The Comparison of CPAP and Oral Appliances in Treatment of Patients With OSA: A Systematic Review and Meta-analysis

Wenyang Li MD, Lin Xiao PhD, and Jing Hu PhD

BACKGROUND: A systematic review and meta-analysis was performed to compare the outcomes of oral appliances (OAs) with those of CPAP in treatment of patients with obstructive sleep apnea (OSA). METHODS: Relevant studies were retrieved from the following electronic databases, up to and including September 2012: MEDLINE, PubMed, EMBASE, and Central Register of Controlled Trials. The main outcomes were Epworth Sleepiness Scale score, health-related quality of life, cognitive performance, blood pressure, apnea-hypopnea index (AHI), arousal index, minimum S_{pO_2}, percent rapid eye movement sleep, treatment usage, side effects, treatment preference, and withdrawals. RESULTS: Fourteen trials were finally included in this review. Our results demonstrated that the effects on Epworth Sleepiness Scale score ($P = .31$ and $.09$ in crossover and parallel-group trials), health-related quality of life, cognitive performance, and blood pressure of OAs and CPAP were similar. Besides, pooled estimates of crossover trials suggested a significant difference in favor of CPAP regarding AHI ($P < .001$), arousal index ($P = .001$), and minimum S_{pO_2} ($P < .001$), while pooled estimates of parallel-group trials showed a significant difference in favor of CPAP regarding AHI ($P < .001$) and percent rapid eye movement sleep ($P = .02$). Moreover, OAs and CPAP yielded fairly similar results in terms of treatment usage ($P = .26$ for hours/night in crossover trials, and $P = .14$ for hours/night and $P = .19$ for nights/week in parallel-group trials), treatment preference, side effects, and withdrawals ($P = .34$ in parallel-group trials). CONCLUSIONS: CPAP yielded better polysomnography outcomes, especially in reducing AHI, than OAs, indicating that OAs were less effective than CPAP in improving sleep-disordered breathing. However, similar results from OAs and CPAP in terms of clinical and other related outcomes were found, suggesting that it would appear proper to offer OAs to patients who are unable or unwilling to persist with CPAP. *Key words: oral appliances; CPAP; obstructive sleep apnea; meta-analysis.* [Respir Care 2013;58(7):1184–1195. © 2013 Daedalus Enterprises]

Introduction

Obstructive sleep apnea (OSA) is characterized by recurrent obstruction of the upper airway, often resulting in oxygen desaturation and arousal from sleep.[1] Excessive daytime sleepiness, snoring, reduction in cognitive function, and the risk of developing long-term vascular consequences are among the common symptoms of this condition.[2] There is now a considerable body of literature documenting the pathophysiology and consequences of OSA; however, the morbidity, benefits of treatment, and optimal mode of management of OSA remain a clinical dilemma.

CPAP has been proposed as the most effective treatment for OSA. Applying CPAP during the night is effec-

tive in reducing symptoms of sleepiness and improving quality of life measurements in people with OSA.[3] However, CPAP is a constraining treatment with frequent local adverse effects that can constitute an obstacle to regular and prolonged use of the apparatus,[4] so individuals may abandon or adhere poorly to this therapy. Of OSA patients in whom CPAP is recommended, 5–50% reject this treatment and 12–25% of the remaining patients can be expected to discontinue CPAP,[5,6] and the milder the symptoms of OSA, the less likely are the subjects to accept CPAP. Besides, although there is no doubt that CPAP is usually very effective, increased healthcare costs with this treatment may be another important factor that affects its adherence in patients with OSA.[7]

Oral appliances (OAs) have emerged as an increasingly popular alternative for CPAP over the past decade.[8] The rationale behind the use of OAs is unclear, but is probably multifactorial, involving both a structural change with enhancement of the caliber of the airway and also triggering of stretch receptors, which activate the airway support muscles.[9] Several studies have demonstrated that OAs can effectively reduce the severity of sleep-disordered breathing and lead to symptomatic improvement, and patients seem to be more adherent to OAs than to CPAP.[10,11]

Many randomized trials have compared the outcomes of OAs versus CPAP in the treatment of patients with OSA,[10-15] most of which indicate that OAs are less effective in reducing AHI but are preferred over CPAP. However, none of these trials has been large enough to confirm the outcomes within subgroups. Therefore, a meta-analysis that allows for the pooling and quantification of results from different studies is required to overcome this limitation. The present systematic review and meta-analysis was performed to compare the outcomes of OAs with those of CPAP.

Methods

Search Strategy

A computerized search of PubMed (1966 to May 2012), EMBASE (1984 to May 2012), and the Cochrane Controlled Trials Register (2nd quarter, 2012) was carried out. The search strategy consisted of a combination of key words concerning the therapies (continuous positive airway pressure, CPAP, oral appliance, OA) and the disease (obstructive sleep apnea, OSA). These key words were used as MESH headings and free text words. All searches were limited to humans, clinical trial, review and meta-analysis. In addition, manual searching of reference lists from potentially relevant papers was performed, based on

QUICK LOOK

Current knowledge

Both CPAP and oral appliances have been used for the treatment of obstructive sleep apnea. CPAP is more effective at reducing the apnea-hypopnea index, but oral appliances are better tolerated.

What this paper contributes to our knowledge

This meta-analysis of 14 comparative trials suggests that CPAP reduces apnea-hypopnea index and is more effective than oral appliances. There were no differences in treatment usage, treatment preference, side effects, or study withdrawals between CPAP and oral appliance.

the computer-assisted strategy, to identify any additional studies that might have been missed.

Selection of Studies

Using a pre-defined protocol, 2 reviewers (LW and XL) independently selected studies for evaluation. Disagreements were resolved through consensus decision. The inclusion criteria were:

- Compared the outcomes of an OA versus CPAP in the treatment of patients with OSA

- Prospective and randomized

- Published in English and full-text available

- All data were included only once (replication was not permitted). Trials with nonclinical outcomes (eg, cephalometry) were excluded.

Data Extraction

Two reviewers independently performed the data extraction. For each trial, the following items were collected: first author, year of publication, design of the study, subject demography (number, mean age, and sex ratio), details of the inclusion criteria, types of OAs, types of CPAP devices, and study duration. The relevant outcomes pooled in this analysis included Epworth Sleepiness Scale (ESS), health-related quality of life, cognitive performance, blood pressure, apnea-hypopnea index (AHI), arousal index, minimum S_{pO_2}, percent rapid eye movement sleep, treatment usage (including nights/week and hours/night), side effects, subject preference, and withdrawals.

