ventilation/ or CPAP.mp.) and (calibration/ or (telemetric or titrat*).mp.)
5. 2 or 3 or 4
6. 1 and 5
7. Limit 6 to (English language and yr = "2000-Current")
8. Limit 7 to "all child (0 to 18 years)"
9. Limit 7 to "all adult (19 plus years)"
10. 8 not 9
11. 7 not 10

CINAHL KQ1 and KQ2:

1. (MH “Sleep Apnea Syndromes/DI/TH”)
2. (MH “Sleep Apnea, Obstructive/DI/TH”)
3. AB (protocol* or algorithm*)
4. (MH “Multidisciplinary Care Team”) OR “MH “Team Nursing”) OR (MH “Total Patient Care Nursing”)
5. (MH “Nursing Practice”) OR (MH “Scope of Nursing Practice”)
6. (MH “Health Personnel”) OR (MH “Allied Health Personnel”)
7. (MH “Sleep Apnea Syndromes/NU”) OR (MH “Sleep Apnea, Obstructive/NU”)
8. “nurse led”
9. “nurse-led”
10. AB (nurse* or nursing or technician*)
11. (MH “Primary Health Care”)
12. (MH “Physicians”)
13. (MH “Referral and Consultation”) OR (MH “Remote Consultation”)
14. (MH “Telemedicine”)
15. AB (electronic adj consult) OR AB consult*
16. (MH “Health Screening”)
17. (MH “Continuous Positive Airway Pressure/MT/NU”)
18. (MH “Polysomnography/NU”)
19. 1 or 2
20. S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18
21. 19 AND 20 (Limits: Published dates 2000 to present, English Language)

CINAHL KQ3:

1. (MH “Sleep Apnea Syndromes/TH”) OR (MH “Sleep Apnea, Obstructive/TH”)
2. (MH “Positive Pressure Ventilation”) OR (MH “Continuous Positive Airway Pressure”) OR (MH “Intermittent Positive Pressure Ventilation”)
3. (MH “Calibration”)
4. AB “telemetric or titra*
5. 3 OR 4
6. 2 AND 5
7. AB titra* AND AB ((manual or conventional or standard or fixed or (auto* or APAP)))
8. (MH “Home Health Care”)
9. 6 OR 7 OR 8
10. 1 AND 9 (Limits: Published dates 2000 to present, English Language)

APPENDIX B. PEER REVIEW COMMENTS/AUTHOR RESPONSES

REVIEWER COMMENT	RESPONSE
1. Are the objectives, scope, and methods for this review clearly described?	
Yes	Thank you
Yes	
Yes	
Yes	
Yes	
Yes	
No - See comments. Methods incomplete. Concepts appeared in Results and Conclusions that were omitted in the Introduction and Methods.	Thank you for the suggestions. We address the specific issues in the comments below.
2. Is there any indication of bias in our synthesis of the evidence?	
No	Thank you
No	
No	
No	
No	
No	
No	
3. Are there any published or unpublished studies that we may have overlooked?	
No	Thank you
Yes - Epidemiology should be updated. Consider ref Peppard Am J Epid 2013. Consider reference to cost associated with sleep apnea (Frost and Sullivan report just released, available on AASM website). The operational partner also has data on staffing, workload, and prosthetic costs related to sleep apnea, which may support the review (looking at alternative provider types to deliver sleep care, burden of disease within VA, etc.)	Thank you for the suggested references. We added the Peppard 2013 data to the Introduction. Although we typically only include data from peer review journals, we have included the Frost & Sullivan report for the AASM.
No	Thank you
No	
No	
No	
No	
4. Additional suggestions or comments can be provided below. If applicable, please indicate the page and line numbers from the draft report.	

I congratulate the sponsor and the ESP team on the conduct of a methodologically rigorous evidence synthesis report on a topic of great clinical importance. (1) The introduction (page 1) suggests that screening is most appropriate for symptomatic patients. I agree with this suggestion. This text, and the related text on page 10, however did not mention the American Heart Association/American Stroke Association (AHA/ASA) Stroke Prevention Guidelines which recommends screening all patients with stroke for OSA (regardless of symptoms) given the robust evidence that OSA is present in the overwhelming majority of stroke patients and the evidence that treatment with PAP improves outcomes. I believe that none of the articles that were the basis of those AHA/ASA guideline recommendations would be included in this ESP review because they did not meet inclusion criteria (e.g., to my knowledge, none compared CPAP to APAP). (2) I do not understand why dashed lines were used in the conceptual model figures from the Intermediate to the Clinical Outcomes boxes. Perhaps a legend might be added. (3) I like the use of italics to distinguish "titration" versus "treatment" in the last question. (4) In non-VA settings, many facilities use standard screening tools (e.g., Berlin) to identify patients for PSG-referral (without the intervening oversight of either a primary care provider or sleep specialist). This alternative model could be mentioned in the Background or Discussion related to question 1. (5) Given that the gaps in literature section may be useful to investigators writing IIR applications, I encourage the authors to be comprehensive in describing studies that would advance the field. For example, there are many alternative approaches to case finding (e.g., screening tools, direct use of APAP as a diagnostic test) and treatment (e.g., remote PAP monitoring, mandibular advancement devices and other approaches for PAP-intolerant patients) that merit investigation.	Thank you 1) Thank you. It is outside the scope of this review to evaluate guidelines and recommendations for screening (for example, after stroke). Our purpose in the introduction was to “set the stage” for our evaluation of alternative care models. 2) A legend (footnote) has been added. 3) Thank you. 4) We have now included this as part of our expanded explanation of ‘screen-detected’ patients and noted the limited data on screen-detected patients in the Applicability and Implementation section. 5) We have expanded the gaps/future research section.
Overall excellent job with the review!! Very comprehensive and highlights important evidence gaps. Under Research Gaps/Future Research: consider suggesting further evaluation on the implementation of the recommendations, not just CER of Key question 1. Key Question 2 has no published data/trials. A large pragmatic approach to studying this within the VA system is not only feasible but should be done since several programs have been using e-consults and never published outcomes. Any way to bridge this for focused HSRD would be helpful. Lastly, consider summarizing the burden of OSA, lack of providers, and clinical need for sleep medicine before wading into outcomes. This would be helpful for folks who skip to the summary directly.	Thank you. We have modified the Executive Summary to include the burden of OSA and to include the discussion of Key Question 2, as we agree with the reviewer statement.
The report is excellent! I thought the Executive Summary was too long.	Thank you. We shortened the Executive Summary.
(1) It is unfortunate that the studies by Berry, Kuna and Rosen were excluded from analysis and discussion (see page 51, line 3). These are three of the most important studies in the area of alternative care models for treatment of OSA, as the models of	1) Thank you for the suggestion. As the reviewer notes, these studies also ‘bundled’ HST vs. PSG with APAP vs. CPAP titration, and comparing HST vs.

care in practice are not restricted to home auto-CPAP versus laboratory-based CPAP titration. Efficiency of care is more strongly impacted by the use of home (HST) versus lab-based (PSG) testing. While this ESP report does not address the evidence supporting HST use as an alternative to PSG, HST use for diagnosis of OSA has become common clinical practice, especially within the VA, and these studies provide strong evidence that the HST-APAP approach achieves equivalent outcomes to the PSG-CPAP approach. This strengthens the current findings of key Question 3 regarding similarity of outcomes with APAP and CPAP. I strongly urge the authors to include a detailed analysis of these papers under Key Question 3. (2) Although it would significantly alter the scope of the report, an additional Key Question asking whether HST is a valid alternative to PSG for diagnosis of OSA would be welcome. Unfortunately, the USPSTF draft report was seriously flawed in this area, in particular failing to take into account the impact of night-to-night variation in OSA severity on PSG-HST comparisons. (3) Finally, the authors correctly observe that in the studies comparing sleep specialist to non-sleep specialist care, the non-specialists generally have had additional training or experience in management of sleep apnea patients. While this important point is clearly stated under Limitations in the Evidence Report (page 50, penultimate paragraph), it is absent from the Executive Summary. As most readers will probably not get beyond the Executive Summary, it is important to include there as well, and also in the Summary of Findings for Key Question #1 on page 22.	PSG was clearly not a goal of our study. However, the APAP vs. CPAP titration data generally suggested the same findings as we found in other studies, so we have expanded our discussion about these articles. Even though they did not strictly meet our inclusion criteria, . 2) We are not able to alter the scope at this time. HST vs PSG could be nominated for a future ESP report. 3) Thank you for the suggestion. We added this (and other) limitations to the Executive Summary.
Very clear layout of questions, only a few things that require clarification, major one is defining RDI- term has been used variably over time so this term needs to be defined for each study that used it, written in a very accessible way for the non-sleep specialist Page 6, line 36: Does this mean that care between the 2 groups was similar or that only the SSP group showed improvement (I think you mean the former but this is a little unclear). Page 7, line 6: RDI not defined prior to this; the definition of RDI has varied over time, for example prior to about the early 2000's RDI was equivalent to AHI, but at time subsequently RDI has included RERAs (Respiratory Effort Related Events). Recommend that this evolution in definition as it pertains to the studies that used the term, be stated somewhere in the document. Page 9, line 54: term "screen-detected pts" is used frequently, may be helpful at beginning of document to define this term- unclear to me if this means screened with validated measure such as a questionnaire or something else Page 28, line 45: The choice of this metric (≥ 4 hours of CPAP use on ≥ 70% of nights) for compliance is not based on any data. Wondering if this should be explicitly stated somewhere as we don't want to give the idea that this is # is sufficient for compliance.	Thank you for the suggestions. We made changes to clarify the text. Page 6. This sentence has been modified. Page 7, line 6: We added what this study reported for their RDI definition. Another study used RDI as part of a pre-test probability determination but did not define RDI. Page 9: We have clarified this term—thank you. Page 28. We comment in several places that measures of compliance varied. We extracted the compliance measures as reported in the included studies. We agree with the reviewer regarding the lack of evdidence for the ' > 4hrs for 70% of nights'

	metric, but given that many studies report this (and it is used for reimbursement—we did add this statement on page 27), we felt compelled to include it.
Page 37, line 10: "Technologist time was higher on titration morning" Does this mean that physician interpretation took less time for APAP than in-lab CPAP titration? This is true and is due largely to the larger amount of data available for the in-lab studies."	Page 37, line 10. This statement has been clarified.
Page 37, line 13: "Physician time for titration study reporting was lower in the home APAP group" This statement is also unclear. Does it mean that physician interpretation requires less time for APAP than CPAP? This is true largely due to increased information available for the in-lab studies.	Page 37, line 13. his statement has been clarified.
Table 6: Does # on front of arrow indicate the number of studies showing this result? Is this information only listed for cases where the data could not be pooled?	Table 6: Yes, the number is the number of studies. All reported outcomes are presented on the arrow tables.
• An abstract would be helpful.	-Thank you for the suggestion. An abstract has been added.
• Introduction, p. 12, lines 30-32: Provide references for estimates for OSA prevalence.	-We revised and added data from Peppard 2013.
• The distinction between intermediate and clinical outcomes is unclear. What is the difference between "weight (sic), BMI, libido, blood pressure, or HBA1c (sic)" (intermediate) and "libido, weight change, BMI, blood pressure, or HbA1c" (clinical)?	-We consider clinical outcomes to be patient-centered outcomes – something the patient can feel. Therefore, we attempt to distinguish between a change in a sleep scale score or weight (for example) and a clinically meaningful change in sleep score or weight.
• I find the limitation to English-speaking and European (do you mean Western European) countries problematic, regardless of an (unstated) interest in making applicable to the VA. Why not high income parts of Asia, Middle East, and Latin America? By North America, do you mean just US and Canada? Similarly, what is the justification for limiting to English language? Google Translate (and other methods) are highly effective.	-As described in the Limitations, our goal was to identify studies most applicable to clinical practice in the US and the VA. Regarding the non-English language studies, Google Translate has been evaluated and has not reached acceptable levels of accuracy. Standard methods for systematic reviews including through AHRQ and the ESP are to limit inclusion to articles published in English language.
• Implicitly, you have restricted to published, peer-reviewed articles. Is this the case?	-Yes. We searched MEDLINE and CINAHL.
• If ESS and SF-36 are all on the same scale (as they presumably are), why were standardized mean differences (and not weighted mean differences) used? SMD is clinically difficult to interpret. (The Cohen reference call out is missing the year). What were your minimum criteria for conducting a meta-analysis (how few studies would you meta-analyze)?	-There is no established minimal important difference for the ESS, so we used SMDs to facilitate interpretation of effect (how large was the effect based on the suggested cut points of 0.2 (small), 0.5 (moderate), and 0.8 (large)) and how precise was the

	estimate (if the upper or lower confidence limit crosses an effect size of 0.5 in either direction this would be considered imprecise). We will present the data as a WMD and an SMD. References have been converted to superscript format. For this review we focused our meta-analysis on ESS and SF-36 scores.
• I would suggest a much more detailed, explicit explanation for assessing the strength of evidence. Owens 2010 gives general guidance. What specifically was done here? Particularly, when and how did you determine the evidence was insufficient? Table 4 (for example) is opaque. It is unclear why strength of evidence is ranked as it is and what were the strengths and weakness of the evidence. Were directness, precision, sparseness, dose-effect, etc. assessed? Also, it would be more helpful to the reader to divide Table 4 into separate tables for each KQ.	-We added more detail in the Methods section. The strength of evidence was insufficient if no studies reported the outcome (*eg,* access to care) or if there was one small study with few events. Directness etc. were assessed and a table has been added (Appendix D). We divided Table 4 into separate tables for each KQ.
• Results summaries. It would be helpful to incorporate strength of evidence into the summaries. The summaries are written in a highly subjective manner, suggesting interpretation by the researchers not objective summary. Examples from KQ 1 include "good" [agreement], "similar", and "may be". • The abbreviation SSP is not used consistently. • Statements such as "At baseline the patients' average age was 55.2 years…" are unclear. This sounds like the description of a single study by appear to be a summary across the 8 studies. Where did 55.2 come from? Were the mean ages meta-analyzed? Is this the median?	-Strength of evidence has been added to the summary statements at the start of each KQ. Similar is a standard term as is "may be" when evidence is very low or even low. -The abbreviation is now used throughout. -The statements are intended to provide an overview of the population in the included studies. Table 1 indicates that the values are means (unless otherwise noted) and reports the number of studies included in determining the mean. The means are weighted means. This has been added to the tables.
• In what ways were high risk of bias (and medium risk of bias) studies likely to be biased? • While the tables succinctly summarize the results of the studies, they provide highly limited data. While the Appendix tables provide the details, they are too difficult to read. I would recommend a separate set of tables in the main text that provide the summary numerical results. Also the call out to "Appendix Tables 2-3" left out which appendix (C). And the formatting of the vertical portions of the tables needs fixing.	-Details on risk of bias are presented in Appendix C, Tables 1 and 7 (Study Characteristics) -Thank you for the suggestion. There were a variety of measures used for the different outcomes (*eg,* different components of the SF-36) and then different reporting of the results (*ie,* mean differences, effect sizes). We thought the "arrow" tables and the Strength of Evidence tables were the best way to convery results in the text. The reference to Appendix C and the vertical alignment have been corrected.
• Figure 3 has "Mental Health" in the wrong place. It should be below the headers. Overall, the figure is unclear. It is not at all clear that the sub-analyses are in fact sub-analyses. Without close inspection, there appear to be 5 studies each for mental health and vitality (3 nurse and 2 PCP). A sub-analysis of a single study (PCP) is uninformative and misleading. The text part of the figure gives no clear distinction between the studies,	-The Figures have been revised for clarity deleting 'sub-analysis" lines and additional information in the legend. WMDs are presented.

sub-analyses, and overall analyses. What are the sample sizes, baseline values, follow-up times of the individual studies? Again, why is this done as a SMD and not a WMD? The figure needs a proper legend. Std diff and CI are not defined.	
• Page 27, lines 19 and following. It is not clear why adverse events are summarized if the RCT did not provide separate data for the two treatment groups. There is no evidence to address the KQ. Also, the term treatment group is unclear for KQ 1 which compares providers not treatments.	-We agree and have modified the statements about adverse events for KQ1 throughout the report. We also changed the “treatment group” wording.
• Tables 2 and 3, with thought, partly distinguish the difference between clinical and intermediate outcomes better than the methods but the distinction seems to be more categorical vs. continuous rather than clinical vs. intermediate. Table 3 should probably better clarify that weight and BP etc. are continuous outcomes.	-As noted above, we considered weight and symptom scores as intermediate outcomes but attainment of a minimally important difference in one of those outcomes as a clinical or patient-centered outcome. We attempted to clarify in the summary statements.
• Limitations: Consider whether you are able to make any determination about possibility of bias due to the agendas of the authors. It seems plausible that studies of non-specialists vs. specialists (or home vs. lab titration which could greatly affect lab income) are being conducted by researchers with an agenda (eg, to promote non-specialists).	-We could not determine bias, but these are by nature unblinded studies. However, the reviewer raises an important issue and we now suggest (in the Executive Summary and full report) that future studies have outcomes collected in a blinded fashion where feasible.
• Research Gaps/Future Research: Consider also talking about within-study gaps, particularly related to possible reporting bias. KQ 1 and 2 are explicitly discussed in this section; why isn't KQ 3?	- We have added some gaps/future research regarding KQ3 in this section (but only in the full report, rather than the Executive Summary—we felt that the gaps in KQ1 and KQ2 were more important for the Executive Summary). Thank you.
• Conclusions: I believe the conclusions section is the first mention of a decreasing supply of sleep physicians.	Thank you for noting this—we now discuss the decreasing supply in the introduction.

