Effectiveness of CPAP Treatment in Daytime Function in Sleep Apnea Syndrome

A Randomized Controlled Study with an Optimized Placebo

JOSEP M. MONTSERRAT, MONTSERRAT FERRER, LOURDES HERNANDEZ, RAMÓN FARRÉ, GEMMA VILAGUT, DANIEL NAVAJAS, JOAN R. BADIA, EVA CARRASCO, JUAN DE PABLO, and EUGENI BALLESTER with the Technical Assistance of Marta Puig

Institut Clínic de Pneumologia i Cirurgia Toràcica and Servei de Psiquiatria, Hospital Clínic, Barcelona, Spain; Health Services Research Unit, Institut Municipal d'Investigació Mèdica (IMIM), Barcelona, Spain; Biophysics and Bioengineering Unit, Barcelona University, Barcelona, Spain

Application of continuous positive airway pressure (CPAP) as the standard treatment for sleep apnea/hypopnea syndrome (SAHS) is a moot point. Studies on the effectiveness of this treatment have been challenged because of the lack of a suitable placebo. The recent description of a true placebo (sham CPAP) prompted us to conduct a randomized trial of CPAP or placebo to assess the effectiveness of CPAP in improving SAHS-related symptoms and daytime function in patients with moderate to severe SAHS. Forty-eight patients, stratified in four groups according to severity, were randomly allocated into two treatment groups (optimal and sham CPAP) for a 6-wk period. Of these, 45 completed follow-up (91% males; age: 54 ± 10 yr; body mass index [BMI]: 32 ± 6 kg/m^2; apnea-hypopnea index [AHI]: 54 ± 19 events/h; and Epworth Sleepiness Scale [ESS]: 16 ± 5). The ESS, a questionnaire on SAHS-related symptoms, Functional Outcomes Sleep Questionnaire (FOSQ), and the Short Form Health Survey (SF-36) were completed at inclusion and after treatment. After 10 d of washout, the placebo group was treated with optimal CPAP and reassessed before and after optimal CPAP. The group receiving optimal CPAP when compared with the group with sham CPAP showed considerably greater improvement in the relief of sleepiness (−9.5 versus −2.3, $p < 0.001$), other SAHS-related symptoms (−18.5 versus −4.5, $p < 0.001$), vigilance (+8.5 versus +3.4, $p = 0.009$), and general productivity (+4.0 versus +0.5, $p = 0.04$) FOSQ scales. Both groups used a similar number of hours for the optimal and the sham CPAP (4.3 versus 4.5, (p = NS). The patients initially treated with placebo CPAP improved significantly more when optimal CPAP was applied for ESS (−2.3 versus −6.7, $p < 0.001$) and other sleep apnea syndrome (SAS)-related symptoms (−4.5 versus −11.2, $p = 0.02$). Our study provides strong evidence of the effectiveness of CPAP treatment in improving symptoms and perceived health status in moderate to severe SAHS.

Keywords: sleep apnea; questionnaires; positive-pressure respiration; treatment outcome; comparative study; controlled study

Ever since its initial description, nasal continuous positive airway pressure (CPAP) has had a broad acceptance as the treatment of choice for the sleep apnea/hypopnea syndrome (SAHS) (1). However, despite extensive use in the last decade, the benefits of this treatment have been challenged (2) given the small number of studies that have adequately evaluated its efficacy. The controlled studies published so far include a comparison of the efficacy of optimal CPAP and conservative management, a pill placebo, or suboptimal CPAP (3–9). One possible weakness of these studies has been the absence of true masking of the placebo group. Our group recently designed a sham CPAP that mimics all the characteristics of a true CPAP (noise, flow, humidity, mask, and psychological constraints) except for the null pressure applied to the upper airway of the patient (10). Moreover, this method does not influence sleep efficiency, arousals, or apnea–hypopnea index (AHI) (10). Using this optimized placebo, we conducted a randomized double-blind placebo-controlled trial to assess the effectiveness of CPAP in improving perceived health status, somnolence, and other sleep apnea syndrome (SAS)-related symptoms in patients with moderate to severe SAHS.

METHODS

Subjects

The patients were recruited in the sleep clinic of a university teaching hospital in Barcelona, Spain. Moderate to severe—symptomatic—SAHS which required CPAP treatment was diagnosed in all of them. None of the patients had received previous treatment with CPAP. Criteria for CPAP treatment were as follows: significant SAHS-related clinical symptoms, which in all cases included excessive daytime somnolence and an AHI > 10. The protocol was approved by the hospital human ethics committee, and written informed consent was obtained from all the patients. Owing to long waiting lists for both diagnosis and treatment with CPAP (2 mo in the latter situation in our hospital), the prescribed treatment with CPAP in the placebo group was delayed only for a period similar to that in which CPAP treatment would normally have been provided.