Heterogeneity

A test for heterogeneity (Cochrane Q) was performed to identify inconsistency in the study results. However, because the test is susceptible to the number of trials included in the meta-analysis, we also calculated I^2. This statistic, which is directly calculated from the Q statistic, describes the percentage of variation across the studies that is due to heterogeneity rather than change. I^2 ranges from 0% to 100%, with 0% indicating the absence of any heterogeneity. Although absolute numbers for I^2 are not available, values < 50% are considered low heterogeneity. When I^2 is < 50%, low heterogeneity is assumed, and the effect is thought to be due to change. Conversely, when I^2 exceeds 50%, then heterogeneity is thought to exist and the effect is random.

Assessment of Risk of Bias

Two independent investigators evaluated the risk of bias of the included studies according to the Collaboration's recommended tool (section 8.5 in chapter 8).[16] Briefly, the risk of bias of each study was assessed by using the following methodological components: randomization and generation of the allocation sequence; allocation concealment; subject blinding and examiner blinding; and description of the follow-up. The details of each methodological item are shown in Table 1. However, the subjects knew which treatment they received, because the appearance of the OAs were obviously different from that of the CPAP devices, and it was impossible to make these treatment devices look alike, so blinding and allocation concealment could not be easily performed, and thus the trials with an adequate method of randomization and clear description of the follow-up were considered to be of low risk of bias. Besides, a particular concern with the crossover trials is the risk of a carry-over effect, which occurs when an intervention given in the first period has an effect that carries over into the second period and may influence subjects' responses in the subsequent period. Therefore, crossover trials with a wash-out period between treatments were generally regarded as having low risk of a carry-over effect.[17]

Statistical Analysis

We conducted the meta-analysis with statistics software (Revman 5.1, Cochrane Collaboration, Oxford, United Kingdom). Results are expressed as risk ratios and/or odds ratios with 95% CIs for dichotomous outcomes, and as mean differences with 95% CIs for continuous outcomes. For crossover trials the analyses for continuous outcome variables were conducted by using the generic inverse variance statistical method, where mean differences and stan-

Table 1. Methodological Variables

Randomization
  Adequate: referred to a random number table; used a computer random number generator; coin toss; shuffled cards or envelopes; threw dice; drew lots; minimization.
  Unclear: insufficient information about the sequence generation process to permit judgment of low risk or high risk.
  Inadequate: sequence generated by odd or even date of birth; sequence generated by some rule based on date of admission; sequence generated by some rule based on hospital or clinic record number.

Allocation concealment
  Adequate: central allocation; sequentially numbered drug containers of identical appearance; sequentially numbered, opaque, sealed envelopes.
  Unclear: insufficient information to permit judgment of low risk or high risk.
  Inadequate: used an open random allocation schedule; assignment envelopes used without appropriate safeguards; alternation or rotation; date of birth; case record number; any other explicitly unconcealed procedure.

Patient blinding
  Adequate: no blinding or incomplete blinding, but the review authors judge that the outcome was not likely to be influenced by lack of blinding; blinding of subjects and key study personnel ensured, and unlikely that the blinding could have been broken.
  Unclear: insufficient information to permit judgment of low risk or high risk.
  Not performed, if the trial was not double blind.

Examiner blinding
  Adequate: no blinding of outcome assessment, but the review authors judge that the outcome measurement was not likely to be influenced by lack of blinding; blinding of outcome assessment ensured, and unlikely that the blinding could have been broken.
  Unclear: insufficient information to permit judgment of low risk or high risk.
  Not performed, if the trial was not double blind.

Withdrawals and Dropouts
  Adequate: the number of and reasons for dropouts and withdrawals in all intervention groups were described, or it was specified that there were no dropouts or withdrawals.
  Unclear: insufficient reporting of attrition/exclusions to permit judgment of low risk or high risk (eg, number randomized not stated, no reasons provided for missing data).
  Inadequate: the number of and reasons for dropouts and withdrawals were not described.

dard errors were entered. A correlation coefficient of 0.5 was used throughout this meta-analysis to estimate the standard errors for some crossover trials, where the appropriate standard deviation of differences was not included in study reports. Sensitivity analyses were performed to assess the impact of the assumed correlation coefficient on the outcomes of meta-analyses by repeating the analyses assuming correlation coefficients of 0.3 and 0.7, respectively. A fixed effects model was initially used; however, we planned to use a random effects model if there was evidence of significant heterogeneity across trials ($P < .10$
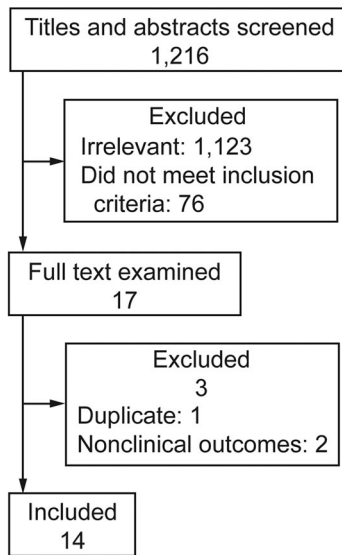
Fig. 1. Flow of study identification, exclusion, and inclusion.

and $I^2 > 50\%$). A sensitivity analysis was performed to explore the potential source of heterogeneity. We also planned to use funnel plot asymmetry to assess for publication bias.

Results

Figure 1 shows the details of study identification, inclusion, and exclusion. The literature search yielded 1,216 articles. By screening the titles and abstracts, 1,123 papers were excluded due to the irrelevance to this topic. In 93 potentially relevant references, 17 papers were taken for a comprehensive evaluation. After retrieving the full articles, one was excluded because of duplicated data[18] and 2 were excluded because of nonclinical outcomes.[19,20] Finally, 14 studies were included in this meta-analysis.[10-15,21-28] Among these studies, 8 trials had a crossover design,[10-12,15,21-23,27] and 6 had a parallel-group design.[13,14,24-26,28] The main characteristics of the included studies are shown in Table 2.

Risk of Bias in These Trials

The assessment of risk of bias in all included studies is shown in Table 3. A method of block randomization was used in 6 trials.[14,22,24-26,28] Adequate method for allocation concealment was applied in 3 trials.[14,27,28] In one trial,[28] the subjects remained blinded to the nature of therapy, and examiner blinding was performed in 3 trials.[22,27,28] The description of follow-up was considered adequate in all included trials. So over half of the included trials, with 3 or more methodological components inadequate or unclear, were regarded as having a high risk of bias. A wash-

out period between treatments was described in 6 of the 8 crossover trials,[10,12,15,21,23,27] and those 6 were considered to have a low risk of carry-over effect.

Clinical Outcomes

Score of Epworth Sleepiness Scale. There were 5 crossover trials[12,15,21-23] and 3 parallel-group trials[13,14,24] reporting the score of ESS. The test for heterogeneity detected a significant heterogeneity across the crossover trials ($P < .001$, $I^2 = 88\%$), while there was no evidence of heterogeneity across the parallel-group trials ($P = .85$, $I^2 = 0\%$). Pooled estimates revealed that there was no significant difference between treatments, both in crossover trials (mean difference 0.74, 95% CI −0.69 to 2.17, $P = .31$) and in parallel-group trials (mean difference 1.33, 95% CI −0.19 to 2.85, $P = .09$) (Fig. 2).