APPENDIX C. EVIDENCE TABLES

Table 1. Study Characteristics for KQ1

Author, year Study Design Location Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/Exclusion Criteria	Patient Characteristics (expressed in mean (sd) unless otherwise noted)	Risk of Bias
Chai-Coetzer 2013[31] RCT Australia Participants screened in primary care	Primary care management (n = 81) Usual care, sleep specialist (n = 74) 6 months	Case finding (diagnosis) and Management Primary care management included physicians and community-based nurses who participated in an education program on obstructive sleep apnea and its management. Patient treated with CPAP, conservative therapy, mandibular advancement splint. One of the 4 nurses had 15 years of experience in a tertiary sleep center.	Inclusion: aged 25-70, high diagnostic likelihood of moderate to severe OSA, defined as a score of ≥ 5 out of 10 points on a 4-item questionnaire and an overnight 3% oxygen desaturation index (≥ 3%ODI) of ≥ 16 events per hour and an ESS score of ≥ 8 or persistent hypertension despite taking ≥ 2 antihypertensive agents Exclusion: severe morbid obesity (BMI > 50), neuromuscular disease, unstable psychiatric disease or cognitive impairment considered likely to interfere, hospitalization in the previous 3 months for MI, unstable angina, cardiac failure, or CVA or New York Heart Association class III or IV symptoms, or lung disease with awake resting oxygen saturation of <92%	N = 155 Primary Care: n = 81 Mean age: 57.2 (10.9) Male gender, %: 85 BMI: 33.1 (5.5) Oximetry ≥ 3% ODI, events/h: 32.7 (18.2) BP, systolic mmHg: 134 BP, diastolic mmHg: 84.5 ESS: 12.8 (3.9) OSA 50 questionnaire score: 8.2 (1.5) Specialist: n = 74 Mean age: 54.5 (11.8) Male gender, %: 77 BMI: 33.7 (5.6) Oximetry: 35.7 events/h (17.4) BP, systolic mmHg: 135.9 BP, diastolic mmHg: 85.23 ESS: 12.5 (3.9) OSA 50 questionnaire score: 8.1 (1.7)	**Sequence generation:** adequate **Allocation concealment:** adequate **Blinding:** NR **Incomplete outcome data:** reasons for dropout reported, ITT analyses (all randomized) but uneven dropouts by arm (21% and 8%) and > 10% dropped out **Selective outcome reporting:** no **Risk of Bias: Medium**

Author, year Study Design Location Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/Exclusion Criteria	Patient Characteristics (expressed in mean (sd) unless otherwise noted)	Risk of Bias
Chamorro 2013[32] Retrospective record review Spain Sleep unit	Primary care pulmonologist vs sleep specialist (n = 96) unclear	Diagnosis Examine concordance between test prescribed by primary care pulmonologist and ideal test recommended by sleep specialist	Inclusion: patients with suspicion of OSAH referred to sleep unit by primary care pulmonologist in 2010	n = 96 Mean age: 58.7 (12.6) Male gender, %: 71 BMI: 30.26 (5.39) ESS: 11.57 (4.7) HTN: 38% Diabetes: 14%	**Selection bias:** inadequate **Blinding of outcome assessment:** inadequate **Intention-to-treat analysis:** adequate **Attrition bias:** adequate **Selective outcome reporting:** adequate **Risk of Bias: High**
Andreu 2012[29] RCT Spain Pulmonology section of the University Hospital	**Group A:** Home respiratory polygraphy and home follow-up by sleep unit nurse (n = 22) **Group B:** Supervised polysomnography and hospital follow-up with sleep unit pulmonologist (n = 22) **Group C:** Home respiratory polygraphy and hospital follow-up with sleep unit pulmonologist (n = 21) 6 months	Treatment All received CPAP	Inclusion: high level of clinical suspicion of OSAS based on an Epworth Sleepiness Scale (ESS) score ≥ 12 and a Sleep Apnoea clinical score (SACS) ≥ 15 Exclusion: impaired lung function (COPD, obesity-hypoventilation, and restrictive disorders), associated pathologies (psychiatric disorders, neoplasms, restless leg syndrome, and other dyssomnias or parasomnias), or previously treated with CPAP	n = 65 Mean age: 52 Male gender, %: 83 BMI: 34 Hypertension: 49% Habitual snoring: 100% SACS: 40 (26) AHI(/hr): 43 (20) ODI(/hr): 44 (26) Neck circumference (cm): 45.5 (3.5)	**Sequence generation**: adequate **Allocation concealment:** adequate **Blinding:** no **Incomplete outcome data:** 58/65 (89%) completed program; intent-to-treat analysis included all but 1 patient (refused PSG) **Selective outcome reporting**: some outcomes not reported by group but overall adequate **Risk of Bias: Medium**

Author, year Study Design Location Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/Exclusion Criteria	Patient Characteristics (expressed in mean (sd) unless otherwise noted)	Risk of Bias
Pamidi 2012[35] Retrospective chart review Chicago University sleep disorders center	Sleep Specialists, initial PSG ordered by a sleep specialist (n = 105) Non-sleep Specialist, Initial PSG ordered by a non-sleep specialist (60% primary care physicians, 8% otolaryngologists, 7% pulmonologists, 6% neurologists, 6% endocrinologists, 5% cardiologists, 3% surgeons, and 4% other) (n = 298) 30 days	Case finding Referred patients received in-lab PSG and CPAP titration done by sleep laboratory personnel, had CPAP set up in homes by a durable medical equipment provider	Inclusion: evaluated medical records of adults who were CPAP naïve and were referred for their first in-laboratory PSG for suspicions of OSA Exclusion: previous CPAP use, requirement for bi-level PAP or adaptive servoventilation, central sleep apnea, and lack of adherence data due to a lack of or faulty wireless modem transmission device	n = 403 Mean age:52.5 (14) Male gender, %: 47 Race, African American: 54% Non-African American: 46% (significantly fewer African Americans in sleep specialist group) BMI: 36.3 (9.1) Hypertension: 58.5% (significantly more hypertensives in sleep specialist group) T2DM: 26% ESS: 9.2 (5.2) CES-D scale: 16 (11) Total sleep time, min: 324, P = .98 Arousal index, events/h: 29, P = .54 AHI(/hr): 36 ODI(/hr): 23 (significantly higher in non-sleep specialist group) SpO2: 80.6 (9.8)	**Selection bias:** adequate **Blinding of outcome assessment:** N/A (objective – wireless transmission of data) **Intention-to-treat analysis:** adequate **Attrition bias:** inadequate **Selective outcome reporting:** adequate **Risk of Bias: Medium**

Author, year Study Design Location Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/Exclusion Criteria	Patient Characteristics (expressed in mean (sd) unless otherwise noted)	Risk of Bias
Lettieri 2011[33] Observational cohort study United States Community-based Hospital and Academic Sleep Center	**Group1 (n = 70):** primarily managed by primary care physician **Group2 (n = 70):** managed by sleep specialist **Group3 (n = 70):** managed by sleep specialist 4-6 week follow-up	Titration All treated with CPAP, titration via home APAP, or in-lab CPAP	Inclusion: diagnosed with OSA by HST or PSG (met criteria for OSAS according to AASM guidelines), eligible for home sleep study program (2 or more high-risk features such as habitual snoring, daytime fatigue, nonrestorative sleep, weight gain, and witnessed apneas; no suspicion of concomitant sleep disorders and no significant underlying comorbidities), diagnosed OSAS defined as AHI > 5 with compatible symptoms Exclusion: not eligible for HST (cardiopulmonary disease, heart failure, CAD, previous cerebrovascular accident, poorly controlled asthma, moderate to severe COPD [FeV1 < 50%], supplemental oxygen requirement)	n = 210 Group 1: Mean age: 50.4 (9.2) Male gender, %: 64.3% BMI: 32.2 (4.8) Baseline ESS: 14.8 (54.8) Baseline fatigue: 6.3 (1.5) AHI(/hr): 20.7 (12.2) Group 2: Mean age: 47.1 (8) Male gender, %: 71.4 BMI: 30 (3.5) Baseline ESS: 14.1 (4.2) Baseline fatigue: 6.7 (1.7) AHI(/hr): 23.1 (13) Group 3: Mean age: 45.5 (5.4) Male gender, %: 68.6% BMI: 28.5 (3) Baseline ESS: 13.9 (4.4) Baseline fatigue: 6.5 (1.4) AHI(/hr): 19.3 (9.4)	**Selection bias:** inadequate; unclear how patients in Groups 2 and 3 were selected to achieve same number as in Group 1; participants in <u>all</u> groups had to meet the same criteria for HST and APAP **Blinding of outcome assessment:** NR (adherence was objective measure) **Intention-to-treat analysis:** adequate after exclusions for Group 1 and selection of equal number for Groups 2 and 3 **Attrition bias:** adequate **Selective outcome reporting:** adequate **Risk of Bias: Medium**

Author, year Study Design Location Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/Exclusion Criteria	Patient Characteristics (expressed in mean (sd) unless otherwise noted)	Risk of Bias
Antic 2009[30] RCT Australia Academic sleep medicine services, after referral for clinical suspicion of OSA	Specialist nurse (n = 100 randomized, 90 analyzed) Sleep physician (n = 95 randomized, 84 analyzed) 3 months	Management Specialist nurse was experienced in sleep disorders, supervising home auto-adjusting positive airway pressure to set therapeutic continuous positive airway pressure (CPAP). Sleep physician group had clinical care supervised by a sleep physician for in-lab CPAP titration and treatment	Inclusion: referred with a clinical suspicion of OSA, ESS score of ≥ 8, history of snoring ''most nights'' or ''every night,'' age 18-75 years, and patient willing to try CPAP Exclusion: unstable cardiovascular diseases (*eg,* recent unstable angina, myocardial infarction, stroke or TIA within the previous 6 months, or severe left ventricular failure), neuromuscular disease affecting or potentially affecting respiratory muscles, moderate to severe respiratory disease or hypoxemia or awake SaO_2<92%, or psychiatric disease that limited the ability to give informed consent or complete the study	n = 195 Nurse-led: n = 100 Mean age: 49.9 (SEM 1.2) Male gender, %: 72 BMI: 35.1 (SEM.7) ≥ 2% oxygen saturation dips, events/h: 49.2 (SEM 2.1) ESS: 13.7 (SEM 0.4) Specialist-led: n = 95 Mean age: 50.3 (SEM 1.3) Male gender, %: 76 BMI: 34 (SEM.6) AHI: 67.9 events/h (SEM 2.82) > / = 2% oxygen saturation dips, events/h: 52.5 (SEM 2.7) ESS: 13.4 (SEM 0.4)	**Sequence generation:** adequate **Allocation concealment**: adequate (explained in online supplement) **Blinding:** open-label but questionnaires and measurements administered by research assistants with no involvement in clinical care of patients and were blinded to patient allocation **Incomplete outcome data**: reasons for dropout reported, # randomized were not included in the analyses **Selective outcome reporting:** No **Risk of bias: Low**
Palmer 2004[34] RCT Scotland Hospital/peripheral clinics	Specialist nurse (n = 87 randomized, 79 at baseline, 68 at follow-up) Hospital-based consultant (n = 87 randomized, 77 at baseline, 71 at follow-up) 3 months	OSA management/ treatment Home visit by specialist nurse or hospital-based consultant review at general respiratory clinic for routine annual review for CPAP users	Inclusion: All patients in Highland who had a diagnosis of SAHS and a CPAP machine on 10/01/2000 Exclusion: Not described(none)	n = 174 randomized, 156 at baseline, 139 at follow-up Nurse: n = 79 at baseline Age: 54 (10) Male gender, %: 84 ESS: 8 (5) Consultant Clinic: n = 77 at baseline Age: 55 (11) Male gender, %: 87 ESS: 9 (6)	**Sequence Generation**: NR **Allocation concealment:** NR **Blinding:** NR **Incomplete outcome data:** unclear, 80% of target population finished both baseline and follow-up questionnaire, some reasons given, not uneven **Selective Reporting:** adequate **Risk of Bias: Medium**

Author, year Study Design Location Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/Exclusion Criteria	Patient Characteristics (expressed in mean (sd) unless otherwise noted)	Risk of Bias
Scharf 2004[36] Retrospective telephone survey and laboratory chart review USA University Specialty Hospital and a laboratory serving the medical community at large	Primary care practitioners (n = 44) Usual care, sleep specialists (n = 59) The mean time from diagnostic PSG to interview for primary care group was 7.0 months and 7.2 months for usual care	Management In primary care group all patients referred by primary care practitioners for usual care patients were seen by sleep specialists. All treated with CPAP	Inclusion: over 18 years old diagnosed with OSA Exclusion: NR	n = 103 Mean age: 49.4 (12.7) Male gender, %: 58.2 BMI: 36.1 (13.4) Diabetes: 19% HTN: 53% Unexplained daytime sleepiness/fatigue: 68% Snoring: 83%	**Selection bias:** adequate groups; unclear regarding possible confounders **Blinding of outcome assessment:** NR **Intention-to-treat analysis:** adequate **Attrition bias:** 37% survey response rate but comparable for 2 sites and non-responders were similar age, gender, RDI **Selective outcome reporting:** adequate **Risk of Bias: Medium**

AHI = apnea/hypopnea index; CES-D = Center of Epidemiology study depression scale; COPD = chronic obstructive pulmonary disorder; ESS = Epworth Sleepiness Scale; ODI = oxygen desaturation index; OSAS = obstructive sleep apnea syndrome; SpO2 = oxygen saturation measured by pulse oximetry; T2DM = type 2 diabetes mellitus

Table 2. Clinical Outcomes for KQ1

Study Intervention (n) Control (n)	All-cause Mortality % (n/N)		Resource Utilization (hospitalization, etc)		Access to Care[a]		MID, Sleep Symptom Scores		MID, Urinary Symptom Scores	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Chai-Coetzer 2013[31] Primary care (n = 81) Usual care (n = 74)	NR	NR	Baseline initiated CPAP: 90% (n = 73) conserva-tive measures: 2% (n = 2) 6 months using CPAP: 63% (n = 51) conserva-tive measures: 9% (n = 7)	Baseline: initiated on CPAP: 70% (n = 52) conserva-tive measures: 24% (n = 18) 6 months using CPAP: 61% (n = 45) conserva-tive measures: 16% (n = 12) RR of using CPAP at 6 m: 1.11 (0.95, 1.31)	NR	NR	NR	NR	NR	NR
Andreu 2012[29] Nurse (n = 22) Sleep pulmonologist (n = 43)	NR	NR	Extra visits 9 Extra calls 24	Extra visits **Group B**: 0 **Group C:** 5 Extra calls **Group B:** 17 **Group C**: 13	NR	NR	NR	NR	NR	NR

Study Intervention (n) Control (n)	All-cause Mortality % (n/N)		Resource Utilization (hospitalization, etc)		Access to Care[a]		MID, Sleep Symptom Scores		MID, Urinary Symptom Scores	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Antic 2009[30] Nurse (n = 100) Usual care (n = 95)	NR	NR	Number of physician visits per patient 0.2 (SEM 0.1) Effect size: 2.24 (1.88, 2.6) Scheduled nursing time per patient 153 min (SEM 3.9) P<.001 Un-scheduled nursing time per patient 8.4 min (SEM 1.5) Effect size: -0.15 (-0.43, 0.13)	Number of physician visits per patient 2.4 (SEM 0.1) P<.001[a] Scheduled nursing time per patient 103 min (SEM 4.2) Effect size: 1.25 (0.94, 1.56) Un-scheduled nursing time per patient 11.4 min (SEM 2.5) P = .31[a]	NR	NR	NR	NR	NR	NR

Study Intervention (n) Control (n)	All-cause Mortality % (n/N)		Resource Utilization (hospitalization, etc)		Access to Care[a]		MID, Sleep Symptom Scores		MID, Urinary Symptom Scores	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Palmer 2004[34] Nurse (n = 68) Consultant Clinic (n = 71)	NR	NR	61% of patients seen by the consultant required onward referral to specialist nurse for practical help Average time spent with nurse: 26 (6) minutes Average time spent with consultant: 10 (6) minutes Effect size: 0.32 (-0.01, 0.66)		NR	NR	NR	NR	NR	NR
Scharf 2004[36] Primary care (n = 44) Usual care (n = 59)	NR	NR	Patients offered CPAP 79% (35/44) RR: 0.92 (0.77, 1.1) P = .367 Accepted treatment with CPAP 83% (29/35) P = NS[a]	Patients offered CPAP 86% (51/59) P = NS[a] Accepted treatment with CPAP 86% (44/51) RR: 0.96 (0.8, 1.16)	NR	NR	NR	NR	NR	NR

[a]between groups

SEM = standard error of the mean; RR = risk ratio; NS = not statistically significant; NR = not reported

Table 3. Clinical Outcomes for KQ1, Continued

Study Intervention (n) Control (n)	Quality of Life		Patient Satisfaction		Remission		Cognitive Symptoms		Other (describe)	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Chai-Coetzer 2013[31] Primary care (n = 81) Usual care (n = 74)	SF-36 Vitality Baseline 43.6 Change at 6 m 16.1 (11.0, 21.2) Adjusted difference: 2.51 (-3.88, 8.9) P<.001 from baseline SF-36 Mental Baseline 66.5 Change at 6 m 7.9 (4.0, 11.8) P<.001 from baseline P = .54[a]	SF-36 Vitality Baseline 34.6 Change at 6 m 19.9 (14.4, 25.4) P<.001 from baseline P = .44[a] SF-36 Mental Baseline 61.6 Change at 6 m 8.4 (4.5, 12.3) P<.001 from baseline Adjusted difference: 1.57 (-3.41, 6.55)	VSQ-9 Small but statistically significant differences in 5/9 items in favor of the primary care group No difference in overall satisfaction Effect sizes for the 9 items were small (range, 0.14-0.41) and may not be clinically significant		NR	NR	NR	NR	NR	NR

Study Intervention (n) Control (n)	Quality of Life		Patient Satisfaction		Remission		Cognitive Symptoms		Other (describe)	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Antic 2009[30] Nurse (n = 100) Usual care (n = 95)	SF-36 Vitality change at 3 m -16.12 (SEM 2.17); n = 89 Effect size: -0.04 (-0.34, 0.26) SF-36 Mental change at 3 m -4.81 (SEM 1.46); n = 89 Effect size: 0.017 (-0.28, 0.32) No significant differences between groups in any of the quality of life indices	SF-36 Vitality change at 3 m -15.31 (SEM 2.06); n = 81 Mean difference: -0.81 (-6.75, 5.12) SF-36 Mental change at 3 m -5.09 (SEM 2.11); n = 81 Mean difference: 0.27 (-4.71, 5.27)	VSQ-9 Total patient satisfaction with treatment was not statistically significantly different between the 2 groups Mean scores, Nurse 3.73 (SD 0.47); n = 89 vs UC 3.76 (SD 0.43); n = 79, P = .68[a] Effect size: -0.06 (-0.37, 0.24)		NR	NR	NR	NR	NR	NR

Study Intervention (n) Control (n)	Quality of Life		Patient Satisfaction		Remission		Cognitive Symptoms		Other (describe)	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Palmer 2004[34] Nurse (n = 68) Consultant Clinic (n = 71)	SF-36 PCS Baseline 39 (16) P = .10 3 months 39 (17) Change -1 (8) Effect size: 0.24 (-0.1, 0.57) MCS Baseline: 51 (11) P = .51 3 months 52 (13) Change 1 (7) No significant difference in any parameter between groups or from baseline for nurse Also reported SF-36 vitality and mental health scores	SF-36 PCS Baseline 34 (17) 3 months 35 (17) Change 2 (9) P = .16 MCS Baseline 49 (14) 3 months 51 (11) Change 2 (10) P = .64 Effect size: 0.08 (-0.25, 0.42) General health and social functioning both significantly improved from baseline P<.025 for consultant group	There were some "preference" data reported that were different (P = .00) by study arm		NR	NR	NR	NR	HADS Anxiety Baseline 6.1 (4.8) P = .5[a] 3 m 5.4 (5) Effect size: 0 (-0.33, 0.33) Change -0.6 (3.1) Depression Baseline 4.4 (4.3) 3 m 4.3 (4.4) Change 0.2 (2.9)	HADS Anxiety Baseline 6.7 (5.2) 3 m 5.4 (4.2) Change -1.1 (4.2) P = .54[a] Depression Baseline 5.5 (4.8) P = .18[a] 3 m 4.7 (4.4) Effect size: -0.09 (-0.42, 0.24) Change -0.6 (3.1) P = .27[a]

