Effects of Continuous Positive Airway Pressure on Quality of Life in Patients With Moderate to Severe Obstructive Sleep Apnea: Data From a Randomized Controlled Trial

Massimiliano M. Siccoli, MD[1]; Justin C.T. Pepperell, MD, FRCP[2]; Malcolm Kohler, MD[1]; Sonya E. Craig, MD, MRCP[1]; Robert J.O. Davies, MD, FRCP[1]; John R. Stradling, MD, FRCP[1]

[1]*Sleep Unit, Centre for Respiratory Medicine, Churchill Hospital, Oxford, UK;* [2]*Somerset Lung Centre, Musgrove Park Hospital, Taunton, UK*

Study Objectives: Previous studies have shown that CPAP has a substantial impact on daytime symptoms and quality of life (QOL). It remains unclear which outcome measures best identify real CPAP effects and carry independent information.

Methods: One hundred-two men with moderate-severe obstructive sleep apnea were randomized to either "real" or "sham" CPAP for one month. Outcome measures were subjective sleepiness (Epworth Sleepiness Scale [ESS]) and QOL measures including SF-36/SF-12 and Calgary Sleep Apnea Quality of Life Index (SAQLI). The bed partner's QOL and rating of patient's response to CPAP were assessed with the Dublin questionnaire. All data were standardized using effect sizes and expressed as real minus sham to remove the nonspecific effects of placebo.

Results: Real CPAP was superior to sham CPAP in almost all outcome measures. ESS, patient's component from Dublin, and social interactions from SAQLI showed the largest differences in effect sizes between real and sham (1.33, 0.98, and 0.92 respectively). ESS carried the highest predictive power of real CPAP response ($P < 0.0001$, $r^2 = 0.21$). Question number 5 from Dublin (partner assessed patient's sleep quality) and question 6 from ESS (dozing while talking) were the best single item predictors of real CPAP response.

Conclusions: Real CPAP reduces subjective sleepiness and improves QOL of both patients and bed partners. ESS is the best score; question number 5 from Dublin and question number 6 from ESS are the best single item predictors of real CPAP response. This information should allow the selection of appropriate questions in clinical practice and research protocols.

Keywords: Obstructive sleep apnea, CPAP, quality of life, health status, randomized, controlled, Epworth score, Dublin questionnaire

Citation: Siccoli MM; Pepperell JCT; Kohler M; Craig SE; Davies RJO; Stradling JR. Effects of continuous positive airway pressure on quality of life in patients with moderate to severe obstructive sleep apnea: data from a randomized controlled trial. *SLEEP* 2008;31(11):1551-1558.

OBSTRUCTIVE SLEEP APNEA (OSA) IS CHARACTERIZED BY REPETITIVE APNEAS AND HYPOPNEAS DURING SLEEP ASSOCIATED WITH OXYGEN DESATURATION, leading to sleep disruption and excessive daytime sleepiness. It has been estimated that OSA is present in 9% to 24%[1] of adults, being clinically significant in 2% to 4%,[1] and becoming more prevalent as the average population body weight rises.[2] Excessive daytime sleepiness related to increased sleep fragmentation and impaired sleep quality clearly has an adverse impact on quality of life and health status.[3]

Nasal continuous positive airways pressure (CPAP) is considered to be the treatment of choice for OSA. Excessive daytime sleepiness remains the main indication for CPAP treatment in OSA,[4] although cardiovascular benefits are thought to result as well. Results from previous studies, including a randomized double-blind controlled parallel trial of therapeutic versus subtherapeutic CPAP showed that CPAP has beneficial effects on excessive daytime sleepiness and on self-reported functioning and well-being, affecting quality of life of OSA patients.[4-8] The cost-effectiveness of this treatment has also been demonstrated.[9]

Submitted for publication February, 2008
Accepted for publication July, 2008
Address correspondence to: Massimiliano M. Siccoli, MD, Oxford Centre for Respiratory Medicine, John Radcliffe Hospital, Churchill Hospital Site, Headington, Oxford, OX3 7LJ, United Kingdom, Tel: +44 1865 225 227; Fax: +44 1865 225 221; E-mail: siccolim@gmx.net

Despite an increasing amount of information on the efficacy of CPAP treatment in OSA patients, data from randomized double-blind controlled trials on treatment outcome measures that included quality of life assessment have been reported in only 2 previous studies.[4,7] Particularly, it remains unclear which aspects of symptomatic improvement are the most sensitive to CPAP treatment and carry independent information in predicting response to CPAP.

To address this uncertainty we investigated, using data from a randomized double-blind controlled trial, which outcome measures best identify the effect of real CPAP compared to nonspecific placebo effects. The overall aim of this study is to improve the current knowledge about the efficacy of CPAP treatment on symptoms and health status. This issue may be relevant in clinical practice and for future treatment trials.

