Continuous Positive Airway Pressure Therapy for Treating Sleepiness in a Diverse Population With Obstructive Sleep Apnea

Results of a Meta-analysis

Sanjay R. Patel, MD; David P. White, MD; Atul Malhotra, MD; Michael L. Stanchina, MD; Najib T. Ayas, MD

Background: Although continuous positive airway pressure (CPAP) has become the standard of care in the treatment of obstructive sleep apnea (OSA), 2 systematic reviews have questioned its utility. Since the publication of these reviews, several randomized controlled trials have been reported. We, therefore, performed a meta-analysis to assess the effect of CPAP on subjective and objective sleepiness.

Methods: We conducted a thorough literature search to identify all published randomized controlled trials of CPAP in patients with OSA. Meta-analyses were performed using a random-effects model. Statistical heterogeneity was assessed using the Q statistic.

Results: Twelve trials of CPAP in patients with OSA meeting our inclusion criteria were found. The Epworth Sleepiness Scale score was reported in 11 studies (706 patients). A meta-analysis found that CPAP reduced the Epworth Sleepiness Scale score an average of 2.94 points more than placebo ($P < .001$). The heterogeneity ($Q_{(H)}=57.7, P < .001$) between studies could not be explained by differences in sex composition, mean age, mean body mass index, or country of study. Trials recruiting subjects with severe OSA plus sleepiness (mean apnea-hypopnea index, $\geq 30$ events per hour; and mean Epworth Sleepiness Scale score, $\geq 11$) had a greater decrease in the Epworth Sleepiness Scale score than the other studies ($4.75$ vs $1.10$; $P < .001$). Objective measures of sleepiness were reported in 8 trials (482 subjects). Continuous positive airway pressure increased sleep onset latency by 0.93 minute ($P = .04$) more than placebo.

Conclusions: Continuous positive airway pressure therapy significantly improves subjective and objective measures of sleepiness in patients with OSA across a diverse range of populations. Patients with more severe apnea and sleepiness seem to benefit the most.

Arch Intern Med. 2003;163:565-571

For editorial comment see page 519

Obstructive sleep apnea (OSA) is a common disorder with a prevalence of approximately 4% in middle-aged men and 2% in middle-aged women. The disorder is characterized by recurrent collapse of the upper airway during sleep, resulting in repetitive episodes of asphyxia and arousal from sleep. This fragmentation of sleep results in daytime sleepiness, decreased quality of life, and an increased risk of injuries from motor vehicle crashes and industrial injuries.

Sleepiness, defined as an increased physiologic drive to fall asleep, can be assessed either subjectively or objectively. The most common measure of subjective sleepiness is the Epworth Sleepiness Scale (ESS).

The ESS is a validated 8-question survey asking the perceived likelihood of falling asleep in various situations. The test yields a score of 0 to 24, with 0 indicating no daytime sleepiness and 24 indicating the most severe sleepiness. Because an ESS score of 10 has been reported to be 2 SDs above the mean in healthy individuals, a score of 11 or greater is commonly used to define significant sleepiness.

The 2 most commonly used surrogate measures of objective sleepiness are the Multiple Sleep Latency Test (MSLT) and the Maintenance of Wakefulness Test (MWT). For the MSLT, the patient is instructed to try to fall asleep in a dark quiet room 4 or 5 times at 2-hour intervals. The MSLT score is the average number of minutes required to fall asleep, as measured by electroencephalography. The MWT is performed in a similar fashion, except the patient is asked to attempt to stay awake.

Continuous positive airway pressure (CPAP) therapy has emerged as the...
primary treatment for OSA. By increasing the intraluminal pressure in the pharynx, CPAP acts as a pneumatic splint, preventing collapse of the pharyngeal airway. However, there is still debate regarding the efficacy of this modality in treating symptoms of excessive daytime sleepiness. In 1997, Wright et al. systematically reviewed the literature regarding CPAP in the treatment of patients with OSA and found a lack of convincing evidence for improvement in sleepiness. However, they found only one randomized controlled trial, consisting of 32 patients, that addressed this issue. Although a meta-analysis from 1998 found a small benefit of CPAP on quality-of-life measures, the researchers found no evidence for improvement in measures of sleepiness. This study was also limited in that it was based on data from a total of only 61 patients. Since the publication of these 2 studies, several randomized clinical trials of CPAP in the management of OSA have been completed.