Health-Related Quality of Life. Two crossover trials[12,22] reported the data on the Functional Outcomes of Sleepiness Questionnaire, and pooled estimates showed no significant difference between groups (mean difference −0.43, 95% CI −1.41 to 0.54, $P = .38$), but with a significant heterogeneity across the trials ($P = .008$, $I^2 = 86\%$) (see Fig. 2). There were 2 crossover trials[12,22] and 3 parallel-group trials[13,14,28] reporting the outcome of the 36-item Medical Outcomes Study Short Form questionnaire (SF-36). In the trial by Engleman et al,[22] CPAP-treated subjects had significantly higher SF-36 scores for the health transition and mental components ($P = .001$ and .008, respectively), while there were no significant difference in the SF-36 physical component scores between treatments. Barnes et al[12] reported that there was no significant difference regarding SF-36 mean scores between the groups. Two parallel-group trials[13,14] reported component scores from the SF-36, and pooled estimates revealed no significant difference regarding each component score between treatments. Aarab et al[28] reported that the changes in the domains of the SF-36 were not significantly different between groups. One crossover trial[22] and one parallel-group trial[14] reported the data on Hospital Anxiety and Depression Scale, both of which showed that the Hospital Anxiety and Depression Scale scores did not differ significantly between treatments.

Cognitive Performance. There were 3 crossover trials reporting the outcome of cognitive performance. In the trial by Engleman et al[22] it was shown that no significant differences in performance intelligence-quotient decrement score, Trail Making Test B, Steer Clear Performance Test, and Paced Auditory Serial Addition Test 2s correct was detected between treatments. Barnes et al[12] reported that there was no significant difference regarding the cognitive function assessed by Paced Auditory Serial Addition

Table 2. Main Characteristics of Included Studies

First Author	Year	Study Design	n	Age y	Male %	Inclusion Criteria AHI, events/h	Oral Appliance	CPAP Device(s)	Study Duration
Ferguson[10]	1996	Crossover	27	46.2	88.9	15–50	Anterior mandibular positioner (Snore-Guard, Hays & Meade, Albuquerque, New Mexico)	Nasal CPAP (REMstar Choice or Tranquility Plus, Respironics, Murrysville, Pennsylvania)	2 × 16 weeks Washout period 2 weeks
Ferguson[21]	1997	Crossover	24	44.0	79.2	15–50	Anterior mandibular positioner (Snore-Guard)	Nasal CPAP (REMstar Choice or Tranquility Plus)	2 × 16 weeks Washout period 2 weeks
Randerath[11]	2002	Crossover	20	56.5	80.0	5–30	ISAD (intraoral sleep apnea device) (IST, Hinz, Herne, Germany)	CPAP (Max II, MAPData, ResMed, San Diego, California) (Somnotron, Weinmann, Hamburg, Germany) (Vector, Hoffrichter, Schwerin, Germany)	2 × 6 weeks Washout period 0 weeks
Engleman[22]	2002	Crossover	48	46.0	75.0	≥ 5 ESS score ≥ 8	Mandibular repositioning splints	CPAP Device not stated	2 × 8 weeks Washout period 0 weeks
Tan[23]	2002	Crossover	24	50.9	83.3	≤ 50	Mandibular advancement splint (Erkodent, Tuttlingen, Germany)	Nasal CPAP (REMstar Choice) (Sullivan Elite, ResMed, San Diego, California)	2 × 16 weeks Washout period 2 weeks
Barnes[12]	2004	Crossover	114	47.0	80.0	5–30	Mandibular advancement splint (Medical Dental Sleep Appliance, RJ and VK Bird, Middle Park, Victoria, Australia)	Nasal CPAP (Sullivan Elite)	3 × 12 weeks Washout period 2 weeks
Hoekema[24]	2007	Parallel	10	47.6	77.7	> 5	Oral appliance (Thornton Adjustable Positioner type 1, Airway Management, Dallas, Texas)	CPAP (PV10, Breas, Mölnlycke, Sweden)	12 weeks
			10	49.7	90.0				
Hoekema[25]	2007	Parallel	21	48.0	100.0	≥ 5	Oral appliance (Thornton) Adjustable Positioner type 1	CPAP (PV10)	12 weeks
			27	51.0	100.0				
Lam[13]	2007	Parallel	34	45.0	76.5	5–40 ESS score > 9	Oral appliance made of dental acrylic modified from a Harvold-type functional activator	CPAP (ARIA LX, Respironics, Murrysville, Pennsylvania)	10 weeks
			34	45.0	79.4				
Hoekema[14]	2008	Parallel	51	48.8	84.3	≥ 5	Oral appliance (Thornton) Adjustable Positioner type 1	CPAP (PV10)	12 weeks
			52	49.4	94.2				
Hoekema[26]	2008	Parallel	15	89.3	49.7	> 20	Oral appliance (Thornton) Adjustable Positioner type 1	CPAP (PV10)	12 weeks
			13						
Gagnadoux[15]	2009	Crossover	59	50.3	86.8	10–60	Mandibular advancement device (AMC, Artech Medical, Pantin, France)	CPAP (Sullivan S6 Elite, ResMed, San Diego, California)	2 × 8 weeks Washout period 1 week
Trzepizur[27]	2009	Crossover	12	46.0	100.0	≥ 15	Mandibular advancement device (AMC)	CPAP (Sullivan S6 Elite)	2 × 8 weeks Washout period 1 week
					100.0				
Aarab[18]	2011	Parallel	21	50.4	81.0	5–45 ESS score ≥10	Mandibular advancement device with adjustable protrusive mandibular positioner at a constant vertical dimension	Nasal CPAP (REMstar Pro)	48 weeks
			22	54.9	68.2				

AHI = apnea and hypopnea index
ESS = Epworth Sleepiness Scale

Test 1.2 between CPAP and OA. Gagnadoux et al[15] reported that CPAP and OAs both improved the trail making test A and trail making test B, but with no significant difference between groups.