Sample Size and Randomization

The trial was estimated to require 48 patients based on the following assumptions: an expected improvement in Epworth Sleepiness Scale (ESS) score defined as 4 points with a standard deviation of 4.2 (11) and a statistical power of at least 90% at a significance level of 5%. To ensure that the main known prognostic characteristics were equally distributed among the treatment groups, a block-randomized assignment was used. Randomization was performed with a computer-generated allocation schedule that had a block size of 12 patients in accordance with severity (Figure 1). Consecutive patients of each of the following four severity groups were recruited: (*1*) AHI ⩽ 50 and ESS < 15; (2) AHI > 50 and ESS < 15; (3) AHI ⩽ 50 and ESS ⩾ 15; and (4) AHI > 50 and ESS AHI ⩾ 15. Three patients were excluded before randomization because they had either a severe or unstable cardiovascular disease or a hazardous job coincidentally with SAHS (professional drivers or handling dangerous machinery). The patients were randomly allocated to one of two groups (A and B), optimal CPAP (24 patients) or placebo (24 patients), respectively. After randomization, three patients dropped out at different follow-up times as shown in Figure 1. All the patients included in the study were encouraged to follow a diet and sleep hygiene regimen, regardless of the treatment group assigned. These conservative measures consisted of the following: as for sleep hygiene, the patients were encouraged to sleep suffi-

(*Received in original form June 7, 2000 and in revised form January 26, 2001*)

Supported by Fondo de Investigaciones Sanitarias (FIS)-00/0575, Sociedad Española Patologia Aparato Respiratorio (SEPAR-1999), and Comisión Interministerial de Ciencia y Tecnología (CICYT, SAF99-001).

Correspondence and requests for reprints should be addressed to Josep M. Montserrat, M.D., Institut Clínic Pneumologia i Cirurgia Toràcica (ICPCT), Hospital Clínic, Villarroel 170, 08034 Barcelona, Spain. E-mail: jcanal@medicina.ub.es

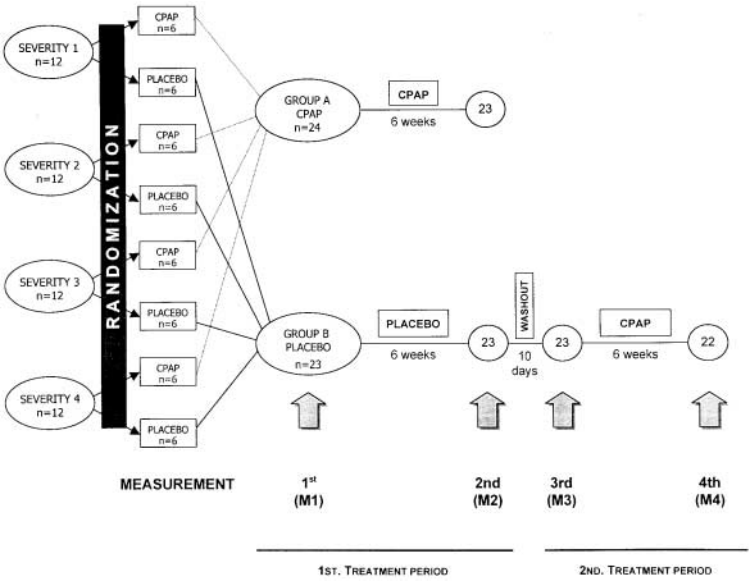

Figure 1. Flow chart describing progress of patients through the randomized trial.

cient hours every night, sleep on their side, avoid sedatives and alcohol consumption. They followed a weight loss program by adopting a home diet prescribed by a dietician. All the patients received written information on recommendations and had a close follow-up through weekly telephone calls by a specialized nurse to encourage compliance with the treatment.

The patients were told to follow a daily treatment which, if necessary, would be readjusted in 6 wk. The patients and the interviewers were unaware of the treatment assignment.

Placebo CPAP

The sham CPAP system has been described previously (10). Briefly, the nasal pressure is eliminated and the reinhalation avoided by increasing the area of the exhalation port from 4 mm to 10 mm. An orifice resistor identical to the original resistor of the mask is placed between the CPAP device and the tubing to load the blower with the same resistance as the true CPAP. The ventilator operating noise and the airflow through the exhalation port remained unchanged, thus ensuring true masking. A Breas PV100 medical device (Breas, Sweden) was used in the optimal CPAP treatment. The same model was modified to operate as a CPAP placebo (10).