Despite the publication of a well-designed placebo-controlled trial by Jenkinson et al., substantial controversy has persisted as to the applicability of its findings. The Jenkinson et al study has been criticized for the homogeneity of its study population (all men with severe symptoms), making generalizability to the diverse population of patients with OSA difficult. The small number of patients studied in many CPAP trials has also made it difficult to make definitive conclusions regarding CPAP efficacy.

To address these concerns and to update the state of knowledge in this area, we conducted a meta-analysis of data from randomized controlled trials of CPAP therapy in patients with OSA. In particular, we wanted to address the question: “Does CPAP therapy improve subjective and objective measures of sleepiness in adults with OSA?”

METHODS

LITERATURE SEARCH

A systematic computerized search of publications in the MEDLINE database from January 1, 1966, to October 31, 2001, was conducted to identify all randomized controlled trials assessing the effect of CPAP on sleepiness in patients with OSA. The search terms used were the following: (apnea.af or apnoea.af or hypopnoea.af or hypopnoea.af) and (CPAP.af or positive airway pressure.af or positive airways pressure.af or positive pressure.af) and (clinical trial.pt or randomized controlled trial.pt). The suffix .af signifies that the term was allowed to be in any field, and the suffix .pt signifies the publication type field. Bibliographies of retrieved articles, previous meta-analyses, and reviews, including a search of the Cochrane database, were also used to further identify relevant publications. Experts in the field were also contacted for information regarding studies that may have been overlooked. Only full-length original articles, and not abstracts, reviews, or duplicate studies, were included.

LITERATURE SELECTION

The following inclusion and exclusion criteria were used: the study population was limited to adults with OSA, the intervention was CPAP therapy for at least 1 week, and the study design was a randomized controlled trial. Only studies that included a measure of change in sleepiness were included. Studies recording the ESS score were included in our analysis of CPAP effects on subjective sleepiness. Studies recording either the MSLT or the MWT score were included in our analysis of CPAP effects on objective sleepiness. Studies were required to have a control arm for inclusion in this review. The control arm could receive either no or minimal therapy (eg, recommendations for weight loss and a reduction in alcohol intake) or a true placebo. We excluded studies comparing CPAP devices with oral appliances and autotitrating devices (devices that automatically vary applied pressure during the night to the minimum required to keep the airway open). Similarly, observational studies without a control group were excluded.

VALIDITY ASSESSMENT

Before data abstraction, studies were assessed using the Jadad scale, a validated scale designed to measure trial quality. This scale rates studies on the basis of the quality of randomization, double-blinding, and inclusion of data on dropouts and withdrawals. The range of the scale is 0 to 5, with 0 denoting poor quality and 5 denoting the highest quality. To minimize selection bias, these criteria were evaluated independently by at least 2 investigators (S.R.P., N.T.A.; there were no disagreements between these investigators with regard to the Jadad score).

DATA ABSTRACTION AND STUDY CHARACTERISTICS

Data abstracted from each article included the year of publication; the study design; the number of subjects in each group; the type of placebo used; subject sex distribution, mean age, mean body mass index (BMI), baseline apnea-hypopnea index (AHI), mean level of CPAP, compliance with CPAP therapy, and length of follow-up; outcomes measured; and results (change in ESS, MSLT, and/or MWT score over time for the treatment and placebo groups).

QUANTITATIVE DATA SYNTHESIS

The measure of treatment effect for subjective sleepiness was defined as the difference between the treatment and placebo groups in the ESS score. The measure of treatment effect for objective sleepiness was defined as the difference between the treatment and placebo groups in either the MSLT or the MWT score. If not provided in the article, authors of the studies were contacted directly to obtain the mean and SD of the differences in ESS, MSLT, and MWT scores. To obtain a pooled estimate of the mean difference, the random-effects model of DerSimonian and Laird was applied. We chose a random-effects model because heterogeneity of studies would be expected because of differences in the study design and populations studied.

The summary estimates, along with the 95% confidence intervals (CIs), were graphed on a Forest plot along with the estimates derived from the original studies. Statistical heterogeneity was assessed using the Q statistic. Random-effects regression modeling was used to explore the sources of heterogeneity in the efficacy of CPAP between individual studies. Variables potentially associated with CPAP efficacy were individually incorporated into a single covariate model, and the significance of the covariate in explaining interstudy heterogeneity was assessed. Predefined variables of interest included sex composition, age, BMI, AHI, severity of sleepiness as assessed by baseline ESS score, length of follow-up, and CPAP compliance rate. Planned subgroup analyses also included divisions based on the country in which the study was conducted, the nature of the placebo used, and the quality of the study. In our
study, we chose to combine the measurements of MSLT and MWT. However, we acknowledge that although MSLT and MWT are both objective measures of daytime sleepiness, they are not well correlated. Therefore, to test for potential heterogeneity, subanalyses using only changes in MSLT and MWT scores were also conducted.