Study Intervention (n) Control (n)	Quality of Life		Patient Satisfaction		Remission		Cognitive Symptoms		Other (describe)	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Scharf 2004[36] Primary care (n = 44) Usual care (n = 59)	NR	NR	NR	NR	NR	NR	NR	NR	Subjective symptoms improve-ment (from diagnostic PSG to interview mean 7 m) 80% (28/35) RR: 1.09 (0.86, 1.4)	Subjective symptoms improve-ment (from diagnostic PSG to interview mean 7.2 m) 74% (36/49) P = NS[a]

[a]between groups

HADS = Hospital Anxiety and Depression Scale; NS = not statistically significant; NR = not reported; MCS = mental component summary (SF-36); PCS = physical component summary (SF-36); PSG = polysomnography; SEM = standard error of the mean; UC = usual care; VSQ-9 = Visit- Specific Satisfaction Questionnaire

Table 4. Intermediate Outcomes for KQ1

Study Intervention (n) Control (n)	Apnea-Hypopnea Index (AHI)		Oxygen Saturation		Sleep Symptom Scores		Other Sleep or Urinary Symptom Scores		Weight Loss	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Chai-Coetzer 2013[31] Primary care (n = 81) Usual care (n = 74)	NR	NR	NR	NR	ESS Baseline 12.8 6 months: 7.0 Change: 5.8 (4.4, 7.2) P<.001 from baseline Adjusted difference in mean change: -0.13 (lower bound -1.5) (for non- inferiority test)	ESS Baseline 12.5 6 months: 7.0 Change: 5.4 (4.2, 6.6) P<.001 from baseline P = .43[a]	FOSQ Baseline 14.7 Change at 6 m 2.8 (2.0, 3.6) P<.001 from baseline Adjusted difference: 0.18 (-0.58, 0.94) SASQ Baseline 71.2 Change at 6 m -29.7 (-23.0, -36.4) P<.001	FOSQ Baseline 14.2 Change at 6 m 2.8 (2.2, 3.4) P<.001 from baseline P = .64[a] SASQ Baseline 72.1 Change at 6 m -31.2 (-23.8, -38.6) P<.001 P = .85[a] Adjusted difference: 0.18 (-0.58, 0.94)	Baseline: 101.9kg Change at 6 m -0.1 (-2.5, 2.3) Adjusted difference: - 0.43 (-3.43, 2.57)	Baseline: 103.2 Change at 6 m 0.3 (-1.5, 2.1) P = .78[a]

Study Intervention (n) Control (n)	Apnea-Hypopnea Index (AHI)		Oxygen Saturation		Sleep Symptom Scores		Other Sleep or Urinary Symptom Scores		Weight Loss	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Andreu 2012[29] Nurse (n = 22) Sleep pulmonologist (n = 43)	NR	NR	NR	NR	ESS Baseline 15 (3) 6 months 6 (5) P<.001 from baseline P = NS[a] Effect size vs B: 0 (-0.59, 0.59) Effect size vs C: 0.22 (-0.38, 0.82)	ESS Baseline **B:** 16 (4) **C:** 16 (3) 6 months **B:** 6 (4) **C:** 5 (4) P<.001 from baseline for both groups	FOSQ Baseline 16 (3) 6 months 18 (2) P<.001 from baseline P = NS Effect size vs B: 0 (-0.6, 0.6) Effect size vs C: -0.63 (-1.24, -0.02)	FOSQ Baseline **B**: 16 (3) **C:** 16 (3) 6 months **B:** 18 (2) **C:** 19 (1) P<.001 from baseline	NR	NR
Pamidi 2012[35] Sleep specialist (n = 105) Non-sleep specialist (n = 298)	Events/hr Baseline 38 Residual 3.7 (median) P<.001[a]	Events/hr Baseline 31 P = .06[a] Residual 4.9	NR	NR	NR	NR	NR	NR	NR	NR

Study Intervention (n) Control (n)	Apnea-Hypopnea Index (AHI)		Oxygen Saturation		Sleep Symptom Scores		Other Sleep or Urinary Symptom Scores		Weight Loss	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Lettieri 2011[33] Group 1: Primary care (n = 70) Group 2: Sleep specialist (n = 70) Group 3: Sleep specialist (n = 70)	NR	NR	NR	NR	ESS **Group 1** Baseline 14.8 (4.8) P = .48[a] Follow-up 9.1 (3.6) P = .39[a] Effect size vs 2: 0.23 (-0.1, 0.56) Effect size vs 3: 0.07 (-0.26, 0.4) Change 38.5% P = .28[a]	ESS **G2:** Baseline 14.1 (4.2) Follow-up 8.4 (2.3) Change 39.8% **G3:** Baseline 13.9 (4.4) Follow-up 8.9 (2.1) Change 36%	NR	NR	NR	NR
Antic 2009[30] Nurse (n = 100) Usual care (n = 95)	NR	NR	NR	NR	ESS Baseline 13.7 Change at 3 m 4.02 (SEM 0.52); n = 90 MD = -0.13 (-1.52, 1.25) P = NS[a]	ESS Baseline 13.4 Change at 3 m 4.15 (SEM 0.47); n = 84	FOSQ Change at 3 m -13.6 (SEM 2.02); n = 89 MD = -0.38 (-5.97, 5.20) P = NS[a] Maintenance of wakefulness test, min Change at 3 m 30.18 (SEM 1.24)	FOSQ Change at 3 m -13.22 (SEM 1.96); n = 81 Maintenance of wakefulness test, min Change at 3 m 31.68 (SEM 1.08) MD -1.49 (-4.76, 1.78) P = NS[a]	NR	NR

Study Intervention (n) Control (n)	Apnea-Hypopnea Index (AHI)		Oxygen Saturation		Sleep Symptom Scores		Other Sleep or Urinary Symptom Scores		Weight Loss	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Palmer 2004[34] Nurse (n = 68) Consultant Clinic (n = 71)	NR	NR	NR	NR	ESS: Baseline 8 (5) P = .24[a] 3 months 8 (6) Effect size: 0 (-0.33, 0.33) Change 0.2 (4) Symptom Score Baseline P = .14[a] 3 months 11 (9) Effect size: -0.3 (-0.63, 0.04) Change -2 (7) P<.025 from baseline	ESS: Baseline 9 (6) 3 months 8 (6) Change -0.9 (4) P = .30 Symptom Score Baseline 17 (12) 3 months 14 (11) Change -3 (9) P<.025 from baseline P = .94[a]	NR	NR	NR	NR

[a]between groups

ESS = Epworth Sleepiness Scale (non-inferiority margin was -2.0); FOSQ = Functional Outcomes of Sleep Questionnaire; MD = mean difference; NS = not statistically significant; NR = not reported; RR = risk ratio; SASQ = Sleep Apnea Symptoms Questionnaire; SEM = standard error of the mean

Table 5. Intermediate Outcomes for KQ1, Continued

Study Intervention (n) Control (n)	BMI		Blood Pressure		HbA1c		Time to Initiation of Therapy		Compliance/Adherence	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Chai-Coetzer 2013[31] Primary care (n = 81) Usual care (n = 74)	NR	NR	Systolic Baseline: 134 mmHg Change at 6 m -2.2 (-6.3, 1.9) Adjusted difference: 1.52 (-4.14, 7.18) Diastolic Baseline: 84.5 mmHg Change at 6 m -1.4 (-4.3, 1.5) P = .48[a]	Systolic Baseline: 136 mmHg Change at 6 m -4.4 (-9.1, 0.3) P = .60[a] Diastolic Baseline: 85 mmHg Change at 6 m -0.5 (-3.6, 2.6) Adjusted difference: -1.32 (-4.97, 2.33)	NR	NR	NR	NR	Hours/night 4.8 (2.1) (n = 51) Effect size: -0.39 (-0.79, 0.02)	Hours/night 5.4 (0.30) (n = 44) P = .11[a]

Study Intervention (n) Control (n)	BMI		Blood Pressure		HbA1c		Time to Initiation of Therapy		Compliance/Adherence	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Andreu 2012[29] Nurse (n = 22) Sleep pulmonologist (n = 43)	NR	NR	NR	NR	NR	NR	NR	NR	6 months Compliant (≥ 4 hours/ night on 70% of nights): 16/22 (73%) P = NS RR vs C: 1.27 (0.81, 2.0) Minutes used: 271 (130) Effect size vs B: -0.16 (-0.43, 0.76)	6 months Compliant **B:** 15/21 (68%) **C:** 12/21 (57%) RR vs B: 1.02 (0.7, 1.48) Minutes used: **B:** 252 (100) **C:** 263 (112) P = NS[a] Effect size vs C: 0.07 (-0.53, 0.66)

Study Intervention (n) Control (n)	BMI		Blood Pressure		HbA1c		Time to Initiation of Therapy		Compliance/Adherence	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Pamidi 2012[35] Sleep specialist (n = 105) Non-sleep specialist (n = 298)	NR	NR	NR	NR	NR	NR	NR	A	Mean CPAP usage, min: 219 (152) Effect size: -0.35 (-0.57, -0.12) % days ≥ 4h CPAP usage: 46% CPAP use ≥ 4hours/ night on ≥ 70% of nights: 98/298 (32.9%) RR: 0.72 (0.55, 0.94)	CPAP usage: 279 (179) P = .005[a] % days ≥ 4h CPAP: 63% P = .004[a] CPAP use ≥ 4hours/ night on ≥ 70% of nights: 48/105 (45.7%) P = .01[a] Consulta-tion with sleep specialist significant predictor of CPAP adherence (1st 30 days of therapy[b])

Study Intervention (n) Control (n)	BMI		Blood Pressure		HbA1c		Time to Initiation of Therapy		Compliance/Adherence	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Lettieri 2011[33] Group 1: Primary care (n = 70) Group 2: Sleep specialist (n = 70) Group 3: Sleep specialist (n = 70)	NR	NR	NR	NR	NR	NR	NR	NR	Hours/night 4.7 (2) P = .98[a] %nights used 70.7% (26) P = .94[a] Effect size vs 2: -0.11 (-0.44, 0.22) Effect size vs 3: -0.07 (-0.4, 0.26) Use > 4 hours/ night for > 70% of nights 54.3% P = .84[a]	Hours/night **G2:** 4.7 (1.1) Effect size: 0 (-0.33, 0.33) **G3:** 4.8 Effect size: -0.05 (-0.4, 0.22) **G2:** 73.2% (18) **G3:** 72.4% (22) **G2:** 51.4% **G3**: 50%
Antic 2009[30] Nurse (n = 100) Usual care (n = 95)	NR	NR	NR	NR	NR	NR	No significant difference between groups in satisfaction with time waiting (P = .71[a]) Effect size: 0.06 (-0.24, 0.36) Patients receiving nurse-led care were more satisfied with their impression of wait time (P = .004[a]) Effect size: 0.46 (0.15, 0.76)		Hours/night 4.11 (SE 0.28) (n = 94)	Hours/night 4.56 (SE 0.30) (n = 81) MD:-0.45 (-1.26, 0.36)

Study Intervention (n) Control (n)	BMI		Blood Pressure		HbA1c		Time to Initiation of Therapy		Compliance/Adherence	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Palmer[34] 2004 Nurse (n = 68) Consultant Clinic (n = 71)	NR	NR	NR	NR	NR	NR	NR	NR	Daily hours of CPAP use Baseline (n = 71) 4.91 (2.85) P = .46[a] Follow-up (n = 63) 5.93 (2.67) Effect size: 0.11 (-0.24, 0.46) Change (n = 58) 0.66 (1.71) P = .5[a] P = .004 from baseline	Daily hours of CPAP use Baseline (n = 71) 5.24 (2.5) Follow-up (n = 63) 5.64 (2.54) P = .54[a] Change (n = 61) 0.45 (1.69) P = .041 from baseline
Scharf 2004[36] Primary care (n = 44) Usual care (n = 59)	NR	NR	NR	NR	NR	NR	Interval between PSG and CPAP study ≤1 m: 8.6% (3/35)	35.3% (18/51) P = .012[a]	Compliant[c] 3 m after onset of treatment all patients: 41% (18/44) RR: 0.8 (0.52, 1.24) Of patients accepting CPAP 62% (18/29) P = NS[a]	Compliant[c] 3 m after onset of treatment, all patients 51% (30/59) P = NS[a] Of patients accepting CPAP 68% (30/44) RR: 0.91 (0.64, 1.29)

[a]between groups;
[b]Mean adherence 58 min higher per day with sleep specialist consultation prior to initial PSG; after adjustment for age, race, BMI, medical insurance, AHI, ESS, CES-D, and education level
[c]compliant defined as use for at least 4h/night 5 nights per week, estimated by the patient over the prior month
MD = mean difference; NS = not statistically significant; NR = not reported; PSG = polysomnography; RR = risk ratio; SE = standard error

Table 6. Intermediate outcomes for KQ1, Continued

Study Intervention (n) Control (n)	Case Finding (describe)		Harms, Adverse Events (describe)		Harms, Adverse Events (describe)		Harms (Overdiagnosis, False Positives/Negatives)		Costs per Patient	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Chai-Coetzer 2013[31] Primary care (n = 81) Usual care (n = 74)	NR	NR	NR	NR	NR	NR	NR	NR	$1819.44	$3067.86
Chamorro 2013[32] Primary care pulmonologist and sleep specialist (n = 96)	Concordance between primary care pulmonologist and sleep specialist kappa = .74, P<.001		NR	NR	NR	NR	NR	NR	NR	NR
Andreu 2012[29] Nurse (n = 22) Sleep pulmonologist (n = 43)	NR	NR	Dryness (54%) Nasal congestion (40%) Abrasions (25%)		NR	NR	NR	NR	Cost per patient €590 (43) Effect size vs B: 8.1 (6.4, 10.1) Effect size vs C 0.75 (0.13, 1.37)	Cost per patient **B:** €849 (11) P<.001 vs A and C **C**: €644 (93) P<.05 vs A
Lettieri 2011[33] Group 1: Primary care (n = 70) Group 2: Sleep specialist (n = 70) Group 3: Sleep specialist (n = 70)	NR	NR	12.9% dis-continued therapy	**Group 2:** 8.6% dis-continued **Group 3**: 10% dis-continued P = .78[a]	NR	NR	NR	NR	NR	NR
Antic 2009[30] Nurse (n = 100) Usual care (n = 95)	NR	NR	NR	NR	NR	NR	NR	NR	Within-trial costs were significantly less with nurse-led care (A$1,111 per patient less)	

Study Intervention (n) Control (n)	Case Finding (describe)		Harms, Adverse Events (describe)		Harms, Adverse Events (describe)		Harms (Overdiagnosis, False Positives/Negatives)		Costs per Patient	
	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control	Non-sleep specialist	Control
Palmer 2004[34] Nurse (n = 68) Consultant Clinic (n = 71)	NR	NR	NR	NR	NR	NR	NR	NR	Total cost (to NHS) of nurse home visit: $83.62 (79.76) Cost to patient was set at $0, no accounting for time off work, *etc*	Total cost to NHS of clinic visit: $9.94 (6.38) Total cost to patient $37.81 (37.13)

[a]between groups
NHS = National Health Service (UK)

Table 7. Study Characteristics for KQ3

Author, year Study design Location / Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/ Exclusion Criteria	Patient Characteristics (expressed in means unless otherwise noted)	Risk of Bias
Bakker 2011[40] Crossover RCT Wellington, New Zealand Sleep clinic	APAP vs CPAP for treatment (n = 12) Outcomes assessed after 6 nights with 4-night washout period	Treatment APAP (5-20 cm H_2O) and CPAP with pressure set during manual titration	Inclusion: English-speaking, PAP naïve, morbidly obese (BMI ≥ 40kg/m^2), ≥ 18 years old, AHI ≥ 15/hour, manually titrated pressure ≥ 14 cmH_2O Exclusion: cardiac, respiratory, psychiatric, sleep co-morbidities (including central sleep apnea and those with irregular sleep patterns)	n = 12 Mean age: 45.9 (range 23-59) Male gender, %: 75 BMI: 49.9 (5.2) Obesity, %: 100 ESS: 17.4 (4.7) AHI (/hr): 75.8 (32.7) Mean O_2 desaturation: 8% (4.2)	**Sequence generation:** adequate **Allocation concealment:** not reported **Blinding:** unclear, blinded during data collection but not data entry, patient was blinded **Incomplete outcome data:** adequate **Selective outcome reporting:** adequate **Risk of Bias: Medium**
Lettieri 2011[33] Observational cohort study United States Community Based Hospital and Academic Sleep Center	**Group 1** (n = 70): unattended Type III home sleep study and home APAP titration (not included for KQ3 comparison) **Group 2** (n = 70): in-lab Type I attended sleep study and in-lab CPAP titration **Group 3** (n = 70): in-lab Type I sleep study, unattended home APAP titration 4-6 week follow-up	Titration All treated with CPAP, titration via home APAP or in-lab CPAP	Inclusion: diagnosed with OSA by HST or PSG (according to AASM guidelines), eligible for home sleep study program (≥ 2 high-risk features such as habitual snoring, daytime fatigue, nonrestorative sleep, weight gain, and witnessed apneas; no suspicion of concomitant sleep disorders and no significant underlying comorbidities), diagnosed OSAS defined as AHI > 5 with compatible symptoms Exclusion: not eligible for HST (cardiopulmonary disease, heart failure, CAD, previous cerebrovascular accident, poorly controlled asthma, moderate to severe COPD (FeV1 < 50%), supplemental oxygen requirement)	n = 140 (groups 2 and 3 only) Group 2: n = 70 Mean age: 47.1 (8) Male gender, %: 71.4 BMI: 30 (3.5) ESS: 14.1 (4.2) Fatigue: 6.7 (1.7) AHI (/hr): 23.1 (13) Group 3: n = 70 Mean age: 45.5 (5.4) Male gender, %: 68.6 BMI: 28.5 (3) ESS: 13.9 (4.4) Fatigue: 6.5 (1.4) AHI (/hr): 19.3 (9.4)	**Selection bias:** inadequate; unclear how patients in Groups 2 and 3 were selected to achieve same number as in Group 1; participants in all groups had to meet the same criteria for HST and APAP **Blinding of outcome assessment:** NR (adherence was objective measure) **Intention-to-treat analysis:** adequate after exclusions for Group 1 and selection of equal number for Groups 2 and 3 **Attrition bias:** adequate **Selective outcome reporting:** adequate **Risk of Bias: Medium**

Author, year Study design Location / Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/ Exclusion Criteria	Patient Characteristics (expressed in means unless otherwise noted)	Risk of Bias
Drummond 2010[43] RCT US VA Medical Center	Empiric APAP (n = 54 randomized, 42 completed) Usual care (n = 55 randomized, 44 completed) 1 month	Treatment APAP: cost-free auto-CPAP unit on day of randomization; returned to clinic at 1 month for assessment; remained on APAP and awaited in-lab PSG and CPAP titration; final assessment 1 month after PSG Usual care: 2nd assessment at 1 month after randomization; waited for in-lab PSG with CPAP titration; returned after 1 month of CPAP for assessment	Inclusion: consecutive patients referred for PSG; ≥ 2 categories of Berlin questionnaire positive Exclusion: age > 80; history of CHF; MI in past 6 months; COPD with FEV_1<60% predicted; stroke; alternative sleep diagnosis; prior diagnosis of OSA	N = 109 randomized, 86 completed protocol Mean age: 55 Male gender, %: 93 Race: African American: 32%, Caucasian: 68% BMI: 35.1 ESS: 14.4	**Sequence generation:** NR **Allocation concealment:** adequate **Blinding**: No **Incomplete outcome data:** 86 (79%) completed protocol; reported using intention-to-treat analysis with last observation carried forward **Selective outcome reporting:** No **Risk of bias**: **Medium**