Overnight Diagnosis Study

All the patients underwent an overnight study by a validated procedure (12). Briefly, this was performed in all the subjects using a Densa Pneumograph (Densa Ltd, Flint, UK), which measures oronasal flow by thermistor and thorax and abdominal motion. Sa_{O_2} was measured using a pulse oximeter (model 504; Critical Care Systems Inc., Waukesha, WI). A body position sensor differentiated between supine and side position. Apnea and hypopnea were defined in accordance with commonly used clinical criteria of airflow cessation lasting 10 s or more for apneas, and 10 s or more of discernible airflow reduction associated with a cyclical Sa_{O_2} dip $\geqslant$ 3% for hypopneas. The studies were made in the sleep laboratory, and sleep time was measured as the total number of events registered divided by registry time. Manual scoring was done in all cases.

Optimal and Placebo CPAP Titration

All the patients were instructed in the use of CPAP before the night in the sleep unit. The best-fitting mask was chosen, and the patients were familiarized with the CPAP treatment with the help of a trained physiotherapist and staff physicians. A short nap with either sham or low-level CPAP increased the patient's confidence in the device. CPAP titration at night was performed using a full polysomnography (PSG) setting. The measurements made in the optimal CPAP group were carried out during a period of stepwise, full-night, PSG-controlled CPAP titration. Electroencephalogram (EEG) (C4/A1, C3/A2); chin electromyogram (EMG); and electrooculogram (EOG) for sleep staging in accordance with standard criteria were recorded. Sa_{O_2} was measured continuously with a finger probe (model 504; Critical Care Systems, Inc., Waukesha, WI). Bands placed over the thorax and abdomen monitored rib cage and abdominal motion. Flow was measured with a pneumotachograph connected to a differential pressure transducer and located between the leak valve and the mask. The parameters were recorded continuously on a polygraph (SleepLab 1000P; Aequitron Medical Inc., Minneapolis, MN). In the optimal CPAP group, the mean AHI was 50.5 $\pm$ 4 and pressure was adjusted until apneas, hypopneas, snoring, thoracoabdominal incoordination, flow limitation, and respiratory-related arousals disappeared. The mean CPAP pressure prescribed was 9.8 $\pm$ 1.3 cm H_2O. In the placebo group, a similar simulated full night set-up with the sham CPAP but without recordings was performed.

Follow-up and Outcome Measures

Figure 1 summarizes the design of the trial. All the patients were evaluated through different questionnaires at inclusion (measurement 1 [M1]), and after 6 wk of treatment (M2). After the first 6 wk, the placebo CPAP group followed a washout period of 10 d and subsequently started a second period of 6 wk with optimal CPAP (obtained by full PSG). In this group of patients, questionnaires were also completed at the end of the washout period (M3), and at the end of the second 6-wk period with optimal CPAP (M4). The following standardized questionnaires were completed under the supervision of a trained interviewer: the ESS (13), the Functional Outcomes Sleep Questionnaire (FOSQ) (14), the 36-item Short Form (SF-36) Health Survey (15), and the questionnaire of symptoms related to SAHS (8). Body weight and hours of CPAP use were also recorded.

The ESS, which is a self-administered questionnaire designed to measure daytime sleepiness, distinguishes patients with primary snoring from those with obstructive sleep apnea syndrome (16). The subjects rate on a scale of 0 to 3 the chances of dozing in eight different situations (13). The final score ranges from 0 (no daytime sleepiness) to 24 (maximum daytime sleepiness).

The FOSQ, developed by Weaver and coworkers, is a self-administered instrument designed to assess the impact of excessive sleepiness on daytime function and to quantify improvement after treatment (14). It contains 30 items divided into five scales: Activity Level, Vigilance, Intimacy and Sexual Relationships, General Productivity, and Social Outcome. A mean score was calculated for each scale ranging from 0 (maximum functional impact) to 24 (no functional impact). As recommended (14), to prevent distortion resulting from a missing response, only items on activities in which the respondent regularly participated were included in the scoring algorithm. Scale scores were

TABLE 1. GENERAL CHARACTERISTICS OF THE STUDY SAMPLE

	CPAP ($n = 23$)	Sham CPAP ($n = 22$)	p Value
Age	55.65 ± 9.41 (28–74)	52.59 ± 10.93 (31–77)	0.319†
AHI	50.52 ± 19.83 (20–81)	57.14 ± 21.14 (19–96)	0.290†
Neck circumference	42.52 ± 3.66 (36–49)	43.72 ± 3.63 (37–52)	0.273†
Alcohol, g/d	21.46 ± 19.31 (0–60)	28.26 ± 30.62 (0–120)	0.705†
BMI			
Baseline	30.31 ± 4.49 (23–41)	33.73 ± 6.62 (21–48)	0.048†
Change at 6 wk (M2 − M1)	+0.22 ± 1.15 (−1.7–+2.7)	–0.27 ± 1.01 (−2.4–+2.7)	0.135†
Smoking habit, %			
Nonsmoker	13.0%	36.4%	0.034‡
Former smoker	39.1%	9.1%	
Smoker	47.8%	54.5%	

* Values are expressed as mean ± standard deviation of the mean, with range in parentheses.
† Mann-Whitney test for independent samples.
‡ Chi-square test.

added to compute a global score ranging from 0 (maximal dysfunction) to 120.