Funnel plots were generated by graphing the logarithm of the mean difference in outcome (ESS, MSLT, or MWT score) against the SE to assess for publication bias. All statistical analyses were performed using Stata software, version 7 (Stata Corp, College Station, Tex).

RESULTS

A total of 211 studies were identified from the literature search. After review of their abstracts, 18 of the studies were obtained and further reviewed. Three studies26-28 were excluded because a change in sleepiness was not an outcome variable. One study was excluded because most of the patients had already been described in a previous publication.29 One crossover study30 was excluded because it failed to randomize the order of treatment, and one study31 was excluded because it assessed only patients with positional sleep apnea. Of the remaining 12 studies,8,10-19,32 11 studies10-19,32 (including 706 patients) recorded change in ESS score and 8 studies8,10,11,13-15,18,32 (including 482 patients) reported change in either MSLT or MWT score.

Characteristics of the 12 studies are shown in Table 1. The quality of the studies varied significantly. Two studies10,15 had a Jadad score of 5, 4 studies14,16-18 had a Jadad score of 3, and the remaining 6 studies8,11-13,19,32 had a Jadad score of 1 or 2. The most common reason for point deduction was the lack of double-blinding, which was impossible in studies using a pill or conservative therapy control group.

Characteristics of each study population are given in Table 2. Substantial variability was found among studies in severity of sleep apnea, as measured by AHI, and in severity of subjective sleepiness, as measured by baseline ESS score. In addition, there were important differences in the nature of the placebo. Three studies12,13,18 administered only conservative measures to the control arm. These included counseling regarding weight loss, abstinence from alcohol and other sedatives, avoidance of the

<table>
<thead>
<tr>
<th>Table 1. Description of Included Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Source</strong></td>
</tr>
<tr>
<td>Barbé et al,15 2001</td>
</tr>
<tr>
<td>Faccenda et al,16 2001</td>
</tr>
<tr>
<td>Henke et al,17 2001</td>
</tr>
<tr>
<td>Monasterio et al,18 2001</td>
</tr>
<tr>
<td>Montserrat et al,19 2001</td>
</tr>
<tr>
<td>Ballester et al,15 1999</td>
</tr>
<tr>
<td>Engleman et al,24 1999</td>
</tr>
<tr>
<td>Jenkinson et al,10 1999</td>
</tr>
<tr>
<td>Engleman et al,25 1998</td>
</tr>
<tr>
<td>Redline et al,31 1998</td>
</tr>
<tr>
<td>Engleman et al,26 1997</td>
</tr>
<tr>
<td>Engleman et al,27 1994</td>
</tr>
</tbody>
</table>

Abbreviations: CPAP, continuous positive airway pressure; CT, conservative therapy; ESS, Epworth Sleepiness Scale; MSLT, Multiple Sleep Latency Test; and MWT, Maintenance of Wakefulness Test.

*The ESS score was measured in only 9 of these 16 patients.

<table>
<thead>
<tr>
<th>Table 2. Patient Characteristics*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Source</strong></td>
</tr>
<tr>
<td>Barbé et al,15 2001</td>
</tr>
<tr>
<td>Faccenda et al,16 2001</td>
</tr>
<tr>
<td>Henke et al,17 2001</td>
</tr>
<tr>
<td>Monasterio et al,18 2001</td>
</tr>
<tr>
<td>Montserrat et al,19 2001</td>
</tr>
<tr>
<td>Ballester et al,15 1999</td>
</tr>
<tr>
<td>Engleman et al,24 1999</td>
</tr>
<tr>
<td>Jenkinson et al,25 1999</td>
</tr>
<tr>
<td>Engleman et al,26 1998</td>
</tr>
<tr>
<td>Redline et al,31 1998</td>
</tr>
<tr>
<td>Engleman et al,27 1997</td>
</tr>
<tr>
<td>Engleman et al,27 1994</td>
</tr>
</tbody>
</table>

Abbreviations: AHI, apnea-hypopnea index; BMI, body mass index (calculated as weight in kilograms divided by the square of height in meters); CPAP, continuous positive airway pressure; ESS, Epworth Sleepiness Scale; NR, not reported.