Table 3. Risk of Bias in the Included Studies

| First Author	Year	Randomization	Allocation Concealment	Patient Blinding	Examiner Blinding	Follow-up
Ferguson[10]	1996	Yes/unclear	Unclear	Unclear	Unclear	Clear report
Ferguson[21]	1997	Yes/unclear	Unclear	Unclear	Unclear	Clear report
Randerath[11]	2002	Yes/unclear	Unclear	Unclear	Unclear	Clear report
Engleman[22]	2002	Yes/adequate	Unclear	Unclear	Yes/adequate	Clear report
Tan[23]	2002	Yes/unclear	Unclear	Unclear	Not performed	Clear report
Barnes[12]	2004	Yes/inadequate	Unclear	Unclear	Unclear	Clear report
Hoekema[24]	2007	Yes/adequate	Unclear	Not performed	Not performed	Clear report
Hoekema[25]	2007	Yes/adequate	Unclear	Not performed	Not performed	Clear report
Lam[13]	2007	Yes/unclear	Unclear	Unclear	Unclear	Clear report
Hoekema[14]	2008	Yes/adequate	Adequate	Not performed	Not performed	Clear report
Hoekema[26]	2008	Yes/adequate	Unclear	Not performed	Not performed	Clear report
Gagnadoux[15]	2009	Yes/unclear	Unclear	Unclear	Unclear	Clear report
Trzepizur[27]	2009	Yes/unclear	Adequate	Unclear	Yes/adequate	Clear report
Aarab[18]	2011	Yes/adequate	Adequate	Blinded to the nature of therapy	Yes/adequate	Clear report
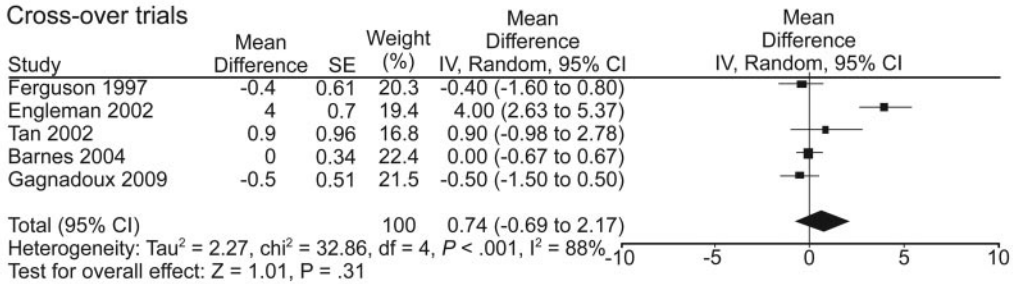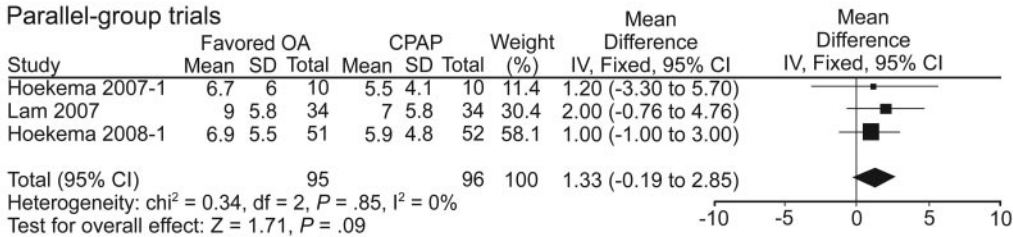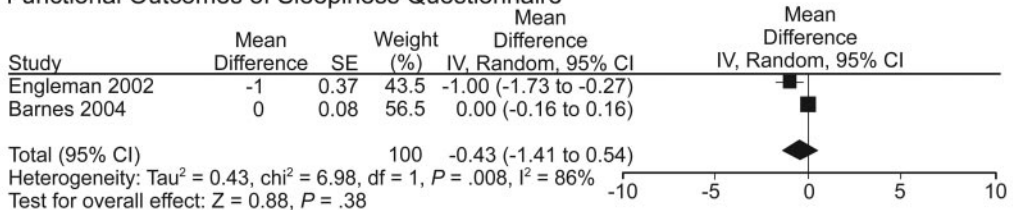
Fig. 2. Forest plot of the meta-analysis of clinical outcomes. IV = inverse variance.

Blood Pressure. Two crossover trials and one parallel-group trial reported the outcome of blood pressure. In the trial by Engleman et al[22] it was reported that, although there was no significant response in the 24-hour mean systolic or diastolic blood pressure between groups, OA-treated subjects had a significantly lower nighttime dia-

stolic blood pressure ($P < .05$). However, Trzepizur et al[27] and Lam et al[13] both reported no significant difference in blood pressure change between OA and CPAP.

Polysomnography Outcomes

Apnea Hypopnea Index. There were 6 crossover trials[10-12,21-23] and 3 parallel-group trials[13,14,28] reporting AHI. The test for heterogeneity revealed a significant heterogeneity across the crossover trials ($P = .008$, $I^2 = 68\%$), while no evidence of heterogeneity was detected across the parallel-group trials ($P = .41$, $I^2 = 0\%$). Overall, OA-treated subjects had significantly more apneas and hypopneas, both in the crossover trials (mean difference 8.25, 95% CI 5.89–10.61, $P < .001$) (Fig. 3) and the parallel-group trials (mean difference 5.96, 95% CI 3.40 to 8.51, $P < .001$) (see Fig. 3).

Arousal Index. Five crossover trials[10-12,21,23] and 2 parallel-group trials[13,28] reported arousal index. Pooled estimates of the crossover trials found a significant difference in favor of CPAP (mean difference 3.10, 95% CI 1.23–4.96, $P = .001$), but with a significant heterogeneity across the trials ($P = .06$, $I^2 = 55\%$) (see Fig. 3). While the pooled estimates of the parallel-group trials revealed no significant difference between treatments (mean difference 3.18, 95% CI –1.17 to 7.52, $P = .15$), the results were robust and there was no heterogeneity across the trials ($P = .32$, $I^2 = 0\%$) (see Fig. 3).

Minimum S_{pO_2}. There were 4 crossover trials[10-12,21] and 4 parallel-group trials[13,14,25,26] reporting minimum S_{pO_2}. Pooled estimates of crossover trials showed that OA-treated subjects had a significantly lower minimum S_{pO_2} (mean difference −5.11%, 95% CI −6.91 to –3.30, $P < .001$), but there was substantial heterogeneity across the trials ($P = .003$, $I^2 = 78\%$) (see Fig. 3). However, pooled estimates of the parallel-group trials found no significant difference between treatments (mean difference −0.94, 95% CI –2.50 to 0.62, $P = .24$), with no evidence of heterogeneity across the trials ($P = .14$, $I^2 = 45\%$) (see Fig. 3).

Rapid Eye Movement Sleep. Four crossover trials[10,11,21,23] and 2 parallel-group trials[14,28] reported the outcome of percent rapid eye movement sleep. The test for heterogeneity detected significant heterogeneity across the crossover trials ($P < .001$, $I^2 = 84\%$), and pooled estimates found no significant difference between treatments (mean difference −0.27, 95% CI –3.75 to 3.22, $P = .88$) (see Fig. 3). While the pooled estimates of the parallel-group trials showed a significant difference in favor of CPAP (mean difference 2.42, 95% CI 0.31 to 4.53, $P = .02$),

there was no heterogeneity across the trials ($P = .97$, $I^2 = 0\%$) (see Fig. 3).