The SF-36 Health Survey is a generic health status quality-of-life instrument. Item responses were coded and transformed into two component summaries (the Physical [PCS], and Mental [MCS]), and eight subscales (Physical Functioning, Role-Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role-Emotional, and Mental Health). The two component summary of the SF-36 (PCS-12 and MCS-12) have a mean of 50 and a standard deviation of 10 in the Spanish general population. All scores above or below 50 are better or worse, respectively (17, 18). Scores of the eight SF-36 subscales range from 0 (minimum well-being) to 100 (maximum well-being).

SAHS-related symptoms were assessed with a list of 15 items about snoring, breathing pauses, nocturia, dry mouth, morning headaches, unrefreshing sleep, and other common symptoms besides sleepiness reported by patients with SAHS. The results range from 15 (absence of symptoms) to 60 (maximally symptomatic patient) (8). Participants were questioned on their smoking habits at the beginning of the study (current, former, or nonsmokers). Information was elicited from active smokers at the end of the study, and stop data were recorded when applicable.

Data Analysis

An independent *t* test, a nonparametric (Mann-Whitney) test, or the chi-square test was applied to compare baseline data between groups. The *t* test for paired samples was applied to test pre- and post-treatment differences. Two different approaches were followed to assess the effectiveness of the treatment. First, the improvement observed during the first treatment period in Group A (optimal CPAP) was compared with Group B (sham CPAP) by an analysis of variance (ANOVA) with repeated measures. Analyses were performed adjusting for tobacco and body mass index (BMI) at baseline. Second, in the group of patients who initially received placebo CPAP, we compared the improvement observed during the first and second treatment periods using a repeated measures ANOVA. Statistical significance was accepted at $p < 0.05$. All the analyses were developed with the Statistical Package for Social Sciences (SPSS, rel. 6.1.3 for Windows; SPSS Inc., Chicago, IL).

RESULTS

Baseline data (Table 1) from the 45 patients with SAHS who completed the follow-up did not show significant differences by randomization groups in age (mean = 54.2 yr, SD = 10.2), AHI (mean = 53.8 events/h, SD = 19.3), neck circumference (mean = 43.1 cm, SD = 3.7), or alcohol consumption (mean = 24.8 g/d, SD = 25.4). Significant differences were found only for BMI (30 versus 34 kg/m^2, $p = 0.048$) and tobacco (48 versus 55% of smokers, $p = 0.034$).

Following the design shown in Figure 1, Table 2 shows the mean and standard error of outcome variables in accordance with randomization group (Group A—optimal-CPAP, and Group B—sham CPAP, during the first treatment period; left side of the table) at different measurements. Moreover, the changes observed during the first treatment period (differences between measurement 1 and measurement 2) of each group and their statistical significance are also provided: for both treatment groups, improvement was significant in ESS, SAHS-related symptoms, most of the FOSQ subscales, and some SF-36 scores. No significant differences were found at baseline (measurement 1) between Group A and Group B for the assessed outcome variables. Significant differences between treatment groups after the first 6 wk were found in the score change in the ESS (< 0.001), SAHS-related symptoms (< 0.001), FOSQ General Productivity, (< 0.004), and the FOSQ Vigilance scale (< 0.009). The total differences in the FOSQ scores during the first treatment period were higher in Group A (optimal CPAP, 25) than in Group B (placebo CPAP, 15) but were not statistically significant. No SF-36 score showed significant differences in accordance with the treatment group.

Table 2 also shows the results of Group B when treated with optimal CPAP after the washout period (second treatment period, right side of the table). A greater improvement was observed during the second treatment period (optimal CPAP) than during the first (sham CPAP) in the ESS (< 0.0001) and in the SAHS-related symptoms scale (< 0.023). Changes in the FOSQ score were similar in both treatment periods. On the other hand, near significant differences were found in accordance with the treatment period on the Mental Summary score (0.066), indicating a greater improvement during the first period with sham CPAP than during the second one with optimal CPAP.

After optimal CPAP the improvement was marked compared with sham CPAP. Group B (initially treated with sham CPAP) after treatment with optimal CPAP achieved values similar to those obtained in the group initially treated with optimal CPAP. This pattern is observed in most of the representative outcome variables studied. A ceiling effect was observed in some FOSQ dimensions, social, intimacy, and sexual relationships being the most important.