*Data are given as mean values unless otherwise indicated.

†Median value.
supine posture, and obtaining sufficient hours of sleep. Five studies\(^8,11,14,16,32\) administered an oral pill as a placebo. Patients were told that the pill might improve upper airway function. The remaining 4 studies\(^8,11,14,16,32\) used a sham CPAP placebo. Sham CPAP uses all the machinery and headgear of CPAP, but with the airway pressure set at a low level, which is adequate to wash out exhaled carbon dioxide but insufficient to eliminate obstructive events. Differences in study design were also apparent that enrolled more severely afflicted patients. This is somewhat surprisingly, neither mean AHI nor mean baseline ESS score for each study significantly predicted the difference in ESS score improvement (\(P=.08\) and \(P=.13\), respectively). However, when the analysis was limited to the 6 studies\(^8,11,14,16,17,19\) that studied patients with severe sleep apnea (mean AHI or oxygen desaturation index, \(\geq 30\) events per hour) and significant sleepiness (mean baseline ESS score, \(\geq 11\)), the pooled average reduction in ESS score was much greater with CPAP compared with control (4.75; 95% CI, 2.97-6.53; \(P<.001\)). That is, the ESS score decreased 4.75 points more in the CPAP arms than in the control arms of the studies that enrolled more severely afflicted patients. This is shown in Figure 2. A meta-analysis of the remaining 5 studies\(^8,10,11,13,15,18,32\) resulted in a mean difference in ESS score of only 1.10 (95% CI, -0.13 to 2.32), which was not significantly different from the null hypothesis of no effect from CPAP (\(P=.08\)).

For the evaluation of objective sleepiness, the difference in mean change in either MSLT or MWT score between the placebo and CPAP groups was chosen as the variable of interest. That is, we assessed whether sleep latency had a greater increase with CPAP or with placebo. From Table 3 and Figure 3, one can see that six\(^8,10,11,13,14,32\) of the eight studies\(^8,10,11,13,15,18,32\) had mean differences greater than 0 (ie, the CPAP group had a greater improvement than the control group). To summarize the data concerning the change in sleep latency (change in MSLT or MWT score), a meta-analysis was performed. Although the Q statistic did not demonstrate significant heterogeneity (\(Q=11.6\), \(P=.12\)), we chose a random-effects model because the Q statistic lacks power to detect differences between studies and one would expect significant heterogeneity given the significant differences in study populations and study designs. The summary estimate for mean ESS score reduction was 2.87 (95% CI, 1.48 to 4.25); and diamond, the pooled effect (with a mean of 2.94).
Our meta-analysis shows that CPAP therapy across diverse populations of patients with OSA results in a significant improvement in subjective and objective measures of sleepiness. On average, CPAP decreases the ESS score by 2.94 points and increases sleep latency by 0.93 minute compared with placebo. We believe this improvement in sleepiness is not only statistically significant, but also clinically important. The magnitudes of changes seen in this study are similar to those observed with other accepted therapies of excessive daytime sleepiness. For example, a large randomized study of modafinil in the treatment of narcolepsy found a mean improvement of 2.6 points and 0.6 minute in ESS and MSLT scores, respectively, at a dose of 200 mg/d, and 3.9 points and 1.1 minutes, respectively, at a dose of 400 mg/d. These same doses of modafinil have improved quality of life in narcoleptic patients when assessed using standard validated measures. Because narcoleptic patients tend to have more severe sleepiness than patients with apnea, the same improvement in measures of sleepiness may not be as clinically meaningful in patients with sleep apnea. However, the finding by Wright et al that CPAP therapy did improve quality-of-life measures supports our notion that the effect of CPAP on sleepiness is clinically relevant.

Our conclusions differ substantially from those of previous reviews. Both of these earlier meta-analyses could not find evidence of a beneficial effect of CPAP therapy on daytime sleepiness, largely because of a lack of randomized placebo-controlled trials. Since these analyses, several clinical trials of CPAP therapy have been published. Our incorporation of these trials into our analysis accounts for the differences between our results and those of other investigators.

All of the randomized trials included in this analysis were performed at either a single site or a few sites within one country, with detailed inclusion and exclusion criteria, resulting in a homogeneous study population within the individual studies. An advantage of our analysis is that by incorporating the different patient populations of each of the individual trials, our findings are based on a more heterogeneous population that better represents the diversity of patients affected by sleep apnea.