Other Related Outcomes

Treatment Usage. There were 2 crossover trials[12,22] reporting treatment usage of hours/night, and 4 parallel-group trials[13,14,25,26] reporting treatment usage of hours/night and nights/week. There was significant heterogeneity across the crossover trials ($P < .001$, $I^2 = 95\%$, hours/night) and across the parallel-group trials ($P < .001$ and $I^2 = 93\%$ for hours/night, and $P = .09$ and $I^2 = 54\%$ for nights/week). The pooled estimates showed no significant difference between treatments both in the crossover trials (mean difference 1.01, 95% CI −0.75 to 2.78, and $P = .26$ for hours/night) (Fig. 4) and in the parallel-group trials (mean difference 0.82, 95% CI −0.27 to 1.91, and $P = .14$ for hours/night, and mean difference 0.16, 95% CI −0.08 to 0.40, and $P = .19$ for nights/week) (see Fig. 4).

Treatment Preference. Five crossover trials reported the outcome of treatment preference. The trials by Ferguson et al[10,21] reported that most subjects who were treated successfully preferred CPAP to OAs, but OA was preferred as a long-term treatment among these subjects. Engleman et al[22] reported that a 5-variable model explained 68% of the variance in treatment preference, and eventually identified 83% and 90% of subjects, respectively, preferring OA and CPAP. Barnes et al[12] reported that the overall percentages of preferred treatment were 30% for OA and 44% for CPAP. While Gagnadoux et al[15] reported that 71.2% of subjects preferred OA, 8.5% preferred CPAP, and 8 subjects had no treatment preference.

Side Effects. There were 6 crossover trials and 2 parallel-group trials reporting side effects. Ferguson et al[10,21] found mild side effects common with OA, including sore teeth, sore jaw muscles, and excessive salivation, while CPAP-treated subjects more commonly had moderate to severe side effects, such as nasal congestion, rhinorrhea, eye irritation, and a sense of suffocation. Randerath et al[11] reported that CPAP-treated subjects often had a sense of pressure on the face, while OA-treated subjects often had early morning discomfort in the mouth. Engleman et al[22] reported that side effects were common both for OAs and CPAP; however, some side effects were treatment-specific, such as dental pain or salivation with OA, and stuffy nose or mask problems with CPAP. Gagnadoux et al[15] reported that the mean side effects scores were similar for OA and CPAP in the subjects who completed the study ($P = .80$). Lam et al[13] and Aarab et al[28] both reported that
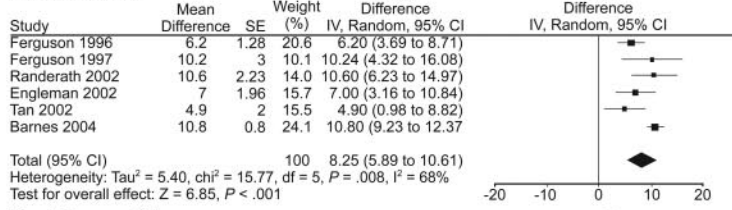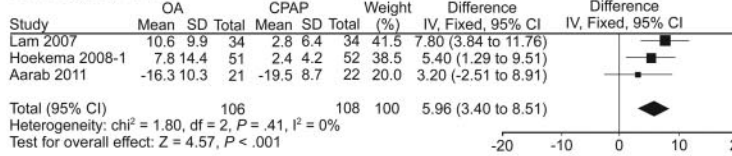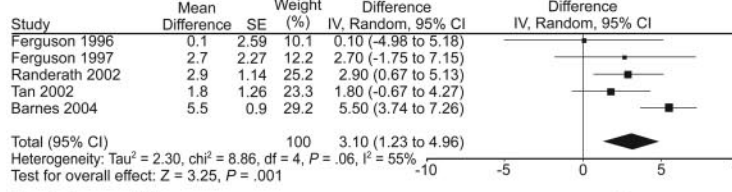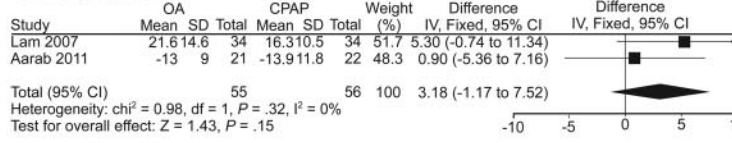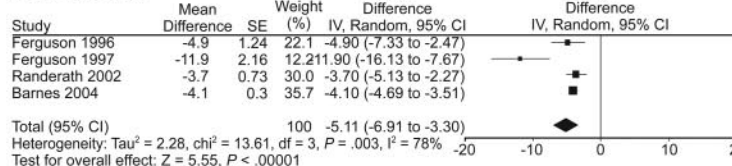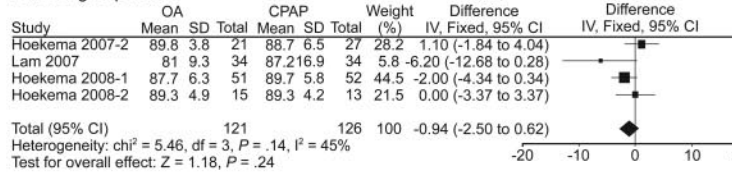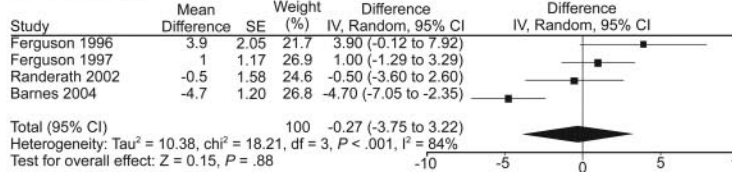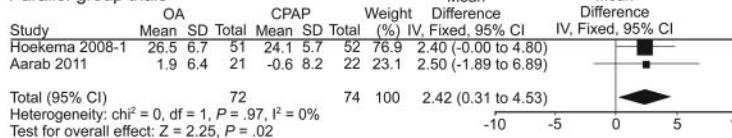
Fig. 3. Forest plot of the meta-analysis of polysomnography outcomes. IV = inverse variance.
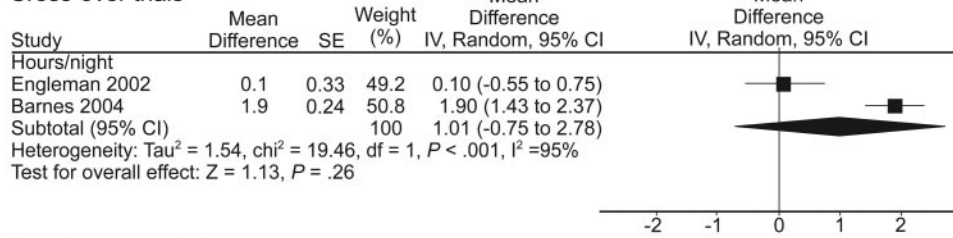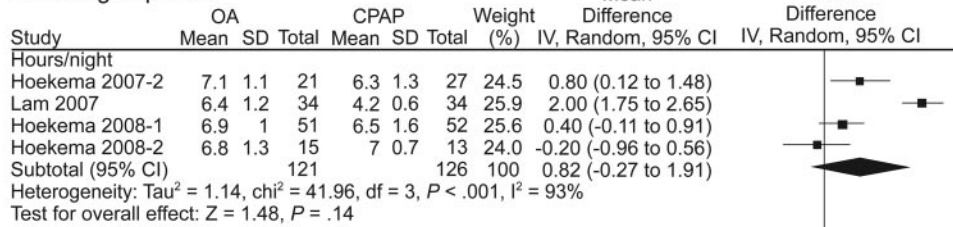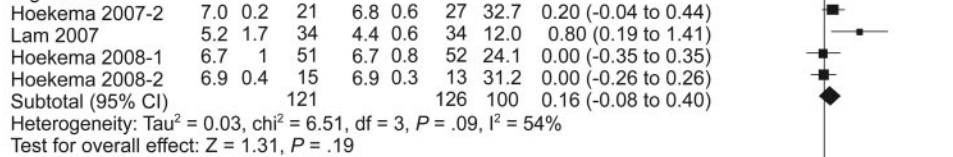
Fig. 4. Forest plot of the meta-analysis of other related outcomes. IV = inverse variance. M-H = Mantel-Haenszel.