DISCUSSION

This true placebo-controlled study demonstrates that the CPAP therapy is effective in improving daytime function. We observed that the relief of sleepiness and other SAHS-related clinical symptoms, as well as the improvement in some FOSQ scales, was much greater in the group receiving optimal CPAP than with placebo CPAP. This difference in outcome of treatment between both groups was statistically significant and substantial.

Despite recent efforts to corroborate the effectiveness of CPAP in patients SAHS (7–9), the type of placebo used has been criticized given that it does not conform to the characteristics of the ideal placebo (true masking), thus making outcome clinical measures difficult to interpret. The requirement of a true placebo such as the one used in this trial has been recently met by applying null pressure and by subjecting the patient to the same physical and psychological constraints, thus mimicking all the features of a conventional CPAP system (exposure to the same noise, humidity, perceived airflow at the mask, and no rebreathing) (10). This has also been supported by the absence of effects in clinical sleep variables (sleep architecture and respiratory events) compared with conventional CPAP. The placebo method has the additional advantage of ease of implementation regardless of the CPAP device used and is therefore a suitable tool for clinical studies.

The first approach used to assess CPAP effectiveness compares the change in the two groups initially randomized. It should be pointed out that both groups of patients used the CPAP device for a similar number of hours (optimal: 4.25 ± 2; sham: 4.5 ± 2). Improvement between Groups A and B differed 7 points in the ESS, 14 points in the SAS-related symptoms scale, and approximately 4 points in the two FOSQ scales. For ESS, the difference was higher than that observed between patients with moderate SAHS and primary snoring (16), approximately 30% of their theoretical range. Although ESS has a poor correlation with AHI, it is a good outcome variable for assessing changes with CPAP therapy. For the FOSQ scales, differences in improvement observed exceeded the theoretical score range by 15%, and were similar to those for individuals with sleep problems and those of similar age and sex with no sleep disorder (14).

The effectiveness of optimal CPAP is supported not only in the first analytical approach of the trial but also in the second analytical approach, when sham CPAP is crossed to optimal CPAP. It is of interest to note that both groups of patients presented at baseline with an ESS mean score equal to that previously described in patients with severe SAHS (16 points). Moreover, both groups presented, after optimal CPAP treatment, a mean ESS score of 6.7 points, similar to that previously reported for patients with primary snoring (8 points) (16). We believe that the agreement with the findings in the two analytical approaches used and also with expected scores lends further support to the results of this study.

A recently published meta-analysis (2) has challenged the effectiveness of CPAP treatment given the lack of conclusive evidence of benefits. The sleep community has thus raised a number of objections (19, 20). The main criticism of these studies is the absence of a suitable placebo. The design and the validation of a true sham CPAP, which mimics all the characteristics of true CPAP (noise, flow, humidity, mask) with the obvious exception of nasal pressure (10), have enabled us to overcome this limitation encountered in published clinical trials. Thus, no differences were found in the respiratory or in the neurological variables when a full PSG was performed in a random order during sham CPAP or a diagnostic setting in the same patients (10). Therefore, we think that the placebo CPAP used meets all the requirements of a true placebo. Moreover, the patients were subjected to the use of CPAP for the first time, being only aware that they were participating in a study in which two different CPAP systems were used.

Although randomization was used to achieve a homogeneous distribution of the possible confounding factors, this cannot be ensured because of the reduced number of patients. To overcome this problem, randomization was blocked for the two main indicators of severity (somnolence and AHI). Moreover, this design allowed us to assess differences in effectiveness by subgroups of severity, as suggested by Wright