Author, year Study design Location / Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/ Exclusion Criteria	Patient Characteristics (expressed in means unless otherwise noted)	Risk of Bias
McArdle 2010[38] RCT Australia Tertiary hospital sleep service	Manual titration (n = 83); in-lab CPAP titration with full PSG monitoring Home APAP (n = 86); used APAP overnight at home Lab APAP (n = 80); APAP titration in-lab without full PSG (outcomes not extracted; in-lab APAP was not comparison of interest) 4 weeks	Titration All received fixed CPAP at pressure determined by sleep specialist	Inclusion: symptoms of OSA (ESS ≥ 8, AHI ≥ 15 events/hr), age 17-85, living near sleep service and no previous treatment for OSA Exclusion: BMI > 45, significant lung or cardiac disease, neuromuscular disease, previous stroke, predominant central sleep apnea, periodic leg movements > 15/hr, severe medical illness or planned surgery, language impairment, or psychiatric illness	Manual: n = 83 Mean age: 50 (12) Male gender,%:75 BMI: 32.4 (5.7) Hypertension: 34% Diabetes: 11% ESS: 14.1 (4) AHI (/hr): 38 Time oxygen saturation < 90%, min: 5 Home APAP: n = 86 Mean age: 50 (12) Male gender, %: 59 (69) BMI: 32.2 (5.2) Hypertension: 29% Diabetes: 15% ESS: 13.8 (4.1) AHI (/hr): 38 Time oxygen saturation < 90%, min: 2	**Sequence generation:** adequate **Allocation concealment:** opaque envelopes **Blinding:** adequate **Incomplete outcomes reporting**: per protocol and ITT analysis, some outcomes not reported as ITT **Selective outcome reporting**: data not given for all outcomes **Risk of bias: Medium**

Author, year Study design Location / Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/ Exclusion Criteria	Patient Characteristics (expressed in means unless otherwise noted)	Risk of Bias
Vennelle 2010[61] RCT (crossover) UK, Sleep center	APAP vs CPAP (n = 192 randomized, 181 completed) Outcomes assessed after each 6-week treatment period No washout period so no data from first 2 weeks of each study arm were included in analysis	Treatment Same CPAP device with 2 modes: fixed-pressure mode (determined during overnight in-lab CPAP titration) and variable pressure mode	Inclusion: diagnosis of OSAHS; ESS ≥ 10 or history of troublesome sleepiness when driving, AHI ≥ 15 on PSG or ≥ 25 apneas and hypopneas per hour in limited sleep study, age 18-70, no previous CPAP use Exclusion: severe neurological deficit sufficient to compromise CPAP usability or understanding; significant comorbidity such as severe COPD, stroke, unstable diabetes, or active angina; coexisting narcolepsy or periodic limb movement syndrome; contraindications to CPAP use including recent pneumothorax	N = 200 randomized, 181 analyzed Mean age: 50 (10) Male gender, %: 77 BMI: 34.5 (7.8) ESS: 14 (3) Mean AHI (/hr): 33 (18) among n = 123 who had PSG	**Sequence generation**: adequate **Allocation concealment**: adequate **Blinding:** adequate (patients, staff involved in data acquisition or analysis) **Incomplete outcomes reporting**: 9.5% did not complete study (19/200) **Selective outcome reporting:** adequate **Risk of Bias: Low**
Damjanovic 2009[41] RCT (controlled parallel group study) Germany	Standard support and APAP (n = 25) Standard support CPAP (n = 25) Intensive support APAP (n = 25) Intensive support CPAP (n = 25) 3- and 9-month follow up	Treatment CPAP pressure was the pressure level with the lowest RDI during polysomnography Intensive support groups visited by specially trained members of authors' sleep lab at 1, 2, 4, 5, and 6 months to optimize treatment and provide support	Inclusion: newly diagnosed OSAS patients, AHI ≥ 15, with or without corresponding daytime symptoms Exclusion: any global respiratory failure, central sleep apnea, severe mental or psychological impairment	n = 100 randomized, 78 at 9 months Mean age: 57 (12) Male gender, %: 78 BMI: 31 (5) APAP: n = 50 ESS: 8.5 (0.8) AHI (/hr): 41.8 (3.5) CPAP: n = 50 ESS: 9.3 (0.7) AHI (/hr): 45.5 (3.6)	**Sequence generation**: NR **Allocation concealment**: NR **Blinding:** NR **Incomplete outcomes reporting**: 22% (22/100) with no follow-up at 9 months; difference between intensive and standard support groups **Selective outcomes reporting**: adequate **Risk of bias: Medium**

Author, year Study design Location / Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/ Exclusion Criteria	Patient Characteristics (expressed in means unless otherwise noted)	Risk of Bias
Galetke 2008[45] RCT (crossover) Germany, university-associated sleep laboratory	APAP vs CPAP (n = 20) Outcomes assessed after 8 weeks of a treatment with full in-lab sleep study after 16 weeks	Treatment Conventional CPAP at fixed pressure obtained during manual titration and APAP therapy (responds to snoring apneas/hypopneas and inspiratory flow limitation), range 4-15 cmH_2O One machine with 2 modes	Inclusion: OSAS newly diagnosed with AHI > 10/h, based on full in-laboratory PSG data and clinical symptoms Exclusion: COPD, CHF, acute neurological or psychiatric disorders, other major intrinsic sleep disorders, or malignant diseases	n = 20 Mean age: 55.5 Male gender, %: 80 BMI: 29.3 ESS: 10.3 (5.7) AHI (/hr): 32.9 (19.1) Arousals/hr: 17.6 (9.2) Snoring, n of epochs: 436.3 (209.6) SaO_2 min, %: 77.8 (8.4)	**Sequence generation**: NR **Allocation concealment**: NR **Blinding:** single blind (patients); data analysis either automated or done by technologists not involved in study **Incomplete Outcome**: adequate **Selective reporting**: adequate **Risk of bias: Medium**

Author, year Study design Location / Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/ Exclusion Criteria	Patient Characteristics (expressed in means unless otherwise noted)	Risk of Bias
Fietze 2007[44] RCT (titration done as crossover, but treatment randomized) Germany	APAP (n = 10) CPAP (n = 11) Outcomes assessed after 6 weeks of treatment	Treatment Titration done as a crossover over 2 nights, both APAP and CPAP titration done in-lab Treatment with second titration device was continued for 6 weeks; APAP set between 4 and 16 cmH_2O	Inclusion: suspected sleep apnea; if AHI > 10/hr in home cardiorespiratory polygraph and symptoms of excessive sleepiness or AHI > 20/hr patients had PSG in lab; patients included if AHI ≥ 10/hr; if AHI > 10/hr at home and AHI <10/hr in lab included if also had excessive sleepiness; BMI <40; age 35-70 Exclusion: other sleep disorders (including leg movements), acute cardiac, pulmonary or other internal medicine-related disorders, acute psychiatric or neurological disorders, or abuse of sleep-inducing agents or other drugs; suspected or confirmed central sleep apnea; previous treatment (*eg,* CPAP, oral devices, or uvulopalatopharyngoplasty)	n = 21 Mean age: 54.2 (11.7) Male gender, %: 95 BMI: 30.9 (5.7) APAP: n = 10 AHI: 43.3 (30.2) Sleep latency, min: 17.7 (13.6) Total sleep time, min: 355.7 (27.9) CPAP: n = 11 AHI: 40.4 (26.1) Sleep latency, min: 11.2 (6.4) Total sleep time, min: 379.5 (63.8)	**Sequence Generation**: NR **Allocation Concealment**: NR **Blinding**: NR **Incomplete Outcomes**: adequate **Selective Reporting**: adequate **Risk of Bias: Medium**

Author, year Study design Location / Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/ Exclusion Criteria	Patient Characteristics (expressed in means unless otherwise noted)	Risk of Bias
Meurice 2007[50] RCT France Sleep laboratories	**Group 1** (n = 17 randomized, 14 at 6 months): fixed CPAP **Group 2** (n = 17 randomized, 13 at 6 months): GK 418P APAP device **Group 3** (n = 17, 15 at 6 months): AutoSet device **Group 4** (n = 17, 12 at 6 months): PV10i device **Group 5** (n = 15, 11 at 6 months): Somnosmart 1 device 6 months	Treatment Fixed CPAP pressure manually determined during laboratory titration, APAPs all set in auto-adjust mode during the titration night All patients treated at home for 6 months with machine they used during titration night	Inclusion: naïve to nasal CPAP, no nasopharyngeal surgery, AHI > 30/hr or > 10 micro-arousals/hr Exclusion: > 20% of respiratory disturbances characterized as central events or taking sedative treatments	n = 83 at randomization, n = 65 at 6 months Mean Age: 56 (10) BMI: 30.8 (5.3) AHI (/hr): 52.3 (17.8)	**Sequence Generation:** adequate **Allocation Concealment:** adequate **Blinding:** unclear **Incomplete outcome reporting:** adequate, > 10% dropped out but balanced and reasons given **Selective outcome reporting**: adequate **Risk of Bias: Medium**

Author, year Study design Location / Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/ Exclusion Criteria	Patient Characteristics (expressed in means unless otherwise noted)	Risk of Bias
Mulgrew 2007[39] RCT Canada Tertiary referral sleep disorders program	In-lab CPAP titration (n = 35 assigned, 30 completed) Auto-CPAP titration (n = 33 assigned, 31 completed) 3-month follow-up	Titration For ambulatory APAP group CPAP set at 95%	Inclusion: adults referred from catchment area of sleep disorders program at University of British Colombia Hospital for assessment of suspected OSA who have clinical suspicion of moderate to severe OSA; high pretest probability of moderate to severe OSA, medically stable, not taking sedative medications Exclusion: pregnant, FEV1 or FVC <70%, known cause for daytime sleepiness, life-threatening comorbid illness, major psychiatric disorder, MVA attributable to hypersomnolence in preceding 5 years, previous treatment for OSA, contraindication for nasal CPAP, inability to provide informed consent	n = 68 CPAP: n = 35 Mean age: 52 (11) Male gender, %: 75 BMI: 38 (8) Median ESS: 14 (11-19IQR) Median RDI: 31 (21-47IQR) Median SACS: 30 (18-42IQR) Auto-CPAP: n = 33 Mean age: 55 (10) Male gender, %: 79 BMI: 39 (9) Median ESS: 14 (12-16IQR) Median RDI: 27 (17-57IQR) Median SACS: 32 (22-48)	**Sequence Generation:** inadequate, block randomized using large envelopes with folded cards inside **Allocation concealment:** inadequate **Blinding:** NR **Incomplete outcome reporting**: adequate **Selective outcome reporting:** adequate **Risk of Bias: Medium**

Author, year Study design Location / Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/ Exclusion Criteria	Patient Characteristics (expressed in means unless otherwise noted)	Risk of Bias
Nolan 2007[51] RCT (crossover) Ireland Respiratory sleep disorders unit	APAP (n = 29) CPAP (n = 29) (n = 34 enrolled, 29 completed) Outcomes assessed after each 8-week treatment period	Treatment CPAP therapy pressure calculated from overnight lab-based autotitration study (95% percentile) Variable pressure set between 4-20 cmH_2O	Inclusion: consecutive patients from sleep disorders unit with newly diagnosed mild to moderate OSAS (AHI ≥ 5 and <30) and compatible clinical features; awaiting a trial of CPAP therapy, ESS score ≥ 7 Exclusion: known cardiovascular disease other than hypertension, previous CPAP therapy, preexisting chronic airways disease, or previous upper airway surgery	n = 29 Mean age: 52.8 (8.3) Male gender, %: 90 BMI: 29.9 (4.7) Blood pressure, mmHg: 132/84 (23/13) Neck circumference (cm): 42 (2)	**Sequence generation**: not reported **Allocation concealment**: adequate **Blinding:** adequate (investigator blinded; patient partially blinded) **Incomplete outcomes**: 5/34 (15%) dropped out, 1 due to side effects, 4 lost to follow-up **Selective reporting**: adequate **Risk of Bias: Medium**
Patruno 2007[55] RCT Italy	Fixed level CPAP (n = 16) APAP (n = 15) Treated for 3 months	Treatment CPAP: pressure determined during titration APAP: pressure set to deliver levels from 4 to 15 cmH_2O Repeat sleep study at end of 3-month treatment	Inclusion: newly diagnosed OSA (AHI > 20/h and diurnal hypersomnolence [ESS score > 12]); free of diseases other than arterial hypertension; never treated for OSA Exclusion: taking treatments other than ACE inhibitors, calcium channel clockers, and diuretics	n = 31 (n = 40 were enrolled, 9 were excluded and not analyzed) Mean age: 48 Male gender, %: 81 BMI: 36.5 ESS: 15 (2.7) AHI (/hr): 46.5 (13.5) SaO_2, mean, %: 90 SaO_2, nadir, %: 72 Hypertensive: n = 17 BP, systolic mmHg: 143 (10) BP, diastolic mmHg: 87 (5) Glucose, mg/dL: 103.9 (6.8)	**Sequence generation**: NR **Allocation Concealment**: NR **Blinding:** compliance recorded by the computer; other outcomes NR **Incomplete Outcome**: not ITT, > 10% attrition with reasons, balance NR **Selective Reporting**: adequate **Risk of Bias: Medium**

Author, year Study design Location / Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/ Exclusion Criteria	Patient Characteristics (expressed in means unless otherwise noted)	Risk of Bias
Richard 2007[58] Retrospective cohort Netherlands	CPAP (n = 78) APAP (n = 96) Follow-up: 2 months to 8 years, unclear when data collected	Treatment	Inclusion: all patients with OSAS (defined as AHI > 5 in overnight PSG accompanied by daytime symptoms) offered nCPAP between Jan 1997 and July 2005; if AHI > 30 – offered nCPAP as 1st treatment; if AHI<30 – offered alternative treatments (oral device, surgery) Exclusion: none reported	n = 174 Mean age: 56.7 Male gender, %: 80.5 BMI: 33 CPAP: n = 78 AHI (/hr): 47.2 (22.3) ESS: 5.6 (4.5) APAP: n = 96 AHI (/hr): 52.0 (23.1) ESS: 7.1 (5.1)	**Selection bias:** unclear **Blinding of outcome assessment:** NR **Intention-to-treat analysis:** inadequate **Attrition bias:** adequate **Selective outcome reporting:** adequate **Risk of Bias: Medium**
Nolan 2006[52] quasi RCT crossover (all patients on CPAP then random crossover assignment of 3 APAP devices) Ireland University sleep disorders center	CPAP vs 3 APAP devices (n = 27) Baseline values based on median of 53 months of CPAP; Outcomes assessed after 4-week home trial with each of 3 APAP devices	Treatment APAP a) Autoset Spirit (reviews shape of inspiratory flow curve on breath-by-breath basis) b) Breas PV 10i (creates model of patient's breathing signal and compares to template to set device) c) RemStar Auto (compares inspiratory flow shape to rolling patient database) CPAP Different devices, used for 37-85 months prior to start of study	Inclusion: attending Respiratory Sleep Disorders Unit, confirmed diagnosis of OSAS, already established on fixed-pressure CPAP with nasal mask and device that downloaded time-coded compliance data Exclusion: malignant or psychiatric disease; on regular narcotics, sedatives, or psychoactive medications	n = 27 Mean age: 53 Male gender, %: 93 BMI: 36.2 Diagnostic AHI(/h): 48 (29-76) Diagnostic ESS: 15 (9-19)	**Sequence generation**: NR (Note: sequence generation only for 3 APAP devices) **Allocation concealment**: adequate **Blinding:** investigator performing analysis and person assigning APAP devices; patients were not informed about APAP technologies but were told they were newer treatment machines **Incomplete outcome data**: No patients lost to follow-up **Selective outcome reporting**: No **Risk of bias: Low**

Author, year Study design Location / Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/ Exclusion Criteria	Patient Characteristics (expressed in means unless otherwise noted)	Risk of Bias
Nussbaumer 2006[54] RCT (crossover) Switzerland Outpatient clinic	APAP vs CPAP (n = 34 randomized, 30 completers) 1 month	Treatment APAP pressure ranged from 5 to 15 cmH_20, CPAP pressure set at 90^{th} percentile	Inclusion: consecutive patients with excess sleepiness and AHI > 10 /hr Exclusion: CHF, chronic rhinitis, other sleep disorders	n = 34, data for 30 completers Mean age: 49 (SE2) Male gender, %: 90 BMI: 31.3 (SE.6) ESS: 12.7 (0.6) AHI (/hr): 41.1 (3.6)	**Sequence generation**: adequate **Allocation concealment**: adequate **Blinding:** double-blind (patients and attending physicians) **Incomplete outcome data**: 4 (12%) did not complete protocol **Selective outcome reporting:** No **Risk of bias: Low**
West 2006[62] RCT UK Outpatient sleep clinic	**Group 1** (n = 31): Auto-titration pressure **Group 2** (n = 33): Fixed pressure **Group 3** (n = 34): Fixed pressure 6 months	Treatment Autotitration pressure (Group 1), fixed pressure determined by the 95% from 1 week of autotitration (Group 2), and fixed pressure determined by algorithm based on neck size and dip rate (Group 3)	Inclusion: aged 18-75 years with excessive daytime sleepiness (ESS > 9) and proven OSA on 1 night respiratory PSG; > 10 dips /hr in SaO_2 of > 4% confirmed as being caused by upper airway obstruction eligible for inclusion; no exclusion on basis of other co-morbidities Exclusion: respiratory failure requiring urgent treatment	n = 98, 86 at 6 months Mean age: 46 Male gender, %: 85 Maintenance of Wakefulness test, mins: 18 4% oxygen saturation dips, events/h: 34 Mean BP mm Hg: 96	**Sequence generation**: adequate **Allocation concealment**: unclear **Blinding:** patients and the outcomes assessors **Incomplete outcome data**: Unclear if the analyses include all randomized. Reasons for dropout were reported. **Selective outcome reporting**: no **Risk of bias: Medium**