nearly all subjects reported side effects, but that all side effects were considered mild and acceptable.
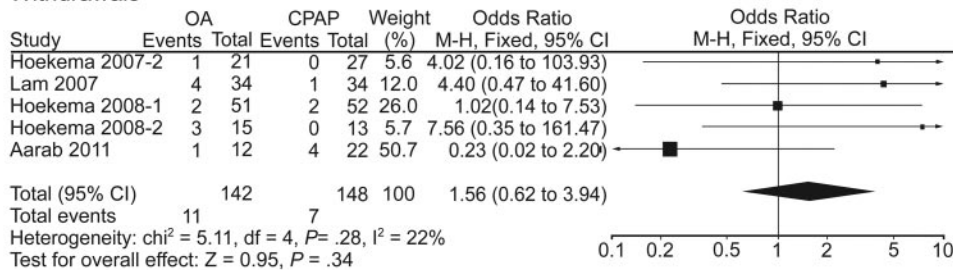
Withdrawals. Ferguson et al[10] reported that one subject dropped out during the wash-in period and one dropped out in the OA treatment period. In another trial by Ferguson et al,[21] one subject withdrew in the OA treatment period and 3 refused to cross over to the CPAP treatment arm. Tan et al[23] and Gagnadoux et al[15] both reported that 2 subjects dropped out in the CPAP treatment period and one withdrew in the OA treatment period. Trzepizur et al[27] reported one dropout from the CPAP group. Five parallel-group trials reported withdrawals,[13,14,25,26,28] and the pooled estimates showed no significant difference between treatments (odds ratio 1.56, 95% CI 0.62–3.94, P = .34), with no evidence of heterogeneity (P = .28, I² = 22%) (see Fig. 4).

Discussion

Although many randomized trials have supported the evidence that CPAP is more effective than OAs in reducing OSA,[10-15] some studies have suggested that the efficacy of OAs in modifying the health risks associated with OSA is somewhat similar to that of CPAP.[29] A previous Cochrane review[8] showed that OAs were less effective than CPAP in reducing AHI and improving minimum S_{pO_2} during sleep; however, subjects seemed to be more adherent to OAs than CPAP in a small part of the included trials. So the prior Cochrane review led to the conclusion that CPAP seemed to be more effective than OAs in improving sleep-disordered breathing, but the difference in symptoms between these 2 treatments was not significant.

However, the results of the prior Cochrane review were not completely convincing. The main reason was that nearly

Table 4.   Sensitivity Analysis for the Heterogeneity of Crossover Trials

	Study That Was the Major Contributor to the Heterogeneity	Heterogeneity After Excluding the Study			Pooled Estimates of the Remaining Studies		
		Chi-square	P	I^2 %	Mean Difference (95% CI)	Z	P
Epworth Sleepiness Scale	Engleman[22]	2.02	.57	0	−0.12 (−0.61 to 0.36)	0.50	.62
Apnea-hypopnea index	Barnes[12]	5.23	.26	24	7.05 (5.41–8.68)	8.43	< .001
Arousal index	Barnes[12]	1.18	.76	0	2.24 (0.76–3.73)	2.96	.003
Minimum S$_{pO_2}$	Ferguson[21]	0.71	.70	0	−4.08 (−4.61 to −3.55)	15.08	< .001
Percent rapid eye movement sleep	Tan[23]	2.90	.23	31	1.07 (−0.61 to 2.74)	1.25	.21

all the included crossover trials in the prior Cochrane review did not report any paired results, though they did provide means and SDs for the outcomes of each treatment. In this case, paired analyses from these crossover trials could only be approximated by assuming a certain degree of correlation between the 2 treatment outcomes. The reviewers used a correlation coefficient of zero, which was equivalent to a parallel group analysis of the results. The limitation by doing this involved the fact that the particular strength of crossover design was ignored, where treatments were evaluated on the same subjects, allowing comparison at the individual rather than the group level.[30] Therefore, the validity of the pooled estimates of the crossover trials in the prior Cochrane review needed further confirmation.

In the present systematic review and meta-analysis, a correlation coefficient of 0.5 was used to estimate the standard errors for some included crossover trials, where the appropriate standard deviations of the differences were not included in the study reports. Our results demonstrated that OA-treated subjects had similar ESS scores, when compared to CPAP-treated subjects, both in crossover ($P = .31$) and parallel-group ($P = .09$) trials. In addition, the effects of OAs and CPAP on health-related quality of life, cognitive performance, and blood pressure were also similar, but CPAP may produce a better outcome than OA. From the above results, it can be concluded that OAs and CPAP yield fairly similar results in terms of these clinical outcomes.