TABLE 2. SYMPTOMS AND DAYTIME FUNCTION AT EACH MEASUREMENT TIME*

	First Treatment Period									Second Treatment Period				
	Group A (CPAP)				Group B (Sham CPAP)				Comparison between Groups p Value‡	Group B (Sham CPAP)				Comparison Before–after p Value§
	Pre-treat (M1)	Post-treat (M2)	Dif	p Value†	Pre-treat (M1)	Post-treat (M2)	Dif	p Value†		Pre-treat (M3)	Post-Treat (M4)	Dif	p Value†	
Epworth‖	16.13 ± 1.03	6.65 ± 0.68	−9.48	0.000	16.86 ± 1.20	14.59 ± 1.08	−2.27	0.013	< 0.001	13.90 ± 1.21	7.18 ± 1.07	−6.72	0.000	< 0.001
SAHS-related symptoms¶	39.70 ± 1.14	21.22 ± 1.13	−18.48	0.000	38.86 ± 1.14	34.41 ± 1.91	−4.45	0.033	< 0.001	33.86 ± 1.52	22.64 ± 1.38	−11.22	0.000	0.023
FOSQ**														
General Productivity	19.18 ± 1.06	23.17 ± 0.27	3.99	0.002	20.25 ± 0.89	20.75 ± 0.72	0.50	0.570	0.044	21.32 ± 0.76	23.17 ± 0.38	1.85	0.020	0.224
Social	18.96 ± 1.26	22.96 ± 0.45	4	0.006	18.00 ± 1.66	22.10 ± 0.98	4.10	0.013	0.908	21.90 ± 0.86	23.45 ± 0.40	1.55	0.057	0.226
Activity Level	16.71 ± 0.81	21.25 ± 0.56	4.54	0.000	17.21 ± 1.23	19.42 ± 0.98	2.21	0.015	0.101	19.76 ± 0.89	22.28 ± 0.40	2.52	0.003	0.774
Vigilance	13.66 ± 1.16	22.18 ± 0.59	8.52	0.000	14.17 ± 1.48	17.61 ± 1.12	3.44	0.003	0.009	18.50 ± 1.04	22.60 ± 0.58	4.10	0.000	0.552
Intimacy and Sexual Relationship	16.40 ± 1.57	18.95 ± 1.76	2.55	0.294	17.43 ± 1.63	20.30 ± 1.42	2.87	0.058	0.769	19.89 ± 1.59	22.53 ± 0.63	2.64	0.037	0.552
Total Score	84.45 ± 4.63	109.43 ± 2.63	24.98	0.000	86.16 ± 5.96	100.66 ± 4.39	14.50	0.008	0.124	101.82 ± 4.29	114.07 ± 1.73	12.25	0.006	0.954
SF-36††														
Physical Component Summary (PCS)	46.53 ± 1.92	50.71 ± 1.58	4.18	0.002	45.54 ± 2.17	47.16 ± 1.87	1.62	0.364	0.230	48.76 ± 1.52	50.67 ± 1.20	1.91	0.143	0.980
Mental Component Summary (MCS)	48.21 ± 2.06	49.53 ± 2.44	1.32	0.606	48.73 ± 2.49	53.65 ± 1.57	4.92	0.006	0.518	52.36 ± 1.62	54.02 ± 1.63	1.66	0.215	0.066
Physical Functioning (PF)	76.96 ± 5.66	82.83 ± 5.16	5.87	0.004	78.18 ± 4.80	78.18 ± 4.89	0.00	1.000	0.057	82.95 ± 3.60	87.05 ± 3.38	4.10	0.080	0.374
Role-Physical (RP)	71.74 ± 8.79	84.78 ± 7.33	13.04	0.162	78.41 ± 8.11	85.23 ± 6.30	6.82	0.378	0.411	92.05 ± 4.76	98.86 ± 1.14	6.81	0.186	0.933
Bodily Pain (BP)	75.26 ± 5.26	85.35 ± 4.02	10.09	0.004	58.32 ± 7.27	78.50 ± 4.30	20.2	0.003	0.247	81.36 ± 4.97	78.18 ± 5.28	−3.18	0.576	0.007
General Health (GH)	60.48 ± 3.05	69.91 ± 3.57	9.43	0.001	61.55 ± 5.61	62.64 ± 4.18	1.09	0.790	0.106	59.45 ± 4.46	70.73 ± 4.09	11.28	0.001	0.084
Vitality (VT)	56.52 ± 5.93	69.35 ± 5.70	12.83	0.053	58.03 ± 6.23	68.41 ± 4.37	10.38	0.053	0.740	66.36 ± 5.46	75.45 ± 3.66	9.09	0.068	0.744
Social Functioning (SF)	82.61 ± 3.91	84.24 ± 5.11	1.63	0.741	82.39 ± 5.37	91.48 ± 4.06	9.09	0.032	0.294	91.48 ± 3.33	92.61 ± 3.36	1.13	0.765	0.024
Role-Emotional (RE)	84.06 ± 7.52	86.96 ± 7.18	2.90	0.760	75.76 ± 8.55	95.45 ± 3.32	19.69	0.024	0.433	92.42 ± 4.87	95.45 ± 4.55	3.03	0.665	0.025
Mental Health (MH)	71.83 ± 3.93	74.96 ± 5.04	3.13	0.502	77.45 ± 3.67	80.36 ± 3.47	2.91	0.306	0.524	80.91 ± 3.70	83.82 ± 3.16	2.91	0.073	0.906

* Crude mean scores of questionnaires ± standard error.

† p Value, *t* test for paired samples to test pre- and post-treatment differences.

‡ p Value, obtained from the ANOVA model of treatment group (adjusting for tobacco and BMI) to assess the treatment group effect (interaction between the treatment group and the repeated measures factor).

§ p Value, obtained from the ANOVA model of treatment group (adjusting for tobacco and BMI) to assess the effect of the treatment period in Group B.

‖ Range: 0 (no daytime sleepiness) to 24 (maximum daytime sleepiness).