Author, year Study design Location / Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/ Exclusion Criteria	Patient Characteristics (expressed in means unless otherwise noted)	Risk of Bias
Hukins 2004[46] RCT (crossover) Australia Hospital's sleep disorders center	APAP (n = 32) CPAP (n = 23) 2-month treatment period with outcomes assessed last 30 days	Treatment Fixed-pressure (CPAP) or autotitrating (APAP) mode of the AutoSet T (default pressure 4-20 cmH_2O) Received each treatment for 2 months, outcomes reported for last 30 days of those 2 months to allow for washout period	Inclusion: AHI ≥ 5 in association with hypersomnolence, optimal CPAP pressure determined by overnight pressure determination PSG, no previous home use of CPAP, and informed consent Exclusion: significant comorbidity (unstable ischemic heart disease, neuromuscular disease, kyphoscoliosis, or severe COPD), significant complication (hypercapnic respiratory failure or right heart failure), presence of non-obstructive sleep apnea, or inability to use masks compatible with Autoset T	n = 55 APAP: n = 32 Mean age: 51 (11.9) Male gender, %: 84 BMI: 35.8 (6.7) ESS: 13 (5) Diagnostic AHI: 59.7 (30.1) CPAP: n = 23 Mean age: 49.3 (12.5) Male gender, %: 91% BMI: 34.3 (6.3) ESS: 11.8 (5.3) Diagnostic AHI: 50.2 (24.9)	**Sequence generation**: adequate (shuffled sealed envelopes) **Allocation concealment**: adequate **Blinding:** attempted to blind patients (used same machine for APAP and CPAP) **Incomplete Outcome**: no, more than 10% attrition, no reasons given **Selective reporting**: unclear **Risk of bias: Medium**
Hussain 2004[47] RCT (crossover) Canada Unclear	Fixed CPAP vs Autotitrating CPAP (n = 10) 4 weeks separated by a 2-week washout period	Treatment APAP set between 3-20 cm H_2O, fixed CPAP pressure determined by overnight titration	Inclusion: CPAP-naïve patients with symptomatic OSAH (AHI > 15/h) Exclusion: none reported	n = 10 Mean age: 44.9 (9.7) Male gender, %: 90 BMI: 35.9 (12.9) ESS: 11.1 (6.4) AHI (/hr): 47.2 (35.6) Snoring: 100% Unrefreshing sleep: 80% Witnessed apnea: 80% Excessive daytime sleepiness: 70% Arousal index: 17.3 (17.7) Desaturation index: 53 (36)	**Sequence Generation:** not reported **Allocation concealment**: not reported **Blinding:** only patients blinded, compliance collected by machine **Incomplete outcome reporting:** adequate **Selective outcome reporting**: adequate **Risk of Bias: Medium**

Author, year Study design Location / Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/ Exclusion Criteria	Patient Characteristics (expressed in means unless otherwise noted)	Risk of Bias
Marrone 2004[48] RCT (crossover) Italy University sleep center	Fixed CPAP vs APAP (n = 22) 1 month	Treatment APAP pressure set to range between 4-18 cmH_20, CPAP level determined during PSG with APAP titration in lab	Inclusion: patients referred for suspected OSAS and consecutive subjects with AHI ≥ 30 and no overt cardiopulmonary disease were requested to participated in study (all accepted) Exclusion: none reported	n = 22 Mean age: 53.4 Male gender, %: 95 BMI: 32.6 ESS: 16.3 (5)	**Sequence generation**: NR **Allocation concealment**: NR **Blinding:** patients **Incomplete outcome data**: NR **Selective outcome reporting**: No **Risk of bias: Medium**
Masa 2004[37] RCT Spain Sleep centers	**Standard Titration** (n = 126 randomized, 107 analyzed): in-lab CPAP titration **AutoAdjusted Titration** (n = 119 randomized, 106 analyzed): at home APAP titration **Predicted Formula Titration** (n = 115 randomized, 102 analyzed): used a formula to calculate optimal pressure (did not extract outcomes for this group) 12 weeks	Titration Standard vs home titration	Inclusion: requiring CPAP treatment (AHI ≥ 30, ESS ≥ 12), age 18-70 Exclusion: psychophysical incapacity to perform questionnaires; chronic disease; drug or alcohol addiction; Cheyne-Stokes syndrome; life-threatening SAHS; previous uvulopalatopharyngoplasty; absence of a partner at home; important chronic nasal obstruction; lack of skill in adjusting nasal mask in daytime CPAP trial; refusal to participate	Standard: n = 107 Mean age: 51 (9.1) Male gender, %: 86.9 BMI: 33.6 (8.4) HTN: 55.4% Sleep, hr/nt: 6.9 (1.1) Habitual snoring: 90.7% Apneas observed: 62.6% Nocturia: 23.4% Restlessness: 47.7% Morning headache: 14% AutoAdjusted: n = 106 Mean age: 52.2 (10.4) Male gender, %: 89.6 BMI: 33.1 (6.3) HTN: 57.4% Sleep, hr/nt:7.0 (1.5) Habitual snoring: 85.8% Apneas observed: 58.5% Nocturia: 31.1% Restlessness: 43.4% Morning headache: 12.3%	**Sequence Generation:** not reported **Allocation concealment**: not reported **Blinding:** not reported **Incomplete outcome reporting:** adequate, > 10% dropout but reasons given and not significantly uneven by groups **Selective Outcome reporting:** adequate **Risk of bias: Medium**

Author, year Study design Location / Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/ Exclusion Criteria	Patient Characteristics (expressed in means unless otherwise noted)	Risk of Bias
Noseda 2004[53] RCT (crossover) Brussels, Belgium Hospital	Auto CPAP vs CPAP for treatment (n = 27 randomized 24 evaluable) 8 weeks	Treatment Fixed CPAP set at the pressure judged to be effective during the titration night at the sleep laboratory, in the auto CPAP mode the pressure was set between 4-14 cm H_2O	Inclusion: high pressure variability during 14 day run-in period on APAP (VI > 2.75cm H_2O), AHI > 20/hr and a microarousal index (MAI) > 30 Exclusion: previous treatment with CPAP, central sleep apnea or Cheyne-Stokes respiration, major facial or pharyngeal anatomic abnormalities likely to require surgery, night or rotating shift work, severe chronic heart failure or COPD, seizure disorder, mental retardation, sedative, hypnotic or antidepressant therapy, previous uvulopalatopharyngoplasy, prolonged hypoventilation during REM sleep	n = 27 Mean Age: 49 (10) BMI: 32.3 (4.9) ESS: 10.7 (2.4) AHI (/hr): 50.9 (25.2) AI: 24.6 (22.6) MAI: 43 (12.9)	**Sequence Generation:** unclear, a randomization table was used **Allocation concealment:** unclear **Blinding:** single blind **Incomplete outcome reporting**: > 10% dropped out but balanced **Selective outcome reporting**: adequate **Risk of Bias: Medium**
Massie 2003[49] RCT (crossover) Multi-site Unclear	CPAP vs APAP as treatment (n = 46 randomized, 44 completed) 6 weeks	Treatment CPAP: fixed pressure as determined by board-certified sleep specialist or equivalent (by AASM standards) APAP: pressure ranged between 4-20 cm H_2O	Inclusion: need for CPAP pressure > 10cm, symptomatic OSAHS with AHI ≥ 15, age 18-65 Exclusion: preexisting lung disease, awake resting SaO_2 <90%, or ≥ 10 central apnea hypopnea events per hour, taking meds known to significantly interfere with sleep or respiration	n = 46 randomized, 44 completed Mean age: 49 (10) Male gender, %: 82 BMI: 32 (4)	**Sequence generation:** not reported **Allocation concealment**: not reported **Blinding**: not reported **Incomplete outcome reporting:** adequate **Selective outcome reporting**: adequate **Risk of bias: Medium**

Author, year Study design Location / Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/ Exclusion Criteria	Patient Characteristics (expressed in means unless otherwise noted)	Risk of Bias
Planès 2003[56] RCT France Four sleep laboratories	Auto-nCPAP, initiated at home (n = 16) Conventional nCPAP, initiated in sleep lab with titrating PSG (n = 14) 2 months	Treatment nCPAP, conventional and auto APAP home pressure set to 2 cmH_2O above to 4 cmH_2O below max pressure delivered by device during 1st week (at least 15 hours) of use	Inclusion: severe obstructive OSAS with AHI ≥ 30 events/hour and obstructive events > 80% of total events, clinical indications for nCPAP according to American Thoracic Society recommendations Exclusion: none reported	N = 35 recruited Mean age: 54.3 Male gender, %: 77 BMI: 32.4 ESS: 14.8 Hypertension: n = 7 No history of nCPAP or surgery for snoring No comorbidities noted n = 30 with outcomes Conventional: AHI (/hr):61.0 SaO_2: 12.7 Auto: AHI (/hr):57.5 SaO_2: 24.9	**Sequence generation:** NR **Allocation concealment**: NR **Blinding**: No **Incomplete Outcome Reporting**: 5/35 didn't complete treatment (unable to tolerate nCPAP) **Selective outcome reporting**: No *patients who didn't tolerate their assigned treatment were allowed to switch interventions… only one did auto to conventional **Risk of bias: Medium**
Senn 2003[59] RCT (cross-over) Switzerland	AutoAdjust LT vs AutoSet T vs Fixed-Pressure CPAP mode (n = 29) 2-week adaptation period with either APAP device then 1 month with each in random order, outcomes assessed at end of each month	Treatment AutoAdjust LT responds to apnea-hypopnea and snoring, AutoSet T responds to apnea-hypopnea, snoring and changes in inspiration flow contour, fixed-pressure CPAP mode is either APAP set in fixed mode	Inclusion: OSAS based on complaints of excessive sleepiness, snoring, and apnea-hypopnea index > 10/hr Exclusion: not naïve to CPAP therapy	n = 31 recruited, 29 completed Mean age: 53 Male gender, %: 79 BMI: 33.3 ESS: 14.2 (0.7) AHI (/hr): 45.8 (4.2) Time with SaO_2 <90%, % time in bed: 12.6 (3.4)	**Sequence generation**: NR **Allocation concealment**: NR **Blinding**:single-blind (patients blinded to study purpose and treatment modes) **Incomplete outcome reporting**: adequate **Selective outcome reporting**: adequate **Risk of bias: Medium**

Author, year Study design Location / Setting	Intervention Groups Follow-up	Aim of Study Treatments Received	Inclusion/ Exclusion Criteria	Patient Characteristics (expressed in means unless otherwise noted)	Risk of Bias
Randerath 2001[57] RCT (crossover) Germany University sleep laboratory	APAP vs CPAP (n = 52 randomized, 47 completed study) Outcomes assessed after each 6-week treatment period	Treatment APAP using forced oscillation set to between 4 and 18 cmH_2O CPAP using the pressure titrated during diagnostic polysomnography	Inclusion: referred to a university sleep laboratory by "pneumologists" and general practitioners, OSA diagnosed (AHI ≥ 10) from PSG, underwent basic lung function examination Exclusion: 1 patient excluded after bronchial carcinoma diagnosed	n = 52 enrolled, 47 completed Mean age: 54.7 (10.1) Male gender, %: 87 BMI: 32.4 (5.8) AHI (/hr): 35.1 (26) Snoring(/hr): 49 (36) Total number of arousals (/hr): 34.0 (21.7)	**Sequence generation**: not reported **Allocation Concealment**: not reported **Blinding:** adequate (patients, physicians, and technicians) **Incomplete outcome reporting**: adequate (5/52 [10%] quit study **Selective outcome reporting**: adequate **Risk of Bias: Medium**
D'Ortho 2000[42] RCT (cross-over) France	APAP vs CPAP (n = 25) Outcomes assessed after each 2-month treatment period	Treatment Constant CPAP or auto-CPAP mode of "REM + Auto" apparatus CPAP titration done in lab to identify effective pressure for constant CPAP APAP range set at 6-16 cmH_2O	Inclusion: clinical suspicion of OSAS confirmed by PSG, AHI > 10/h of sleep with obstructive events > 80% of total events and clinical indication for CPAP treatment according to American Thoracic Society recommendations Exclusion: restless legs, cardiac failure, cerebrovascular disease, or lung disease	n = 25 Mean age: 57 (11) Male gender, %: 88 BMI: 32 (5) ESS: 12.7 (5.3) Sleep onset latency, min:37 Total sleep time, min: 375 (65) Arousal/awakening index, events(/h): 45.6 (25.8) AHI (/hr): 57.8 (5.8) Mean SaO_2, %: 93 (3.0)	**Sequence generation**: NR **Allocation concealment**: NR **Blinding**: single blinded (patients) **Incomplete outcome reporting:** adequate **Selective outcome reporting:** adequate **Risk of Bias: Medium**
Teschler 2000[60] RCT (crossover) Australia	APAP vs CPAP as treatment (n = 10) Outcomes assessed and PSG every 2 months	Treatment APAP was CPAP device operated in auto mode CPAP pressure determined during manual titration night following diagnostic night	Inclusion: newly diagnosed moderate to severe OSAS (AHI > 20/hour); residence within 50 km of clinic Exclusion: primary diagnosis of asthma, emphysema, allergic rhinitis, or cardiac failure	n = 10 Mean age: 52.2 (2) Male gender, %: 100 BMI: 33.8 (1.3) AHI (/hr): 52.9	**Sequence generation**: unclear **Allocation concealment**: unclear **Blinding:** double-blind (patients, staff) **Incomplete outcome data**: No reported missing data **Selective outcome reporting**: No **Risk of bias: Medium**

CPAP = continuous positive airway pressure; CVA = cerebrovascular accident; ESS = Epworth Sleepiness Scale (cut-off score of 8 or more suggests the presence of at least mild daytime sleepiness); MI = myocardial infarction; ODI = oxygen desaturation index; OSA = obstructive sleep apnea; OSAH = obstructive sleep apnea-hypopnea; OSAHS = obstructive sleep apnea-hypopnea syndrome; OSAS = obstructive sleep apnea syndrome; PSG = polysomnogram; RDI = respiratory disturbance index; AASM = American Academy of Sleep Medicine; VI = variability index; MAI = micro arousal index

Table 8. Clinical Outcomes for KQ3

Study Intervention (n) Control (n)	All-cause Mortality % (n/N)		Resource Utilization (hospitalization, etc)		Access to Care		MID, Sleep Symptom Scores		MID, Urinary Symptom Scores	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Drummond 2010[43] APAP (n = 54) Usual care (n = 55)	No deaths		Hospitalized for chest pain 5/43 (12%)	4/44 (9%)	NR	NR	NR	NR	NR	NR
McArdle 2010[38] Manual CPAP titration (n = 83) Home APAP titration (n = 86) -all outcomes per protocol NOT ITT unless specified	NR	NR	Humidifiers/ pt, 4 weeks 0.65 (0.61) Chin straps, 4 weeks 0.26 (0.47) Staff time/pt (min) Technologist:, titration morning 14 (9.1) Physician, titration study reporting 12.7 (4.9)	0.49 (0.53) ES: -0.3 (-0.02, -0.58) P = NS 0.63 (0.66) ES: -0.65 (-0.96, -0.34) P = .001 10.1 (6.8) ES: 0.48 (0.18, 0.79) P = .01 1.3 (4.5) ES: 2.42 (2.02, 2.80) P<.001 All other measures of staff time/pt: P = NS	NR	NR	NR	NR	NR	NR
Vennelle 2010[61] APAP and CPAP (n = 192 randomized, 181 analyzed)	One person died for reasons unrelated to the trial		Patients sought help from sleep center 13 times	25 times (P = .70)	NR	NR	NR	NR	NR	NR

Study Intervention (n) Control (n)	All-cause Mortality % (n/N)		Resource Utilization (hospitalization, etc)		Access to Care		MID, Sleep Symptom Scores		MID, Urinary Symptom Scores	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Richard 2007 APAP (n = 96) CPAP (n = 78)	5/232 patients who returned questionnaires died before post-treatment evaluation (group not reported)		NR	NR	NR	NR	NR	NR	NR	NR
West 2006[62] Group 1, APAP (n = 31) Group 2, 1 wk titration CPAP (n = 33) Group 3, algorithm CPAP (n-34)	NR	NR	There was no difference between the groups in terms of the number of extra calls or extra visits made to the sleep nurses because of CPAP problems		NR	NR	NR	NR	NR	NR
Hukins 2004[46] APAP and CPAP (n = 55)	NR	NR	Unplanned contacts: 21 (15 clinic, 6 phone) Total duration of unplanned contacts: 440 minutes	18 (11 clinic, 7 phone) P = .73 455 minutes P = .56	NR	NR	NR	NR	NR	NR
Senn 2003[59] Auto Adjust and AutoSet and Fixed (n = 29)	NR	NR	NR	NR	NR	NR	ESS: clinically relevant change defined as change by 2 points All treatment modalities improved scores by > 5 points.		NR	NR

[a]APAP: a) RemStar Auto, b) Autoset Spirit, c) Breas Pv 10i

ES = effect size; ESS = Epworth Sleepiness Scale; NR = not reported

Table 9. Clinical Outcomes for KQ3, Continued

Study Intervention (n) Control (n)	Quality of Lfe		Patient Satisfaction		Remission		Cognitive Symptoms		Other (describe)	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Bakker 2011[40] CPAP and APAP (n = 12)	NR	NR	6 preferred APAP, 3 preferred CPAP, 3 had no preference		NR	NR	NR	NR	NR	NR
McArdle 2010[38] Manual CPAP titration (n = 83) Home APAP titration (n = 86) - all outcomes per protocol NOT ITT unless specified	SF-36 Physical Baseline: 58 (median) (n = 62) Week 4: 73 (n = 59) Change: 7.8 (18.6) (n = 58) SF-36 Mental Baseline: 53 (median) (n = 62) Week 4: 70 (n = 59) Change: 11.4 (15.0) (n = 58)	SF-36 Physical Baseline: 57 (median) (n = 62) Week 4: 66 (n = 60) P = NS between groups Change: 7.5 (13.5) (n = 60) ES: 0.02 (-0.34, 0.38) SF-36 Mental Baseline: 54 (median) (n = 62) Week 4: 68 (n = 60) P = NS between groups Change: 8.4 (14.2) (n = 60) ES: 0.21 (-0.15, 0.57)	NR	NR	NR	NR	Trails A, sec Baseline: 28 (median) (n = 60) Week 4: 26 (n = 60) Trails B, sec Baseline: 74 (median) (n = 59) Week 4: 73 (n = 58)	Trails A, sec Baseline: 28 (median) (n = 62) Week 4: 26 (n = 61) P = NS between groups Trails B, sec Baseline: 71 (median) (N = 62) Week 4: 73 (n = 61) P = NS between groups	NR	NR

Study Intervention (n) Control (n)	Quality of Lfe		Patient Satisfaction		Remission		Cognitive Symptoms		Other (describe)	
	APAP	**CPAP**	**APAP**	**CPAP**	**APAP**	**CPAP**	**APAP**	**CPAP**	**APAP**	**CPAP**
Vennelle 2010[61] APAP and CPAP (n = 192, n = 181 analyzed)	SF-36 58 (SEM 0.1) P = NS difference between groups for any SF-36 components	SF-36 58 (SEM 0.1) P = .9	Preferred by 69/181 (38%)	Preferred by 72/181 (40%) P = NS 40/181 (22%) had no preference Significant order effect (P = .009)	NR	NR	NR	NR	NR	NR
Galetke 2008[45] APAP and CPAP (n = 20)	NR	NR	Preferred by 13/20 (65%)	Preferred by 7/20 (35%) P<.01	NR	NR	NR	NR	NR	NR
Fietze 2007[44] APAP n = 10 CPAP n = 11	SF-36 Psychic (Mental Health): Baseline: 50.7 (6.5) 6-week: 52.3 (9.1) Bodily (Physical Health): Baseline: 46.4 (11.8) 6-week: 49 (10.2) Did not differ between CPAP and APAP groups		NR	NR	NR	NR	NR	NR	NR	NR
Meurice 2007[50] At 6 months Group 1, n = 14 Group 2, n = 13 Group 3, n = 15 Group 4, n = 12 Group 5, n = 11	SF36 (emotional) **Group 2** initial: 49.5 (8.3) 6-month: 46 (12.9) **Group 3** initial: 45.7 (7.9) 6-month: 46.2 (13.3) **Group 4**	SF36 (emotional) **Group 1** initial: 43.1 (9.4) 6 month: 47.3 (8.7) P = NS from baseline	NR	NR	NR	NR	NR	NR	NR	NR