As for the polysomnography (PSG) outcomes, the pooled estimates of the crossover trials suggested a significant difference in favor of CPAP with regard to AHI ($P < .001$), arousal index ($P = .001$), and minimum S$_{pO_2}$ ($P < .001$). The pooled estimates of parallel-group trials showed that, compared with CPAP-treated subjects, OA-treated subjects had significantly more apneas and hypopneas ($P < .001$). Moreover, a significant difference in favor of CPAP was detected regarding percent rapid eye movement sleep ($P = .02$). The above results may lead to a conclusion that CPAP yielded better PSG outcomes than OAs,

especially in reducing AHI. Treatment success with OA, defined as an AHI of < 5 events/hour, was found in 19–75% of the subjects. An AHI of < 10 events/hour was reported in 30–94% of the subjects. However, CPAP reduced AHI more efficiently and gave a higher success rate in all these studies.[10-15,21-23] Overall, these PSG outcomes (especially hypoxia) are of crucial importance with regard to survival and morbidity in subjects with OSA, which emphasizes the relevance of optimal suppression of respiratory disturbances and argues against OA's treatment effect on OSA in terms of these respiratory parameters,[31,32] indicating that OAs can be given only for those who refuse CPAP.

As far as other related outcomes were concerned, the pooled estimates revealed no significant difference between OA and CPAP with regard to treatment usage ($P = .26$ for hours/night in the crossover trials, and $P = .14$ for hours/night and $P = .19$ for nights/week in the parallel-group trials) and withdrawals ($P = .34$ in the parallel-group trials).

Side effects were common with both OAs and CPAP, with a similar severity across treatments. Moreover, although the subjects generally preferred OA to CPAP, similar preferences for OA and CPAP were reported. The above results lead us to draw the conclusion that OA-treated subjects have similar results regarding all these related outcomes, when compared with CPAP-treated subjects.

Although a substantial heterogeneity was detected in all pooled estimates of crossover trials, it could be eliminated by the sensitivity analysis. After excluding the major contributors to the heterogeneity, the pooled estimates still got the same results regarding all these outcomes (Table 4), indicating that the heterogeneity did not have a significant effect on the pooled estimates of the crossover trials. In addition, a sensitivity analysis performed by repeating the analyses assuming correlations of 0.3 and 0.7 revealed that the assumed correlation of 0.5 did not affect the pooled estimates of the crossover trials with regard to all these outcomes (Table 5). While the heterogeneity across the

Table 5. Sensitivity Analysis for Crossover Trials, With 2 Assumed Correlation Coefficients

	Number of Trials	Number of Subjects	Pooled Estimates			Heterogeneity		
			Mean Difference (95% CI)	Z	P	Chi-square	P	I^2 %
Assuming a Correlation Coefficient of 0.3								
Epworth Sleepiness Scale	5	272	0.77 (–0.67 to 2.20)	1.05	.30	54.99	< .001	93
Functional Outcomes of Sleepiness Questionnaire	2	165	–0.41 (–1.38 to 0.55)	0.84	.40	5.13	.02	81
Apnea-hypopnea index	6	260	8.23 (5.85–10.62)	6.76	< .001	18.93	.002	74
Arousal index	5	209	3.59 (2.25–4.93)	5.26	< .001	6.40	.17	38
Minimum S_{pO_2}	4	185	–5.01 (–6.60 to –3.42)	6.17	< .001	15.41	.002	81
Percent rapid eye movement sleep	4	95	–0.24 (–3.69 to 3.21)	0.13	.89	28.57	< .001	90
Treatment usage, h/night	2	165	1.01 (–0.76 to 2.77)	1.12	.26	31.24	< .001	97
Assuming a Correlation Coefficient of 0.7								
Epworth Sleepiness Scale	5	272	0.72 (–0.72 to 2.15)	0.98	.33	23.37	< .001	83
Functional Outcomes of Sleepiness Questionnaire	2	165	–0.46 (–1.43 to 0.52)	0.92	.36	10.54	.001	91
Apnea-hypopnea index	6	260	8.27 (5.93 to 10.62)	6.91	< .001	13.67	.02	63
Arousal index	5	209	2.94 (1.08 to 4.80)	3.10	.002	14.12	.007	72
Minimum S_{pO_2}	4	185	–5.14 (–7.08 to –3.21)	5.20	< .001	12.17	.007	75
Percent rapid eye movement sleep	4	95	–0.30 (–3.80 to 3.21)	0.17	.87	13.41	.004	78
Treatment usage, h/night	2	165	1.02 (–0.74 to 2.78)	1.13	.26	14.06	< .001	93

parallel-group trials was slight, most of the evidence from the analyses should be considered robust. But there was substantial heterogeneity in the analysis of treatment usage, the major contributor to heterogeneity was the study by Lam et al.[13] By removing this study the heterogeneity was eliminated ($P = .16$ and $I^2 = 46\%$ for hours/night, and $P = .47$ and $I^2 = 0\%$ for nights/week), and the pooled estimates indicated a significant difference in favor of OA in treatment usage of hours/night ($P = .04$), but there was still no significant difference in treatment usage of nights/week between treatments ($P = .29$).

Nevertheless, the present systematic review and meta-analysis still had several potential limitations. One potential limitation was that the degree of mandibular advancement by the OAs and the OA designs were variable among the included trials, which caused uncertainty regarding comparisons between these studies. A second potential limitation involved the fact that most of the included trials had high risk of bias, due to 3 or more unclear or inadequate methodological components. Moreover, a few of the included crossover trials[11,22] had a high risk of a carryover effect, due to the absence of a wash-out period between treatments.

A third limitation was the small sample sizes of all the included trials and the small number of studies. A funnel plot for pooled estimates to assess the potential publication bias was not performed, and unpublished studies with negative results could not be identified, so there might be publication bias as well, which could result in overestimation of the effectiveness of the interventions.

Conclusions

CPAP has yielded better PSG outcomes than OA, especially lower AHI, indicating that OAs are less effective than CPAP in improving sleep-disordered breathing. If sleep-disordered breathing is left inadequately controlled, the long-term risk of systemic morbidity associated with OSA may be substantial, which may also suggest that CPAP is more reliable than OA in treatment of OSA. However, similar results from OA and CPAP in terms of clinical and other related outcomes were found, including ESS score, health-related quality of life, and treatment usage. Based on this evidence it would appear proper to offer OA to patients who are unable or unwilling to persist with CPAP.

REFERENCES

1. Ip MS, Lam B, Tang LC, Lauder IJ, Ip TY, Lam WK. A community study of sleep-disordered breathing in middle-aged Chinese women in Hong Kong: prevalence and gender differences. Chest 2004;125(1): 127-134.
2. Cistulli PA, Gotsopoulos H, Marklund M, Lowe AA. Treatment of snoring and obstructive sleep apnea with mandibular repositioning appliances. Sleep Med Rev 2004;8(6):443-457.
3. Giles TL, Lasserson TJ, Smith BJ, White J, Wright J, Cates CJ. Continuous positive airways pressure for obstructive sleep apnoea in adults. Cochrane Database Syst Rev 2006;(1):CD001106.
4. Pépin JL, Leger P, Veale D, Langevin B, Robert D, Lévy P. Side effects of nasal continuous positive airway pressure in sleep apnea syndrome. Study of 193 patients in two French sleep centers. Chest 1995;107(2):375-381.