¶ Range: From 15 (absence of symptoms) to 60 (maximally symptomatic patients).

** The five scales of the FOSQ range from 0 (maximum functional impact) to 24 (no functional impact). The total score ranges from 0 (maximal dysfunction) to 120.

†† The Component summaries of the SF-36 (PCS-12 and MCS-12) have a mean of 50 and a standard deviation of 10 on the general population. The eight subscales range from 0 (minimal well-being) to 100 (maximal well-being).

and coworkers (2). No effect of severity was found when this variable was included in the model, indicating that CPAP effectiveness does not differ in accordance with severity groups included in our study. However, because only patients with moderate to severe SAHS were recruited, our results cannot be extrapolated to a population of patients with mild SAHS.

Improvement in the group with sham CPAP was statistically significant for Epworth, SAHS-related symptoms, and for most of the FOSQ scores (Social, Activity Level, Vigilance, and Total). The Epworth improvement was similar to the placebo effect observed by Jenkinson and coworkers with suboptimal CPAP (9). However, in our study, this improvement in the group treated with sham CPAP could in part be explained by the complementary sleep hygiene regimen. Given that SAHS has been closely associated with obesity and tobacco (20, 21), and that conservative measures are essential to the treatment, this sleep hygiene regimen was implemented during both optimal and sham CPAP. On the other hand, differences at baseline in these risk factors or differences in the intensity of sleep hygiene regimen between the two groups should be considered: (*1*) differences for BMI and tobacco between groups at baseline were observed, and they were included in the model as possible confounding factors to be adjusted; (*2*) the effect of giving up tobacco was tested, and the results showed no statistically significant influence in improvement (only three patients gave up tobacco); and (*3*) groups of treatment did not differ owing to the decrease in BMI during the first treatment period of 6 wk.

The lack of statistical significance of some of the outcome variables merits further comment. Nonsignificant differences between the A and B groups in the FOSQ total and SF-36 scores after optimal or sham CPAP may be due to the small number of patients included in the study, which was calculated according to the expected differences in the ESS and its standard deviation. Despite this, for the FOSQ total score, the differences observed in the changes between the groups during the first treatment period were substantial (10 points). No effect was detected by the FOSQ in the second approach, but the ceiling effect is a likely explanation for this, given that in most of the FOSQ scales more than 70% of patients presented the best possible score in measurement 4. For the SF-36, the negative results could seem surprising given the sensitivity to changes, particularly on the Vitality subscale, in the patients with SAHS (22–24). However, the Vitality scores observed in our patients after optimal CPAP treatment (mean of 69 in Group A and 75 in Group B) are very close to previous findings (22–24), and also to the SF36 norms (18). The main difference concerns the Vitality score observed by Jenkinson and coworkers after treatment with suboptimal CPAP, which was worse than in the group of patients with optimal CPAP (50.9 versus 73.0) (9). However, in our study, the group of patients with sham CPAP achieved Vitality scores that were very similar to those observed in the group with optimal CPAP (mean of 68). As stated previously, this Vitality improvement in the group with sham CPAP could be attributed to the adoption of a complementary sleep hygiene regimen.

Although the short period of study could invite some criticism, we do believe that a longer follow-up would not have been ethical given the ample evidence of treatment benefits (3–8). On the other hand, the beneficial effects of CPAP are likely to be maintained over time compared with the disappointing outcome commonly observed with the conservative measures in other studies.

In conclusion, this study provides solid evidence of the effectiveness of CPAP treatment in moderate to severe SAHS to alleviate somnolence and other SAHS-related symptoms and to improve health status. The use of CPAP in symptomatic patients is now adequately supported, and the indication of CPAP in these patients should be considered as firmly established.

***Acknowledgment*:** The authors thank the nurses of the respiratory ward and the technicians of the sleep laboratory for their support.