Study Intervention (n) Control (n)	Quality of Lfe		Patient Satisfaction		Remission		Cognitive Symptoms		Other (describe)	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
	initial: 43.3 (12.3) 6-month: 50.8 (7.1) **Group 5** initial: 47.5 (7) 6-month: 42.5 (10.5) All P = NS from baseline SF36 (physical) **Group 2** initial: 42.7 (13.8) 6-month: 50.5 (7.7) **Group 3** initial: 48.6 (5.3) 6-month: 47.8 (8.7) **Group 4** initial: 46.2 (7.8) 6-month: 48.9 (6.3) **Group 5** initial: 46.7 (8.7) 6-month: 48.8 (7.7) All P = NS from baseline	SF36 (physical) **Group 1** initial: 45.6 (8.6) 6 month: 47.5 (9) P = NS from baseline								

Study Intervention (n) Control (n)	Quality of Lfe		Patient Satisfaction		Remission		Cognitive Symptoms		Other (describe)	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Mulgrew 2007[39] CPAP titration (n = 35) APAP titration (n = 33)	SAQLI Median (IQR) Baseline: 3.5 (2.8, 4.1) 3 months: 5.8 (4.9, 6.3)	Baseline: 2.8 (2.1, 4.2) 3 months: 5.5 (4.8, 6.2) Difference at 3 months: -0.19 (95% CI -0.7, 0.3), P = .41	6% would have preferred lab CPAP All patients expressed overall satisfaction	62% would have preferred home manage-ment	NR	NR	NR	NR	NR	NR
Nolan 2007[51] APAP and CPAP (n = 29)	NR	NR	Preferred by 13/29 (45%)	Preferred by 13/29 (45%) 3/29 (10%) did not express a preference Observed order effect – preferred machine received for first leg of trial	NR	NR	NR	NR	NR	NR
Richard 2007 APAP (n = 96) CPAP (n = 78)	NR	NR	VAS 10 point scale 7.5 (2.3) (n = 95)	VAS 10 point scale 7.5 (1.9) (n = 76) P = .88	NR	NR	NR	NR	NR	NR
Nolan 2006 CPAP and 3APAPs (n = 27)[a]	SF-36 No significant differences between 3 APAP devices or between APAP devices and CPAP		14/27 (52%) preferred APAP 13/27 (48%) preferred CPAP Preferred RemStar Auto: 6/14 (43%) Preferred Autoset Spirit: 5/14 (36%) Preferred Breas PV 10i: 3/14 (21%)		NR	NR	NR	NR	NR	NR

Study Intervention (n) Control (n)	Quality of Lfe		Patient Satisfaction		Remission		Cognitive Symptoms		Other (describe)	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Nussbaumer 2006[54] APAP and CPAP (n = 30)	SF-36 Physical Baseline[d] 82 (SE 4) At 1 month 84 (SE 4) Mental Baseline[d] 65 (SE 4) At 1 month 76 (SE 3)	SF-36 Physical At 1 month 85 (SE 4) Mental At 1 month 73 (SE 3) All P = NS vs baseline and between groups	Preferred APAP: 26/30 (87%)	Preferred CPAP: 4/30 (13%) P<.001	NR	NR	NR	NR	NR	NR

Study Intervention (n) Control (n)	Quality of Lfe		Patient Satisfaction		Remission		Cognitive Symptoms		Other (describe)	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
West 2006[62] Group 1, APAP (n = 31) Group 2, 1 wk titration CPAP (n = 33) Group 3, algorithm CPAP (n-34)	Data reported as median (5th/95th%) *SAQLI* Baseline 3.9 (1.7/6.0) P = .4 Change, 6 m 1.6 (-4.8/4.3) P = .7 P<.05 baseline *SF-36 MC* Pre CPAP 57.2 (19.8/87.5) P = .9 6 m: 79.3 (30.5/94.1) P = .9 P<.05 baseline *SF-36 PC* Pre CPAP 62.5 (17.2/93.2) P = .6 6 m: 78.8 (20/96.2) P = .5	Data reported as median (5th/95th%) *SAQLI* **Group 2:** Baseline 3.1 (1.5/5.8) Change, 6 m 1.5 (-5.6/4.7) **Group 3**: Baseline 3.5 (1.9/6.1) Change, 6 m 1.4 (-5.2/3.4) *SF-36 MC* **Group 2:** Pre CPAP 56.8 (26/89.4) 6 m: 81.5 (27.8/95) **Group 3:** Pre CPAP 56.6 (16.6/88.7) 6 m: 82.7 (35.4/95.4) *SF-36 PC* **Group 2:** Pre CPAP 65.7 (25.2/90) 6 m: 85.2 (26.6/95.3) **Group 3:** Pre CPAP 62.6 (25.9/92.5) 6 m: 83.3 (40.9/98.3)	NR	NR	NR	NR	NR	NR	NR	NR

Study Intervention (n) Control (n)	Quality of Lfe		Patient Satisfaction		Remission		Cognitive Symptoms		Other (describe)	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Hukins 2004[46] APAP and CPAP (n = 55)	SF-36, significant improvements in both treatment modes in the Role Physical and Vitality domains all P<.05), but no difference between groups; other domains P = NS		Subjective ease of CPAP use 7.15 (2.41) Attitude to CPAP 7.33 (2.05)	Ease of use 6.84 (2.54) P = .47 Attitude 9.91 (2.02) P = .20 Both used VAS 0-10	NR	NR	NR	NR	NR	NR
Hussain 2004[47] CPAP and APAP (n = 10)	NR	NR	Preferred by 10% of patients Patients reported similar satisfaction with therapy	Preferred by 60% of patients P = .06	NR	NR	NR	NR	NR	NR
Marrone 2004[48] APAP and CPAP (n = 22)	NR	NR	Preferred APAP: 14/22 (64%)	Preferred CPAP: 4/22 (18%) No preference: 4/22 (18%)	NR	NR	NR	NR	NR	NR

Study Intervention (n) Control (n)	Quality of Lfe		Patient Satisfaction		Remission		Cognitive Symptoms		Other (describe)	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Masa 2004[37] Standard titration (n = 107) Autoadjusted titration (n = 106)	SF 36 Physical Pre: 45.9 (8.6) Post: 47.3 (7.8) Change at 3 months: -1.4 (7.7) SF 36 Mental Pre: 47.5 (10.4) Post: 51.8 (9.2) Change at 3 months:-4.0 (10.8)	SF 36 Physica Pre: 44.3 (8.7) Post: 48.6 (7.3) Change at 3 months: -4.3 (6.9) P<.01 between groups SF 36 Mental Pre: 45.6 (12.2) Post: 49.4 (10.4) Change at 3 months: -3.9 (10.0) P = NS between groups	NR	NR	NR	NR	NR	NR	NR	NR
Noseda 2004[53] CPAP and APAP (n = 24)	NR	NR	Preferred by 16 patients	Preferred by 8 patients	NR	NR	NR	NR	NR	NR
Massie 2003[49] CPAP and APAP (n = 44)	SF-36 Vitality: 65 (20) SF Mental health: 80 (14) P<.05	SF-36 Vitality: 58 (23) P<.05 SF Mental Health: 75 (18) P > .07 for all other domains	NR	NR	NR	NR	NR	NR	NR	NR
Planès 2003[56] Auto (n = 16) Conventional (n = 14)	NR	NR	NR	NR	NR	NR	NR	NR	Tolerance score: 22.9 (5.8)	Tolerance score: 18.8 (10.5) P = NS

Study Intervention (n) Control (n)	Quality of Lfe		Patient Satisfaction		Remission		Cognitive Symptoms		Other (describe)	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Senn 2003[59] Auto Adjust and AutoSet and Fixed (n = 29)	SF-36 health transition, vitality, social functioning, and mental component summary scores were significantly improved by all treatments, no significant between group differences SF-36 vitality reached clinically relevant treatment effect (change of at least 10 points) for all treatments Effect sizes AutoSet: 0.63 AutoAdjust: 0.65 Fixed: 0.78		Treatment preference 21/29 (72%) no preference 4/29 (14%) preferred an auto mode 4/29 (14%) preferred a fixed mode 17/29 (59%) preferred one auto device brand 11/29 (38%) preferred the other brand 1/29 (3%) had no preference		NR	NR	NR	NR	NR	NR
Randerath 2001[57] APAP and CPAP (n = 47 completed)	NR	NR	Preferred by 35/47 (74%)	Preferred by 12/47 (26%) P<.01	NR	NR	NR	NR	NR	NR
D'Ortho 2000[42] APAP and CPAP (n = 25)	NR	NR	Preferred mode: 15/25 (60%)	Preferred mode: 8/25 (32%) (2/25 [8%] unable to tolerate either)	NR	NR	NR	NR	NR	NR

[a]APAP: a) RemStar Auto, b) Autoset Spirit, c) Breas Pv 10i

NR = not reported; NS = not statistically significant; PCS = physical component summary; MCS = mental component summary; SAQLI = Sleep Apnea Quality of Life Index; SEM = standard error of the mean; VAS = visual analog scale

Table 10. Intermediate Outcomes for KQ3

Study Intervention (n) Control (n)	Apnea-Hypopnea Index (AHI)		Oxygen Saturation		Sleep Symptom Scores		Other Sleep or Urinary Symptom Scores		Weight Loss	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Bakker 2011[40] CPAP and APAP (n = 12)	On-treatment (by machine at home): 13.2 (10.2) On-treatment (during PSG): 9.8/hr (9.5)	On-treatment (by machine at home): 8.0 (6.4) P = .06 On-treatment (during PSG): 7.3/hr (6.6) P = .35	Mean SpO2 94.8 (2.1) ODI 4% 6.8/hr(7.9) ODI3% 16.1/hr (16.6) Time <90% SpO2 1.5% (2.6)	Mean SpO2 95.5% (1.4) P = .03 ODI 4% 5.1/hr (5.1) P = .21 ODI 3% 11.4/hr (10.8) P = .15 Time <90% SpO2 1.2% (2.2) P = .50	ESS Baseline: 17.4 (4.7) post-APAP: 10.8 (5.7)	ESS Baseline: 17.4 (4.7) post-CPAP: 10.1 (6.1) P = .12	Total sleep time, min 382 (53.4)	Total sleep time, min 393.1 (44.9) P = .74	NR	NR
Lettieri 2011[33] Group 2: PSG+lab CPAP titration (n = 70) Group 3: PSG+APAP titration (n = 70)	NR	NR	NR	NR	ESS Baseline: 13.9 (4.4) Follow-up: 8.9 (2.1) Change: 36%	ESS Baseline: 14.1 (4.2) Follow-up: 8.4 (2.3) Change: 39.8%	NR	NR	NR	NR
Drummond 2010[43] APAP (n = 54) Usual care (n = 55)	NR	NR	NR	NR	ESS Baseline 14.8 (4.9) 1 month 11.6 (5.4) P = NS from baseline	ESS Baseline 14.1 (5.0) 1 month 12.7 (5.0) P = .25	FOSQ Baseline 14.0 (3.6) 1 month 15.5 (3.0)	FOSQ Baseline 14.2 (3.3) 1 month 14.6 (3.8) P = .17	NR	NR

Study Intervention (n) Control (n)	Apnea-Hypopnea Index (AHI)		Oxygen Saturation		Sleep Symptom Scores		Other Sleep or Urinary Symptom Scores		Weight Loss	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
McArdle 2010[38] Manual CPAP titration (n = 83) Home APAP titration (n = 86) -all outcomes per protocol NOT ITT unless specified	Baseline: 38.0 4 weeks: 8.0	Baseline: 38.0 P = NS 4 weeks: 7.1 P = NS	SaO_2 <90%, time, min Baseline: 2 4 weeks: 0 Avg SaO_2 4 weeks: 96% (2.0)	Baseline: 5 4 weeks: 0 P = NS 95.8% (2.0) P = NS	ESS Baseline: 13.8 (4.4) (n = 61) 4 weeks: 8.3 (4.5) (n = 62) Change: -5.5 (5.2) (n = 62) ITT ESS Baseline: 14.0 (4.2) (n = 70) 4 weeks: 8.5 (4.4) (n = 69) Change: -5.6 (5.0) (n = 69)	ESS Baseline 14.4 (4.0) (n = 62) 4 weeks: 8.7 (5.1) n = 61 P = NS Change: -5.7 (5.3) n = 61 P = NS ITT ESS Baseline: 14.3 (4.0) (n = 71) 4 weeks: 8.6 (5.1) (n = 70) P = NS Change: -5.7 (5.6) (n = 70) P = NS	Arousal Index, events/h 4 weeks: 22 (9) TST, min 4 weeks: 364 (71)	Arousal Index, events/h 4 weeks: 23 (12) P = NS TST, min 4 weeks: 351 (79) P = NS	NR	NR
Vennelle 2010[61] APAP and CPAP (n = 192, 181 analyzed)	Residual A+H/h 6.7 (SEM 0.4) (NOTE: only 70/181 had baseline A+H/h measure with mean of 49 [20])	Residual A+H/h 6.3 (SEM 0.4) P = .17	NR	NR	ESS Baseline: 14 (3) (overall) At 6 weeks: 9.5 (SEM 0.4)	ESS At 6 weeks: 10.0 (SEM 0.3) P = .031	NR	NR	NR	NR

Study Intervention (n) Control (n)	Apnea-Hypopnea Index (AHI)		Oxygen Saturation		Sleep Symptom Scores		Other Sleep or Urinary Symptom Scores		Weight Loss	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Damjanovic, 2009[41] Standard APAP (n = 25) Standard CPAP (n = 25) Intensive APAP (n = 25) Intensive CPAP (n = 25)	3 months: 4.8 (0.7) 9 months: 3.6 (0.8)	3 months: 6.7 (0.9) 9 months: 5.4 (1.4) P = NS	ODI Baseline: 35.6 (3.9) 3 months: 2.1 (0.3) 9 months: 2.9 (0.7)	ODI Baseline: 41.1 (3.8) 3 months: 4.1 (0.7) 9 months: 4.8 (1.3) P = NS	ESS 3 months: 6.4 (0.7) 9 months: 5.9 (0.7)	ESS 3 months: 7 (0.7) 9 months: 6.6 (0.7) P = NS	ARI Baseline: 30.6 (3.3) 3 months: 12.3 (1.3) 9 months: 12.9 (1.5)	ARI Baseline: 34.5 (3.1) 3 months: 16.4 (1.4) 9 months: 13.2 (1.5) P = NS	NR	NR
Galetke 2008[45] APAP and CPAP (n = 20)	Baseline 32.9 (19.1) (combined group) After 8 weeks: 5.6 (3.6)	After 8 weeks: 4.6 (2.9) P = NS	SaO_2 min, % Baseline 77.8 (8.4) (combined group) After 8 weeks: 86.5 (5.2)	SaO_2 min, % After 8 weeks: 88.3 (3.6) P = NS	ESS Baseline 10.3 (5.7) (combined group) After 8 weeks: 4.9 (4.6)	ESS After 8 weeks: 6.6 (4.8) P = NS	Arousals/h Baseline: 17.6 (9.2 (combined) After 8 weeks: 13.6 (8.6) Snoring, n of epochs Baseline: 436.3 (209.6) After 8 weeks: 54.9 (108.5)	Arousals/h After 8 weeks: 12.6 (7.3) Snoring, n of epochs After 8 weeks: 78.8 (88.3) P = NS	NR	NR
Fietze 2007[44] APAP n = 10 CPAP n = 11	Baseline: 43.3 (30.2) At 6 weeks: 4.4 (3.4)/hr P = NS	Baseline: 40.4 (26.1) At 6 weeks: 3.9 (4.3)/hr P<.05 form baseline	NR	NR	ESS baseline 12.9 (5.6) 6 weeks: 6.5 (4.3) P<.01 Did not differ at any point in time between the CPAP and APAP groups		NR	NR	NR	NR

Study Intervention (n) Control (n)	Apnea-Hypopnea Index (AHI)		Oxygen Saturation		Sleep Symptom Scores		Other Sleep or Urinary Symptom Scores		Weight Loss	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Meurice 2007[50] At 6 months Group 1, n = 14 Group 2, n = 13 Group 3, n = 15 Group 4, n = 12 Group 5, n = 11	**Group 2** Baseline: 49.9 (16.5) 6 m: 3.7 (3.9) **Group 3** Baseline: 53.4 (15.1) 6 m: 2.3 (2.9) **Group 4** Baseline: 48.1 (18.7) 6 m: 8.6 (10) **Group 5** Baseline: 54.5 (17.7) 6 m: 8.5 (9.7)	**Group 1** Baseline: 56.1 (21.4) 6 m: 2.4 (3.4)	*Mean SaO_2* **Group 2** Baseline: 92.5% (2.3) 6 m: 95% (1.5) **Group 3** Baseline: 91.8% (2) 6 m: 94.2% (1.5) **Group 4** Baseline: 92.6% (3.8) 6 m: 94.6% (2) **Group 5** Baseline: 91.9% (2.8) 6 m: 94.3% (1.8) *SaO_2<90%, min* **Group 2** Baseline: 16.2% (16.4) 6 m: 0.3% (0.6) **Group 3** Baseline: 23.3% (17.6) 6 m: 0.3% (0.6) **Group 4** Baseline: 11.7% (15.1) 6 m: 3% (5.3) **Group 5** Baseline: 16.7% (18.8) 6 m: 1.1% (2)	Mean SaO_2 **Group 1** Baseline: 90.8% (6.9) 6 m: 94% (1.8) P = NS SaO_2 <90%, min **Group 1** Baseline: 19.8% (28.7) 6 m: 2.2% (7.7) P = NS	ESS **Group 2** Baseline: 11.2 (5.6) 6 m: 6.5 (4.1) **Group 3** Baseline: 12.9 (4.3) 6 m: 5.2 (4.1) **Group 4** Baseline: 11.3 (3.8) 6 m: 7.2 (4) **Group 5** Baseline: 10 (6.2) 6 m: 7.5 (5.7)	ESS **Group 1** Baseline: 10.6 (5.2) 6 m: 5.9 (5.1) P = NS	TST, min **Group 2** Baseline: 400.6 (74.4) 6 m: 382.7 (94.7) **Group 3** Baseline: 390.1 (65.1) 6 m: 370.5 (65.2) **Group 4** Baseline: 372.5 (87.8) 6 m: 377.4 (65.9) **Group 5** Baseline: 371.7 (96.1) 6 m: 356.1 (43.9)	TST, min **Group 1** Baseline: 373.8 (91.8) 6 m: 376.6 (50.9)	NR	NR