As expected given the heterogeneity of study designs and patient populations among the individual trials, substantial heterogeneity in the mean difference of changes in subjective sleepiness was identified ($Q_{10} = 57.7, P < .001$). We explored several patient characteristics as potential explanations for this heterogeneity, but could not identify a consistent determinant. In particular, differences in sex composition, nation of the study, mean age, and BMI were not predictors of the mean difference. Thus, on the basis of the available data, CPAP therapy seems equally effective in patients with sleep apnea, regardless of these variables.
The length of follow-up in the included studies was relatively short. The longest follow-up was 24 weeks, and most studies were limited to 4 to 6 weeks in duration. Within this time frame, we could not find a relationship between length of treatment and mean difference in ESS score. However, longer-term studies are clearly needed.

As in all meta-analyses, one potential weakness of this study is the possibility of publication bias. If negative trials remain unpublished, this would result in an overestimate of the treatment effect of CPAP. Although we cannot completely exclude this possibility, we doubt this is the case for several reasons. First, we constructed a funnel plot of the treatment effect vs the sample size; this did not indicate a substantial effect of publication bias. That is, there was not a clear predominance of positive effects in studies with a small sample size. Second, because large trials are likely to be published regardless of the results, publication bias is more likely to exist in small trials. We, therefore, calculated mean differences after excluding the 2 smallest studies from the analysis. No substantial change in the results was found.

Although neither mean AHI nor mean ESS score alone predicted CPAP effectiveness, an analysis limited to studies of patients with severe sleep apnea (AHI, \( \geq 30 \)) and severe sleepiness (ESS score, \( \geq 11 \)) demonstrated a much greater effect of CPAP (mean ESS score change, 4.75) than the remainder of the trials (mean ESS score change, 1.10). Because this subgroup was defined post hoc and the results are based on study means as opposed to data from individual patients, these results need to be interpreted cautiously.

Nevertheless, these data suggest that CPAP may be more effective in individuals with severe OSA and sleepiness. There may be several reasons for this. First, in some sleepy patients with mild sleep apnea, sleepiness may be partially due to a concomitant sleep disorder, such as periodic limb movements, chronic sleep deprivation, occult narcolepsy, or idiopathic hypersomnolence. Consequently, treatment with CPAP may not impact daytime sleepiness substantially in these individuals. Second, patients with severe sleep apnea and a normal ESS score may not demonstrate a substantial improvement in ESS score because of a “floor effect.” If the score is already low at baseline, there is little room for it to decrease further. Third, sleepier patients may be motivated to use CPAP more often, increasing the potential benefits of therapy.

Our analysis supports the claim that CPAP improves subjective and objective daytime sleepiness in patients with OSA. These benefits were independent of study population sex composition, age, BMI, or nationality. However, many issues concerning CPAP therapy in patients with OSA remain unanswered. Further research is necessary to compare standard CPAP with other treatment options, such as oral appliances, upper airway surgery, and autotitrating devices. The effect of CPAP on different subgroups of patients with OSA (such as those with mild disease or asymptomatic patients) needs to be better defined. Also, the effect of CPAP on other health outcomes (such as cardiovascular disease) needs to be clarified. However, this analysis does show a clinically important improvement in sleepiness with CPAP therapy in a diverse group of patients with apnea.

Accepted for publication June 19, 2002.

This study was supported by grants HL07633 (Drs Patel and Stanchina), HL48531 (Dr White), and HL60292 (Dr White) from the National Institutes of Health, Bethesda, Md; the Medical Research Council of Canada, Ottawa, Ontario (Dr Malhotra); and the American Heart Association, Dallas, Tex (Dr Malhotra).

We thank Catherine S. Berkey, DSc, for her statistical expertise and Yvonne Gilreath for her assistance in the preparation of the manuscript.

Corresponding author: Najib T. Ayas, MD, Division of Pulmonary Medicine, Vancouver General Hospital, 2775 Heather St, Vancouver, British Columbia, Canada V5Z 3J5 (e-mail: ayasnajib@pol.net).

REFERENCES


**Correction**

**Error in Letter.** In the letter by Raoul A. Walsh, BA, Dip Ed, PhD, titled “Nicotine Lozenge Trial: A ‘Real-World’ Perspective,” published in the December 9/23, 2002, issue of the *ARCHIVES* (2002;162:2632-2633), an error occurred in the second sentence of the second paragraph. The parenthetical statement “(during which 9% dropped out)” should have read “(during which 13% dropped out).”