References

1. Sullivan CE, Issa F, Berthon-Jones M, Eves L. Reversal of obstructive sleep apnea by continuous positive airway pressure applied through the nares. *Lancet* 1981;1:862–865.
2. Wright J, Johns R, Watt I, Melville A, Sheldon T. Health effects of obstructive apnea and the effectiveness of continuous positive airway pressure: a systematic review of the research evidence. *BMJ* 1997;314: 851–860.
3. Engleman HM, Martin SE, Douglas NJ. Effect of CPAP therapy on daytime function in patients with mild sleep apnea/hypopnoea syndrome. *Thorax* 1997;52:114–119.
4. Engleman HM, Martin SE, Douglas NJ. Effect of continuous positive airway pressure on daytime function in sleep apnea/hypopnoea syndrome. *Lancet* 1994;343:572–575.
5. Engleman HM, Asgari-Jirhandeh N, McLeod AL, Ramsay CF, Deary IJ, Douglas NJ. Self-reported use of CPAP and benefits of CPAP therapy. *Chest* 1996;109:1470–1476.
6. Engleman HM, Chesire KE, Deary IJ, Douglas NJ. Daytime sleepiness, cognitive performance and mood after continuous positive pressure for the sleep apnea/hypopnoea syndrome. *Thorax* 1993;48:911–914.
7. Redline S, Adams N, Strauss ME, Roebuck T, Winters M, Rosenberg C. Improvement of mild sleep disordered breathing with CPAP compared with conservative therapy. *Am J Respir Crit Care Med* 1998;57: 858–865.
8. Ballester E, Badia JR, Hernandez L, Carrasco E, De Pablo J, Fornas C, Rodriguez-Roisin R, Montserrat JM. Evidence of the effectiveness of CPAP in the treatment of sleep apnea/hypopnea syndrome. *Am J Respir Crit Care Med* 1999;159:495–501.
9. Jenkinson C, Davies RJO, Mullins R, Stradling JR. Comparison of therapeutic and subtherapeutic nasal continuous positive airway pressure for obstructive sleep apnoea: a randomised prospective parallel trial. *Lancet* 1999;353:2100–2105.
10. Farre R, Hernandez L, Montserrat JM, Rotger M, Ballester E, Navajas D. Sham continuous positive airway pressure for placebo-controlled studies in sleep apnea. *Lancet* 1999;353:1154.
11. Ferrer M, Vilagut G, Monasterio C, Montserrat JM, Mayos M, Alonso J. [Measurement of the perceived impact of sleep problems: the Spanish version of the Functional Outcomes Sleep Questionnaire and the Epworth Sleepiness Scale] Medida del impacto de los trastornos del sueno: las versiones españolas del cuestionario del impacto funcional del sueño y de la escala de somnolencia de Epworth. *Med Clin (Barc)* 1999;113:250–255.
12. Lloberes P, Montserrat JM, Ascaso A, Parra O, Granados A, Alonso P, Vilaseca I, Rodriguez-Roisin R. Comparison of partially attended night time respiratory recordings and full polysomnography in patients with suspected sleep apnea/hypopnoea syndrome. *Thorax* 1996;51:1043–1047.
13. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* 1991;14:540–545.
14. Weaver TE, Laizner A, Evans LK, Maislin G, Chugh DK, Lyons K, Smith P, Schwartz AR, Redline S, Pack AI, *et al.* An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep* 1997;20:835–843.
15. McHorney CA, Ware JE, Lu JF, Sherbourne CD. The MOS 36-item Short Form Health Survey (SF-36): tests of data quality, scaling assumptions, and reliability across diverse patients groups. *Med Care* 1994;32:40–66.
16. Johns MW. Daytime sleepiness, snoring, and obstructive sleep apnea: the Epworth Sleepiness Scale. *Chest* 1993;103:30–36.
17. Ware JE Jr, Kosinski M, Gandek B, Aaronson NK, Apolone G, Bech P, Brazier J, Bullinger M, Kaasa S, Leplège A, *et al.* The factor structure of the SF-36 Health Survey in 10 countries: results from the IQOLA Project. *J Clin Epidemiol* 1998;51:1159–1165.
18. Alonso J, Regidor E, Barrio G, Prieto, L, Rodríguez C, de la Fuente L. [Population reference values of the Spanish version of the Health Questionnaire SF-36] Valores poblacionales de referencia de la versión española del Cuestionario de Salud SF-36. *Med Clin (Barc)* 1998; 111:410–416.

19. Stradling J. Sleep apnea and the misuse of evidence-based medicine. *Lancet* 1997;349:201–202.
20. Engleman HM, Martin SE, Deary IJ, Douglas NJ. Obstructive sleep apnea: some criticisms of studies are unfounded (letter). *BMJ* 1997;15:369.
21. Levinson PD, McGarvey ST, Carlisle CC, Eveloff SE, Herbert PN, Millman RP. Adiposity and cardiovascular risk factors in men with obstructive sleep apnea. *Chest* 1993;103:1336–1342.
22. Jenkinson C, Stradling J, Petersen S. Comparison of three measures of quality of life outcome in the evaluation of continuous positive airways pressure therapy for sleep apnoea. *J Sleep Res* 1997;6:199–204.
23. Jenkinson C, Stradling J, Petersen S. How should we evaluate health status? A comparison of three methods in patients presenting with obstructive sleep apnoea. *Qual Life Res* 1998;7:95–100.
24. D'Ambrosio C, Bowman T, Mohsenin V. Quality of life in patients with obstructive sleep apnea: effect of nasal continuous positive airway pressure—a prospective study. *Chest* 1999;115:123–129.