Study Intervention (n) Control (n)	Apnea-Hypopnea Index (AHI)		Oxygen Saturation		Sleep Symptom Scores		Other Sleep or Urinary Symptom Scores		Weight Loss	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Mulgrew 2007[39] CPAP titration (n = 35) APAP titration (n = 33)	AHI Median (IQR) 3 m: 2.5/hr (0.9, 10.1)	AHI 3 m: 3.2/hr (1.7, 8.4) Difference at 3 m: 0.8 (95% CI -0.9, 2.3) P = .31	NR	NR	ESS Median (IQR) 3 m: 5.0 (3.0, 9.0)	ESS 3 m: 5.0 (2.0, 8.0) Difference at 3 m: 0.0 (95% CI -2.0, 2.0) P = .86	NR	NR	NR	NR
Nolan 2007[51] APAP and CPAP (n = 29)	Baseline: 14.7 (8) (combined) At 8 weeks: 2.7 (2.1) P = .15	At 8 weeks: 3.5 (2.5)	Mean SaO_2 Baseline: 92% (2.1) At 8 weeks: 93.2% (1.8) P = .44	Mean SaO_2 At 8 weeks: 93.3% (1.7)	ESS Score Baseline 12.3 (4) (combined) At 8 weeks: 8.6 (4.0) Total sleep time, min Baseline: 343 (48) (combined) At 8 weeks: 335 (43)	ESS Score At 8 weeks: 7.7 (4.6) P = .35 Total sleep time, min At 8 weeks: 349 (55) P = .09	Total snore events Baseline: 313 (259)/h (combined) At 8 weeks: 16 (11) P = .72 Respiratory arousals Baseline: 16 (14) (combined) At 8 weeks: 2 (3) P = .03	Total snore events At 8 weeks: 17 (16) Respiratory arousals At 8 weeks: 5 (4)	Body weight (88.1 (13.3)kg) did not change during the course of the study	
Patruno 2007 APAP (n = 15) CPAP (n = 16)	AHI/h 6 (2.3) P<.001	AHI/h 2 (1.6) Significantly reduced from baseline in both groups	ODI/h 4.8 (2.1) P<.001 SaO_2, mean 95.7% (17.4) SaO_2, nadir 88.1% (1.6)	ODI/h 1.1 (1.3) Significantly reduced in both groups SaO_2, mean 96.3% (0.8) SaO_2, nadir 90.8% (1.3) P = NS	NR	NR	NR	NR	NR	NR

Study Intervention (n) Control (n)	Apnea-Hypopnea Index (AHI)		Oxygen Saturation		Sleep Symptom Scores		Other Sleep or Urinary Symptom Scores		Weight Loss	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Nolan 2006 CPAP and 3 APAPs (n = 27)[a]	NR	NR	NR	NR	No further significant change in ESS after treatment with any APAP device	Fixed pressure CPAP reduced ESS from baseline of 15 (3-110) to 5 (3-11) P = .002	NR	NR	NR	NR
Nussbaumer 2006[54] APAP and CPAP (n = 30)	Baseline 41.1 (SE 3.6) At 1 month 4.6 (0.7)	At 1 month 5.4 (1.2) P = NS	ODI Baseline[b] 29 (SE 4) At 1 month 4.2 (0.7)	ODI At 1 month 4.1 (0.7) P = NS	ESS Baseline 12.7 (SE 0.6) At 1 month 6.6 (SE 0.6)	ESS At 1 month 6.6 (SE 0.6) P = NS	NR	NR	NR	NR
West 2006[62] Group 1, APAP (n = 31) Group 2, 1 wk titration CPAP (n = 33) Group 3, algorithm CPAP (n-34)	All data reported as median (5th/95th centile) 6 months 5.2 (1.5/13.2) P = .3	All data reported as median (5th/95th centile) 6 months **Group 2:** 3.6 (0.5/15.9) **Group 3:** 3.8 (0.7/26.1)	NR	NR	All data reported as median (5th/95th centile) ESS Pre CPAP 16.0 (10.6/23.0) P = .7 6 m: 6.0 (0.45/13.8)	All data reported as median (5th/95th centile) ESS **Group 2:** Pre CPAP 17.0 (10.4/22.6) 6 m: 5.0 (0/15.5) **Group 3:** Pre CPAP 16.5 (10.5/22.3) 6 m: 5.0 (0.5/12.5) P = .8	All data reported as median (5th/95th centile) MWT (mins) Pre CPAP 19.4 (1.4/40) P = .8 6 m: 40 (11.6/40) P = .2	All data reported as median (5th/95th centile) MWT (mins) **Group 2:** Pre CPAP 19.5 (2.9/40) 6 m: 40 (14.5/40) **Group 3:** Pre CPAP 15.7 (2.1/40) 6 m: 40 (2.2/40)	NR	NR

Study Intervention (n) Control (n)	Apnea-Hypopnea Index (AHI)		Oxygen Saturation		Sleep Symptom Scores		Other Sleep or Urinary Symptom Scores		Weight Loss	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Hukins 2004[46] APAP and CPAP (n = 55)	NR	NR	NR	NR	ESS Both groups improved from baseline (P<.001) P = NS		NR	NR	NR	NR
Hussain 2004[47] CPAP and APAP (n = 10)	Baseline: 47.2 (35.6) Follow-up: 13.1 (8.3)	Baseline: 47.2 (35.6) Follow-up: 9.6 (5.4) P = NS	Desaturation index Baseline: 53 (36) Follow-up: 15 (14) Basal O_2 Saturation Baseline: 91.3 (4.5) Follow-up: 94.6 (1)	Desaturation index Baseline: 53 (36) Follow-up: 10 (13) P = NS Basal O_2 Saturation Baseline: 91.3 (4.5) Follow-up: 95.3 (1.6) P = NS Minimum oxygen saturation (P<.05)	ESS Baseline: 11.1 (6.4) Follow-up: 8 (5.7)	ESS Baseline: 11.1 (6.4) Follow-up: 6.6 (5.9) P = NS	Arousal Index Baseline: 17.3 (17.7) Follow-up: 5.9 (6.5) Total sleep time, min Baseline: 381 (92) Follow-up: 346 (87)	Arousal Index Baseline: 17.3 (17.7) Follow-up: 4.9 (3.7) P = NS Total sleep time, min Baseline: 381 (92) Follow-up: 360 (108) P = NS	NR	NR
Marrone 2004[48] APAP and CPAP (n = 22)	NR	NR	NR	NR	ESS Baseline 16.3 (5) 1 month 3.9 (2.8)	ESS Baseline 16.3 (5) 1 month 4.9 (3.7) P = NS	NR	NR	NR	NR

Study Intervention (n) Control (n)	Apnea-Hypopnea Index (AHI)		Oxygen Saturation		Sleep Symptom Scores		Other Sleep or Urinary Symptom Scores		Weight Loss	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Masa 2004[37] Standard titration (n = 107) Autoadjusted titration (n = 106)	Pre-treatment: 62.8 (22.8) Post: 4.9 (7.6) Change at 3 months; 57.9 (22.6)	Pre: 61.8 (22.0) Post: 5.1 (6.8) Change at 3 months: 56.6 (21.0) P = NS	SaO_2 <90% of TST Pre: 29.9 (27.3) Post: 1.4 (4.1) Change at 3 months: 28.2 (26.1)	SaO_2 <90% of TST Pre: 25.3 (25.0) Post: 3.0 (13.9) Change at 3 months: 22.0 (28.1) P = NS	*ESS* Pre: 15.2 (3.5) Post: 7.2 (4.4) Change at 3 months: 8.1 (5.4) P = NS *FOSQ* Pre: 94.4 (1.07) Post: 108.0 14.3) Change - 14.2 (17.2)	*ESS* Pre: 15.9 (3.5) Post: 7.9 (4.6) Change at 3 months: 8.0 (4.8) P = NS for change between groups *FOSQ* Pre: 84.4 (22.8) Post: 105.1 (16.0) Change - 20.8 (20.1) P = NS	Arousal Index Pre: 55.5 (19.3) Post: 12.0 (8.5) Change at 3 months: 8.1 (5.4)	Arousal Index Pre: 55.2 (18..0) Post: 12.3 (10.0) Change at 3 months: 8.0 (4.8) P = NS	NR	NR
Noseda 2004[53] CPAP and APAP (n = 24)	NR	NR	NR	NR	ESS Baseline: 10.7 (2.4) 8 w: 5.1 (2.8) Sleep latency 12 min (12) Self-estimated effective sleep 6.1 (1.3)	ESS Baseline: 10.7 (2.4) 8 w: 6.1 (2.8) P<.01 Sleep latency 14 (12) P = NS Self-estimated effective sleep 6.2h (1.3) P = NS	AI Baseline: Mean 24.6 (22.6) On APAP: Median 0.45/h	AI Baseline: Mean 24.6 (22.6) On APAP: Median 0.4/h P = NS	NR	NR

Study Intervention (n) Control (n)	Apnea-Hypopnea Index (AHI)		Oxygen Saturation		Sleep Symptom Scores		Other Sleep or Urinary Symptom Scores		Weight Loss	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Massie 2003[49] CPAP and APAP (n = 44)	Residual AHI 9.6 (5.3)	Residual AHI 10.7 (6.6) P = NS	NR	NR	ESS 8 (4)	ESS 9 (4) P = NS	P<.006 more restful sleep, overall better sleep quality, less discomfort from pressure, and less trouble getting to sleep	NR	NR	NR
Planès 2003[56] Auto (n = 16) Conventional (n = 14)	Events/h Baseline: 57.5 (16.5) 2 months: 7.6 (6.9)	Events/h Baseline: 61.0 (17.4) 2 months: 10.4 (12.5) P = NS	SaO_2<90% (time spent) Baseline: 24.9 (21.6) 2 months: 0.3 (0.6)	SaO_2<90% (time spent) Baseline: 12.7 (12.8) 2 months: 1.9 (5.0) P = NS	ESS score Baseline: 15.5 (4.7) 2 months: 7.5 (3.4) P<.0001 from baseline	ESS score Baseline: 14.7 (3.9) 2 months: 7.6 (3.4) P = NS	NR	NR	NR	NR
Senn 2003[59] Auto Adjust and AutoSet and Fixed (n = 29)	Mean over treatment period AutoSet 7.8 (0.9) AutoAdjust 6.6 (1.3)	7.4 (1.3) P = NS	Time with SaO_2<90%, % time in bed AutoSet 0.9 (0.7) AutoAdjust 2.7 (1.9)	Time with SaO_2<90%, % time in bed 1.1 (0.7) P = NS	ESS score AutoSet" 9.0 (0.6) AutoAdjust 8.0 (0.8)	ESS score 8.2 (0.7) P = NS	"Overall benefit from CPAP therapy" 5 pt Likert scale AutoSet 4.3 (0.1) AutoAdjust 4.1 (0.2)	"Overall benefit from CPAP therapy" 4.3 (0.2) P = NS	NR	NR

Study Intervention (n) Control (n)	Apnea-Hypopnea Index (AHI)		Oxygen Saturation		Sleep Symptom Scores		Other Sleep or Urinary Symptom Scores		Weight Loss	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Randerath 2001[57] APAP and CPAP (n = 52, 47 completed)	Baseline: 35.1 (26) combined At 6 weeks: 5.0 (5.2)	At 6 weeks: 4.3 (6.3) P<.001	NR	NR	ESS Baseline: 11.1 (5.1) At 6 weeks: 7.8 (4.7) Total sleep time Baseline: 319 (55) (combined) At 6 weeks 324 (52)	ESS At 6 weeks: 8.8 (4.6) Total sleep time At 6 weeks: 330 (43)	Snoring Baseline: 49 (36)/h (combined) At 6 weeks: 13 (20)/h Total number of arousals Baseline: 34.0 (21.7)/h At 6 weeks: 10.9 (5.7)	Snoring At 6 weeks: 6 (13) P<.001 Total number of arousals At 6 weeks: 12.6 (8.3) P<.001	NR	NR
D'Ortho 2000[42] APAP and CPAP (n = 25)	Baseline: 57.8 (5.8) (combined) After 2 months: 10.6 (9.3)	After 2 months: 9.7 (1.9)	Mean SaO_2, % Baseline: 93.0 (3.0) (combined) After 2 months: 95.6 (1.6)	Mean SaO_2, % After 2 months: 95.9 (1.5)	ESS Baseline: 12.7 (5.3 (combined) After 2 months: 9.3 (4.8)	ESS After 2 months: 9.2 (5.5) P = NS	NR	NR	NR	NR
Teschler 2000[60] CPAP and APAP (n = 10)	Baseline: 52.9 (8.1) At home: 4.0 (0.3)	At home: 3.7 (0.3) P = NS, no order effect, no tx by order interaction	NR	NR	NR	NR	NR	NR	NR	NR

[a]APAP: a) RemStar Auto, b) Autoset Spirit, c) Breas Pv 10i; [b]Data from Visual Analog Scale (0-100) with lower values "better"
ESS = Epworth Sleepiness Scale (non-inferiority margin was -2.0); FOSQ = Functional Outcomes of Sleep Questionnaire; ODI = oxygen desaturation index; MD = mean difference; SASQ = Sleep Apnea Symptoms Questionnaire, TST = Total sleep time, Ar/AwI = arousals and awakening index, ARI = arousal index; SaO_2 = oxygen saturation; MWT = maintenance of wakefulness test; NS = not statistically significant; NR = not reported

Table 11. Intermediate Outcomes for KQ3, Continued

Study Intervention (n) Control (n)	BMI		Blood Pressure		HbA1c		Time to Initiation of Therapy		Harms (False Positives/Negatives)	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
McArdle 2010[38] Manual CPAP titration (n = 83) Home APAP titration (n = 86) -all outcomes per protocol NOT ITT unless specified	NR	NR	SBP Baseline: 128.4 (14.9) 4 weeks: 126.5 (15.4) DBP Baseline: 81.5 (10.1) 4 weeks: 77.0 (10.3)	SBP Baseline: 125.2 (17.3) 4 weeks: 122.4 (17.1) ES: 0.25 (-0.05, 0.56) P = NS DBP Baseline: 79.1 (9.9) 4 weeks: 76.1 (8.6) ES: 0.09 (-0.2, 0.4) P = NS	NR	NR	NR	NR	NR	NR
Damjanovic, 2009[41] Standard APAP (n = 25) Standard CPAP (n = 25) Intensive APAP (n = 25) Intensive CPAP (n = 25)	BMI did not significantly change in the course of the study		NR	NR	NR	NR	NR	NR	NR	NR
Nolan 2007[51] APAP and CPAP (n = 29)	NR	NR	No significant change in blood pressure with either APAP or CPAP during the course of the study		NR	NR	NR	NR	NR	NR

Study Intervention (n) Control (n)	BMI		Blood Pressure		HbA1c		Time to Initiation of Therapy		Harms (False Positives/Negatives)	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Patruno 2007 APAP (n = 15) CPAP (n = 16)	"not significantly affected by treatments"		SBP Baseline: 142 (12) 3 m: 136 (6) P = NS, change from baseline DBP Baseline: 87.5 (4) 3 m: 86 (4) P = NS, change from baseline	SBP Baseline: 144 (10) 3 m: 132 (8) P<.001, change from baseline DBP Baseline: 88 (4) 3 m: 79 (6) P<.001, change from baseline	NR	NR	NR	NR	NR	NR
West 2006[62] Group 1, APAP (n = 31) Group 2, 1 wk titration CPAP (n = 33) Group 3, algorithm CPAP (n-34)	NR	NR	All data reported as median (5th/95th centile) Mean BP Pre CPAP 95.8 (77/122) P = .9 6 m: 99.6 (77.3/119) P = .5	All data reported as median (5th/95th centile) Mean BP **Group 2:** Pre CPAP 95.2 (77.3/118.3) 6 m: 96.7 (82.7/119) **Group 3:** Pre CPAP 96.2 (75.0/120.6) 6 m: 96.4 (73.3/114.3)	NR	NR	NR	NR	NR	NR

Study Intervention (n) Control (n)	BMI		Blood Pressure		HbA1c		Time to Initiation of Therapy		Harms (False Positives/Negatives)	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Planès 2003[56] Auto (n = 16) Conventional (n = 14)	"in neither group did BMI change significantly during the study"		NR	NR	NR	NR	11.8 (15.5) days P<.01	47.2 (46.5) days	NR	NR

NR = not reported; NS = not statistically significant; BP = blood pressure; SBP = systolic blood pressure; DBP = diastolic blood pressure

Table 12. Intermediate Outcomes for KQ3, Continued

Study Intervention (n) Control (n)	Harms (Overdiagnosis)		Harms, Adverse Events (describe)		Harms, Adverse Events (describe)		Costs per Patient, US dollars (unless otherwise noted)		Compliance/Adherence	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Bakker 2011[40] CPAP and APAP (n = 12)	NR	NR	NR	NR	NR	NR	NR	NR	6.3 (1.8) hrs/night	5.8 (2.8) hrs/night (P = .11)
Lettieri 2011[33] Group 2: PSG+lab CPAP titration (n = 70) Group 3: PSG+APAP titration (n = 70)	NR	NR	10% dis-continued therapy	8.6% dis-continued therapy P = NS between groups	NR	NR	Diagnostic PSG cost: $704.28 CPAP titration cost: $753.76 (Medicare reimbursement rates)		% nights used 72.4 (22) hrs/night, nights used 4.8 (1.7) Use of PAP > 4h/ night for > 70% of nights: 50%	% nights used 73.2 (18) ES: -0.04 (-0.4, 0.3) hrs/night, nights used 4.7 (1.1) ES: 0.07 (-0.2, 0.4) Use of PAP > 4h/ night for > 70% of nights 51.4%

Study Intervention (n) Control (n)	Harms (Overdiagnosis)		Harms, Adverse Events (describe)		Harms, Adverse Events (describe)		Costs per Patient, US dollars (unless otherwise noted)		Compliance/Adherence	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
McArdle 2010[38] Manual CPAP titration (n = 83) Home APAP titration (n = 86) - all outcomes per protocol NOT ITT unless specified	NR	NR	NR	NR	NR	NR	Staff/pt A$70.74 Equipment etc./pt A$61.35 Direct costs/pt A$132.09 Travel/pt A$26.91	Staff/pt A$250.95 Equipment etc./pt A$93.19 Direct costs/pt A$817.84 Travel/pt A$15.04	% patients continuing CPAP use at 4 weeks 85% (n = 63) ITT: % using CPAP at 4 weeks 81% (n = 70) CPAP use (hrs) at 4 weeks: 4.39 (2.2) (n = 61) ITT: CPAP use, hrs, at 4 weeks 4.24 (2.2) (n = 68)	% patients continuing CPAP use at 4 weeks 87% (n = 62) P = NS ITT: 86% (n = 71) P = NS CPAP use (hrs) at 4 weeks: 4.36 (2.2) (n = 63) ES: 0.014 (-0.34, 0.37) P = NS ITT 4.38 (2.2) (n = 70) ES:-0.06 (-0.4, 0.27) P = NS
Vennelle 2010[61] APAP and CPAP (n = 192, 181 analyzed)	NR	NR	NR	NR	NR	NR	NR	NR	Mean CPAP use 4.2 (SEM 0.2)h/night	4.0 (SEM 0.2)h/night P = .047