5. Engleman HM, Wild MR. Improving CPAP use by patients with the sleep apnoea/hypopnoea syndrome (SAHS). Sleep Med Rev 2003; 7(1):81-99.

6. McArdle N, Devereux G, Heidarnejad H, Engleman HM, Mackay TW, Douglas NJ. Long-term use of CPAP therapy for sleep apnea/hypopnea syndrome. Am J Respir Crit Care Med 1999;159(4 Pt 1):1108-1114.

7. Almeida FR, Henrich N, Marra C, Lynd LD, Lowe AA, Tsuda H, et al. Patient preferences and experiences of CPAP and oral appliances for the treatment of obstructive sleep apnea: a qualitative analysis. Sleep Breath 2012 [Epub ahead of print].

8. Lim J, Lasserson TJ, Fleetham J, Wright J. Oral appliances for obstructive sleep apnoea. Cochrane Database Syst Rev 2009;(3): CD004435.

9. Ng AT, Gotsopoulos H, Qian J, Cistulli PA. Effect of oral appliance therapy on upper airway collapsibility in obstructive sleep apnea. Am J Respir Crit Care Med 2003;168(2):238-241.

10. Ferguson KA, Ono T, Lowe AA, Keenan SP, Fleetham JA. A randomized crossover study of an oral appliance vs nasal-continuous positive airway pressure in the treatment of mild-moderate obstructive sleep apnea. Chest 1996;109(5):1269-1275.

11. Randerath WJ, Heise M, Hinz R, Ruehle KH. An individually adjustable oral appliance vs continuous positive airway pressure in mild-to-moderate obstructive sleep apnea syndrome. Chest 2002; 122(2):569-575.

12. Barnes M, McEvoy RD, Banks S, Tarquinio N, Murray CG, Vowles N, et al. Efficacy of positive airway pressure and oral appliance in mild to moderate obstructive sleep apnea. Am J Respir Crit Care Med 2004;170(6):656-664.

13. Lam B, Sam K, Mok WY, Cheung MT, Fong DY, Lam JC, et al. Randomised study of three non-surgical treatments in mild to moderate obstructive sleep apnoea. Thorax 2007;62(4):354-359.

14. Hoekema A, Stegenga B, Wijkstra PJ, van der Hoeven JH, Meinesz AF, de Bont LG. Obstructive sleep apnea therapy. J Dent Res 2008; 87(9):882-887.

15. Gagnadoux F, Fleury B, Vielle B, Pételle B, Meslier N, N'Guyen XL, et al. Titrated mandibular advancement versus positive airway pressure for sleep apnoea. Eur Respir J 2009;34(4):914-920.

16. Higgins JPT, Altman DG, Sterne JAC; Cochrane Statistical Methods Group and Cochrane Bias Methods Group. Assessing risk of bias in included studies. In: Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions, version 5.1.0. (updated March 2011). The Cochrane Collaboration; 2011. http://www.cochrane-handbook.org. *Accessed April 18, 2013.*

17. Elbourne DR, Altman DG, Higgins JP, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: methodological issues. Int J Epidemiol 2002;31(1):140-149.

18. Aarab G, Lobbezoo F, Heymans MW, Hamburger HL, Naeije M. Long-term follow-up of a randomized controlled trial of oral appliance therapy in obstructive sleep apnea. Respiration 2011;82(2):162-168.

19. Doff MH, Hoekema A, Pruim GJ, Huddleston Slater JJ, Stegenga B. Long-term oral-appliance therapy in obstructive sleep apnea: a cephalometric study of craniofacial changes. J Dent 2010;38(12):1010-1018.

20. Arya D, Tripathi A, Singh SV, Tripathi S, Nagar A, Mishra A. A pilot study to evaluate posttreatment cephalometric changes in subjects with OSA. J Prosthet Dent 2010;103(3):170-177.

21. Ferguson KA, Ono T, Lowe AA, al-Majed S, Love LL, Fleetham JA. A short-term controlled trial of an adjustable oral appliance for the treatment of mild to moderate obstructive sleep apnoea. Thorax 1997; 52(4):362-368.

22. Engleman HM, McDonald JP, Graham D, Lello GE, Kingshott RN, Coleman EL, et al. Randomized crossover trial of two treatments for sleep apnea/hypopnea syndrome: continuous positive airway pressure and mandibular repositioning splint. Am J Respir Crit Care Med 2002;166(6):855-859.

23. Tan YK, L'Estrange PR, Luo YM, Smith C, Grant HR, Simonds AK, et al. Mandibular advancement splints and continuous positive airway pressure in patients with obstructive sleep apnoea: a randomized cross-over trial. Eur J Orthod 2002;24(3):239-249.

24. Hoekema A, Stegenga B, Bakker M, Brouwer WH, de Bont LG, Wijkstra PJ, et al. Simulated driving in obstructive sleep apnoea-hypopnoea; effects of oral appliances and continuous positive airway pressure. Sleep Breath 2007;11(3):129-138.

25. Hoekema A, Stel AL, Stegenga B, van der Hoeven JH, Wijkstra PJ, van Driel MF, et al. Sexual function and obstructive sleep apnea-hypopnea: a randomized clinical trial evaluating the effects of oral-appliance and continuous positive airway pressure therapy. J Sex Med 2007;4(4 Pt 2):1153-1162.

26. Hoekema A, Voors AA, Wijkstra PJ, Stegenga B, van der Hoeven JH, Tol CG, et al. Effects of oral appliances and CPAP on the left ventricle and natriuretic peptides. Int J Cardiol 2008;128(2):232-239.

27. Trzepizur W, Gagnadoux F, Abraham P, Rousseau P, Meslier N, Saumet JL, et al. Microvascular endothelial function in obstructive sleep apnea: impact of continuous positive airway pressure and mandibular advancement. Sleep Med 2009;10(7):746-752.

28. Aarab G, Lobbezoo F, Hamburger HL, Naeije M. Oral appliance therapy versus nasal continuous positive airway pressure in obstructive sleep apnea: a randomized, placebo-controlled trial. Respiration 2011;81(5):411-419.

29. Chan AS, Cistulli PA. Oral appliance treatment of obstructive sleep apnea: an update. Curr Opin Pulm Med 2009;15(6):591-596.

30. Gotzsche PC. Patients' preference in indomethacin trials: an overview. Lancet 1989;1(8629):88-91.

31. Randerath WJ, Verbraecken J, Andreas S, Bettega G, Boudewyns A, Hamans E, et al. Non-CPAP therapies in obstructive sleep apnoea. Eur Respir J 2011;37(5):1000-1028.

32. Marklund M, Verbraecken J, Randerath W. Non-CPAP therapies in obstructive sleep apnoea: mandibular advancement device therapy. Eur Respir J 2012;39(5):1241-1247.