Study Intervention (n) Control (n)	Harms (Overdiagnosis)		Harms, Adverse Events (describe)		Harms, Adverse Events (describe)		Costs per Patient, US dollars (unless otherwise noted)		Compliance/Adherence	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Damjanovic 2009[41] Standard APAP (n = 25) Standard CPAP (n = 25) Intensive APAP (n = 25) Intensive CPAP (n = 25)	NR	NR	NR	NR	NR	NR	NR	NR	Daily usage (h) 3 m: 5.4 (0.2) 9 m: 5.2 (0.4) Percent of days 3 m: 76% (3.9) 9 m: 67.9% (5) Hours used/sleep time 3 m: 73.4 (3.1) 9 m: 72.5 (5)	Daily usage (h) 3 m: 5.4 (0.3) 9 m: 5.1 (0.3) Percent of days 3 m: 75% (4.1) 9 m: 69.2% (4.9) Hours used/sleep time 3 m: 81.4 (5.8) 9 m: 72.1 (5.2) No difference between groups in adherence
Galetke 2008[45] APAP and CPAP (n = 20)	NR	NR	NR	NR	NR	NR	NR	NR	382 (107) min/night	383 (116) min/night P = NS
Fietze 2007[44] APAP n = 10 CPAP n = 11	NR	NR	NR	NR	NR	NR	NR	NR	Overall use: 77% (25%) of nights Nocturnal usage: APAP: 5.0 (1.6)h CPAP: 4.2 (2.2)h No significant differences in the course of compliance between the 2 patient groups	

Study Intervention (n) Control (n)	Harms (Overdiagnosis)		Harms, Adverse Events (describe)		Harms, Adverse Events (describe)		Costs per Patient, US dollars (unless otherwise noted)		Compliance/Adherence	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Meurice 2007[50] At 6 months Group 1, n = 14 Group 2, n = 13 Group 3, n = 15 Group 4, n = 12 Group 5, n = 11	NR	NR	NR	NR	NR	NR	NR	NR	CPAP use, h/night **Group 2** 6 m: 5.5 (1.4) **Group 3** 6 m: 6.1 (1.6) **Group 4** 6 m: 5.1 (1.6) **Group 5** 6 m: 7.0 (1.9)	CPAP use **Group 1:** 6 m: 6.5 (1.8) P = NS
Mulgrew 2007[39] CPAP titration (n = 35) APAP titration (n = 33)	NR	NR	NR	NR	NR	NR	NR	NR	CPAP adherence (h/night) Median (IQR) 3 months: 6.0 (5.1, 7.1)	CPAP adherence 3 months: 5.4 (3.7, 6.4) Difference at 3 months: -1.12 (95% CI -2.0, 0.2) P = .02
Nolan 2007[51] APAP and CPAP (n = 29)	NR	NR	All patients experienced some side effects on each treatment, there was no significant difference between APAP and CPAP in terms of side effects (dry mouth, blocked/runny nose, pressure felt too high, claustrophobic)		NR	NR	NR	NR	Nights used 79% (29) Mean hrs used per night used 4.9 (2.1)	Nights used 81% (25) P = .87 4.9 (1.9) P = .94
Patruno 2007 APAP (n = 15) CPAP (n = 16)	NR	NR	NR	NR	NR	NR	NR	NR	6.2 (0.8) h/day P = NS	6.0 (1.0) h/day

Study Intervention (n) Control (n)	Harms (Overdiagnosis)		Harms, Adverse Events (describe)		Harms, Adverse Events (describe)		Costs per Patient, US dollars (unless otherwise noted)		Compliance/Adherence	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Richard 2007 APAP (n = 96) CPAP (n = 78)	NR	NR	NR	NR	NR	NR	NR	NR	Nights/wk[b] 6.3 (1.4) (n = 96) Hrs/night[b] 6.3 (1.8) (n = 95) Used ≥ 4h/night, ≥ 5d/wk: (74/96) 77.1%	Nights/wk 6.4 (1.4) (n = 76) P = .57 Hours/night 6.5 (1.5) (n = 75) P = .64 Used ≥ 4h/night, ≥ 5d/wk: (64/78) 82.1%
Nolan 2006 CPAP and 3APAPs (n = 27)[a]	NR	NR	Nasal symptoms: greater problem for c); P<.05 vs a) Throat/mouth symptoms: greater problem for c); P<.05 vs a) Pressure discomfort: greater problem for c); P<.05 vs a) and b)		NR	NR	NR	NR	Nights used (%) **a)** 100 (79-100) P = NS vs CPAP **b)** 96 (42-100) P = NS vs CPAP **c)** 59 (17-83) P<.01 vs CPAP, a, and b Hrs/night **a)** 7.1 (5.3-8.1) P = NS vs CPAP **b)** 6.8 (5.9-8.0) P = NS vs CPAP **c)** 5.0 (3.8-5.6) P<.01 vs CPAP, a, and b	Nights used (%) 100 (94-100) Hrs/night 6.6 (5.9-7.9)

Study Intervention (n) Control (n)	Harms (Overdiagnosis)		Harms, Adverse Events (describe)		Harms, Adverse Events (describe)		Costs per Patient, US dollars (unless otherwise noted)		Compliance/Adherence	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Nussbaumer 2006[54] APAP and CPAP (n = 30)	NR	NR	No side effects requiring interruption of therapy or consultations		Nasal stuff-iness[d] 24 (SE 6) Sore/dry mouth/ throat 27 (SE 5) Discom-fort with air pressure 8 (SE 2)	Nasal stuffi-ness 24 (SE 6) P = NS Sore/dry mouth/ throat 34 (SE 6) P = NS Discom-fort with air pressure 27 (SE 5) P<.05	NR	NR	% of days with > 4 hours use 72% (SE 4) Hrs/night 5.1 (SE 0.3)	% of days with > 4 hours use 68% (SE 5) P = NS Hrs/night 4.8 (SE 0.3) P = NS
West 2006[62] Group 1, APAP (n = 31) Group 2, 1 wk titration CPAP (n = 33) Group 3, algorithm CPAP (n-34)	NR	NR	NR	NR	NR	NR	NR	NR	All data reported as median (5th/95th centile) Hrs used/night 6 m: 5.49 (0/7.5) P = .23 Nights used (%) 6 months 100 (5/100) P = .21	All data reported as median (5th/95th centile) Hrs used/night 6 months **Group 2:** 4.9 (0/7.2) **Group 3:** 4.0 (0/8.3) Nights used (%) 6 months **Group 2:** 98.3 (61/100) **Group 3:** 92.6 (33/100)

Study Intervention (n) Control (n)	Harms (Overdiagnosis)		Harms, Adverse Events (describe)		Harms, Adverse Events (describe)		Costs per Patient, US dollars (unless otherwise noted)		Compliance/Adherence	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Hukins 2004[46] APAP and CPAP (n = 55)	NR	NR	Nasal irritation or obstruction5 Pressure intolerance2 Partner dislike 0 Total number of side effects15	Nasal irritation or obstruction 10 P = .27 Pressure intolerance 5 P = .44 Partner dislike 1 P = .99 Total number of side effects 28 P = .02	NR	NR	NR	NR	Avg. nightly use 5.05 (2.38) hours per night P = .14 Percentage of nights treatment used 83.3% (23.3%)	Avg. nightly use 4.86 (2.65) hours per night Percentage of nights treatment used 78% (32.6%) P = .29
Hussain 2004[47] CPAP and APAP (n = 10)	NR	NR	NR	NR	NR	NR	NR	NR	H/night 4.3 (1.9)	H/night 3.7 (2.6) P = NS
Marrone, 2004[48] APAP and CPAP (n = 22)	NR	NR	NR	NR	NR	NR	NR	NR	Hours/day 4.9 (1.7) Days of machine use 88.8% (15.2%)	Hours/day 4.4 (1.9) P = NS Days of machine use 83.9% (18.6%) P = NS
Masa 2004[37] Standard titration (n = 107) Autoadjusted titration (n = 106)	NR	NR	"no important differences" between standard and autoadjusted titration; tendency for more side effects (*eg*, rhinitis, mask intolerance, aerophagia, headache, smothering sensation, bed partner intolerance) in autoadjusted group		NR	NR	NR	NR	Use hr/day 5.3 (1.9)	Use hr/day 5.2 (2.0) ES: 0.05 (-0.2, 0.32) P = NR

Study Intervention (n) Control (n)	Harms (Overdiagnosis)		Harms, Adverse Events (describe)		Harms, Adverse Events (describe)		Costs per Patient, US dollars (unless otherwise noted)		Compliance/Adherence	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Noseda 2004[53] CPAP and APAP (n = 24)	NR	NR	NR	NR	NR	NR	NR	NR	Median percentage of nights 96.5% Mean use per effective night, h 5.3 (1.9) P = NS	Median percentage of nights 95.5% P = NS Mean use per effective night, h 5.5 (1.5)
Massie 2003[49] CPAP and APAP (n = 44)	NR	NR	NR	NR	NR	NR	NR	NR	% of nights used 92% (11) Minutes used/24hrs 306 (114) P<.005	% of nights used: 88% (15) P = NS Minutes used/24hrs 271 (115)
Planès 2003[56] Auto (n = 16) Conventional (n = 14)	NR	NR	NR	NR	NR	NR	Hospital care: €602 P<.001 between groups Tele-communica-tion: €155 Total cost: €1264 P<.01	Hospital care: €1220 Tele-communica-tion: €9 P<.001 between groups Total cost: €1720	Compliance (defined as 3hrs/night) 13/16 (81%) Mean duration 4.5h (1.7)	Compliance (defined as 3hrs/night) 14/14 (100%) Mean duration 5.3h (1.4) P = NS between groups

Study Intervention (n) Control (n)	Harms (Overdiagnosis)		Harms, Adverse Events (describe)		Harms, Adverse Events (describe)		Costs per Patient, US dollars (unless otherwise noted)		Compliance/Adherence	
	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP	APAP	CPAP
Senn 2003[59] Auto Adjust and AutoSet and Fixed (n = 29)	NR	NR	Dry mouth AutoSet 8/29 (28%) AutoAdjust 7/29 (24%) Skin irritation AutoSet 9/29 (31%) AutoAdjust 8/29 (28%) Nasal irritation AutoSet 5/29 (17%) AutoAdjust 8/29 (28%)	Dry mouth 3/29 (10%) Skin irritation 10/29 (34%) Nasal irritation 8/29 (28%) All P = NS all “mild”	NR	NR	NR	NR	Mean h/night AutoSet 5.5 (0.2) Auto Adjust 5.5 (0.2) Nights with > 2.5h, % AutoSet 83% (3) Auto Adjust 79% (4)	Mean h/night 5.6 (0.2) Nights with > 2.5h, % 82% (3) P<.01 for all compared to baseline, P = NS
Randerath 2001[57] APAP and CPAP (n = 47 completed)	NR	NR	Side effects were mild under both modes, and no significant differences were observable		NR	NR	NR	NR	Usage: 98.4% of days (APAP and CPAP) Minutes/ day APAP: 315.4 (94.7) CPAP: 315.4 (97.4)	
D’Ortho 2000[42] APAP and CPAP (n = 25)	NR	NR	NR	NR	NR	NR	NR	NR	Hours of use per night 4.1 (1.8) CPAP use ≥ 4h/night 18/25 (72%)	Hours of use per night 4.7 (1.8) P = .20 CPAP use ≥ 4h/night 19/25 (76%)

[a]APAP: a) RemStar Auto, b) Autoset Spirit, c) Breas Pv 10i
[b]Excludes failures
ES = effect size; ITT = intent to treat (analysis); NS = not statistically significant; NR = not reported; PSG = polysomnography; SE = standard error

APPENDIX D. STRENGTH OF EVIDENCE

OUTCOME	Strength of Evidence Elements[a]					Summary of Findings	
	Risk of Bias	Consistency	Directness	Precision	Publication Bias	Description of Effect	Strength of Evidence Direction[b]
KQ1: SLEEP PHYSICIAN CARE COMPARED TO MANAGEMENT BY PRIMARY CARE, SLEEP-SPECIALIST NURSES OR OTHER NON-SLEEP PHYSICIANS							
Access to care						We found no evidence for this outcome.	Insufficient
Epworth Sleepiness Scale (ESS)	Moderate	Consistent	Direct	Precise	Unclear	Based on 4 RCTs (n = 568), improvement from baseline in ESS scores was similar for patients being managed by primary care/sleep-specialist nurses compared to sleep specialist physicians (SMD = 0.06 [95% CI -0.15, 0.26]). One observational study also found ESS scores were similar between groups.	Moderate Similar
Quality of life	Moderate	Consistent	Direct	Precise	Unclear	Based on 3 RCTs (n = 524), quality of life measures were similar for patients being managed by primary care/ sleep-specialist nurses compared to sleep specialist physicians. SMDs for SF-36 Vitality and Mental Health scores were -0.04 [95% CI -0.22, 0.15]) and -0.04 [95% CI -0.22, 0.14], respectively	Moderate Similar
Compliance, hours per night	Moderate	Consistent	Direct	Precise	Unclear	Based on 4 RCTs (n = 568), compliance was similar for patients being managed by primary care/sleep-specialist nurses compared to sleep specialist physicians (WMD = -0.29 [95% CI -0.71, 0.12]). One observational cohort study also found compliance was similar between groups but one study based on retrospective chart review reported compliance was greater in the sleep specialist physician group compared to the non-sleep specialist group.	Moderate Similar
Adverse events	Moderate	Unknown	Direct	Imprecise	Unclear	Based on one RCT (n = 65) that did not report adverse events by treatment arm, the evidence is insufficient to draw conclusions.	Insufficient
KQ3: HOME APAP TECHNOLOGY VERSUS STANDARD IN-CENTER MANUAL CPAP TITRATION							
Access to care						We found no evidence for this outcome.	Insufficient

OUTCOME	Strength of Evidence Elements[a]					Summary of Findings	
	Risk of Bias	Consistency	Directness	Precision	Publication Bias	Description of Effect	Strength of Evidence Direction[b]
Epworth Sleepiness Scale (ESS)	Moderate	Consistent	Direct	Precise	Unclear	Based on 2 RCTs (n = 414) with moderate risk of bias, we found improvement from baseline in ESS scores was similar for patients allocated to home APAP titration compared to patients allocated to in-center CPAP titration (SMD = 0.0 [95% CI -0.22, 0.21]). One moderate risk of bias RCT (n = 68) found median change in EES scores from baseline was also similar between groups (MD -1 [95% CI --1, -4]). One observational cohort study also found ESS scores were similar between groups.	Moderate Similar
Quality of life	Moderate	Consistent	Direct	Precise	Unclear	Based on two RCTs (n = 414) with moderate risk of bias, we found quality of life measures were similar for patients allocated to home APAP titration compared to patients allocated to in-center CPAP titration. The SMDs for SF-36 Mental Health and Physical Health scores were 0.08 [95% CI -0.14, 0.29] and -0.21 [95% CI -0.61, 0.20], respectively. Results for the Physical Health scores were imprecise. One moderate risk of bias RCT (n = 68) found median improvement from baseline in the SAQLI was similar between groups (median difference = 0.17 [95% CI -0.6, 0.9])	Moderate Similar
Compliance, hours per night	Moderate	Inconsistent	Direct	Precise	Unclear	Based on two RCTs (n = 414) with moderate risk of bias, we found compliance was similar for patients allocated to home APAP titration compared to patients allocated to in-center CPAP titration (WMD = 0.02 [95% CI -0.41, 0.45]). One moderate risk of bias RCT (n = 68) found median compliance was better in the APAP group versus the CPAP group (MD -1.1 [95% CI -2.0, -0.2]). One observational cohort study found compliance was similar between groups	Low Similar

OUTCOME	Strength of Evidence Elements[a]					Summary of Findings	
	Risk of Bias	Consistency	Directness	Precision	Publication Bias	Description of Effect	Strength of Evidence Direction[b]
Adverse events	Moderate	Unknown	Direct	Imprecise	Unclear	Based on the findings of one RCT (n = 245) that reported no "important differences" in adverse events between the home APAP and in-lab CPAP and groups, the evidence is insufficient to draw conclusions.	Insufficient
KQ3: APAP VERSUS CPAP TREATMENT							
Access to care						We found no evidence for this outcome.	Insufficient
Epworth Sleepiness Scale (ESS)	Moderate	Consistent	Direct	Precise	Unclear	Based on four parallel group RCTs (n = 327) with aggregate moderate risk of bias, we found improvement from baseline in ESS scores was similar for patients allocated to APAP treatment compared to patients allocated to CPAP treatment (SMD = 0.18 [95% CI -0.06, 0.43]). Two parallel group trials not pooled (reported as a median or data not shown) also found improvement from baseline in ESS scores similar between groups. Ten crossover RCTs (n = 269) reported similar improvements between groups and two (N = 227) reported greater improvement with APAP.	Moderate Similar
Quality of life	Moderate	Consistent	Direct	Precise	Unclear	Based on 3 parallel group RCTs (n = 202) with aggregate moderate risk of bias, we found quality of life measures (SF-36, SAQLI) were similar for patients allocated to APAP treatment compared to patients allocated to CPAP treatment (data were not pooled due to variation in reporting of results, *ie,* reported as medians). Six crossover RCTs (n = 393) also reported no differences in most of the quality of life measures between the treatment groups.	Moderate Similar
Compliance, hours per night	Moderate	Consistent	Direct	Precise	Unclear	Based on 5 parallel group RCTs (n = 279) with aggregate moderate risk of bias, we found compliance was similar for patients allocated to APAP treatment compared to patients allocated to CPAP treatment (WMD = -0.08 [95% CI -0.55,	Moderate Similar

OUTCOME	Strength of Evidence Elements[a]					Summary of Findings	
	Risk of Bias	Consistency	Directness	Precision	Publication Bias	Description of Effect	Strength of Evidence Direction[b]
						0.38]). One parallel group RCT reporting median compliance, most of the remaining crossover RCTs, and one observational study also found compliance was similar between groups.	
Adverse events	Moderate	Consistent	Direct	Imprecise	Unclear	Adverse events were infrequently reported. One parallel group RCT (n = 109) reported adverse events, chest pain in 12% and 9% of APAP and CPAP patients, respectively. Five crossover trials reported adverse events for both APAP and CPAP treatments. One trial (n = 55) reported a higher frequency of total number of events and another trial (n = 34) reported a higher incidence of pressure discomfort with CPAP therapy arm compared with the APAP treatment. Three trials (n = 112) reported no differences in adverse events between the treatment groups.	Low Similar

AHI = apnea–hypopnea index ; APAP = Auto-adjusted (autoregulated) continuous positive airway pressure; CPAP = continuous positive airway pressure; RCT = randomized controlled trial; SF-36 = Short Form-36; SMD = standardized mean difference; WMD = weighted mean difference

[a]Strength of Evidence Elements[28]

Precision: Degree of certainty surrounding an effect estimate; in meta-analysis, the confidence interval around the summary effect size

Consistency: Degree to which reported effect sizes appear to have the same direction of effect

Directness: Whether the evidence links the interventions directly to health outcomes

Risk of bias: Degree to which includes studies have a high likelihood of protection against bias; 2 main elements are study design and aggregate quality of the studies

[b]Direction of difference between groups