METHODS

Patients

Patients with possible obstructive sleep apnea referred to the Oxford Sleep Unit, Oxford Centre for Respiratory Medicine, UK, by general practitioners, ear, nose, and throat surgeons or other hospital consultants were considered for inclusion. Patients were eligible for the trial if they were males aged between 20 and 75 years who had excessive daytime sleepiness (Epworth Sleepiness Scale [ESS] score[10] ≥10) and proven obstruc-

tive sleep apnea with > 10 oxygen desaturations > 4% per hour (oxygen desaturation index [ODI] > 10/h). All eligible patients were offered participating in the study unless they required urgent CPAP therapy because of respiratory failure, driving or job-related issues. Of the 102 randomized patients in this trial, 52 had been involved in a previously published study evaluating the effect of CPAP on ambulatory blood pressure.[11]

The study was approved by the Oxford research ethics committee (COREC No 96.127), and written informed consent was obtained from all participants before inclusion.

Sleep Study

OSA was assessed with a one-night in-hospital respiratory polysomnographic sleep study. Patient's body movements, heart rate, and pulse transit time (PTT) changes were recorded as measures of arousal from sleep. Pulse oximetry, snoring, and increases in the respiratory swing in PTT were used as markers of breathing pattern and respiratory effort (Win-Visi monitoring system, Stowood Scientific Instruments, Oxford, UK) as previously described and validated.[11-13]

The results of the sleep study were scored automatically, with manual review to ensure accuracy of the data. OSA was diagnosed from review of all data and the severity was quantified as the number of oxygen desaturations > 4% per hour of study (ODI).

Daytime Sleepiness

Excessive subjective daytime sleepiness was assessed by using the ESS,[10] the most widely used index to measure subjective sleepiness in OSA. Objective sleepiness was measured using a modified maintenance of wakefulness test, the Osler test.[14] Mean time to sleep onset was considered a measure of the objective sleepiness and integrated in the CPAP/placebo comparison. The test was carried out at the same time of day on the 2 occasions patients attended for the study. Patients were asked to refrain from caffeine on the day of examination.

Self-Reported Health Status and Quality of Life

Self-reported health status was assessed by using different questionnaires. The change in each variable was calculated using effect sizes. Effect sizes are defined as the mean change divided by the original standard deviation (SD) of the population and therefore represent the number of SDs the population has shifted. Individual effect sizes are defined as the individual change divided by the original standard deviation of the population. An effect size of 0.2 was considered small, 0.5 medium, and 0.8 or higher large. The sham effect size was then subtracted from the real effect size to provide a measure of specific CPAP effect.

The Medical Outcome Study 36-Item Short-Form Health Survey (SF-36)[15,16]

This is a 36-item questionnaire which measures 8 domains of health: physical functioning, physical problems, emotional problems, social functioning, mental health, energy/vitality, pain, as well as general perception of health. Scores obtained from each domain are coded, summed and transformed onto a scale ranging from 0 (worst possible) to 100 (best possible) to calculate two summary scores expressing physical (physical component summary) and emotional (mental component summary) well-being. The domain and summary scores are standardized such that a mean score of 50 with a standard deviation of 10 would reflect the mean score of the Oxford population established from previous lifestyle surveys.[17] This questionnaire has been widely used to assess quality of life and health status in different disorders, including obstructive sleep apnea, showing high reliability, validity, and responsiveness.[15,18]

The 12-Item Short-Form Health Survey (SF-12)

This is a shorter health survey derived from the SF-36, allowing faster assessment of patients and producing physical and emotional component summaries without any substantial loss of information compared to the SF-36. The summary scores were calculated using the procedure indicated by the developers and published in detail elsewhere[19-22] and transformed onto a scale ranging from 0 (worst possible health) to 100 (best possible health), analogous to SF-36 scores. The validity of SF-12 has been also evaluated in patients with sleep apnea under CPAP treatment, showing results virtually identical to those of the SF-36.[22]

The Calgary Sleep Apnea Quality of Life Index (SAQLI)

This test was originally developed as a disease-specific instrument to evaluate health-related quality of life in patients with sleep apnea for use in clinical trials, and successfully tested for validity, consistency, and reliability in patients with OSA under CPAP treatment.[23,24] It consists of 45 items organized in 6 domains: daily functioning (domain A), social interactions (domain B), emotional functioning (domain C), patient selected or nominated symptoms potentially due to OSA (domain D), and treatment-related symptoms (domain E). The score of each item ranges from 0 (worst possible) to 7 (best possible). The scores for each domain are expressed as mean scores of all items in the same domain, and the total score is the mean score of all domains. Since the domain of treatment-related symptoms (E) could be assessed only at follow-up, and because it does not represent a change in the related measure, it was not considered in the calculation of effect sizes.

The Dublin Bed Partner's Assessment[25,26]

This questionnaire is a tool designed to provide a subjective assessment of bed partner's change in health status including sleep quality, daytime alertness, mood, and overall quality of life (questions 1 to 4, partner's component). In addition, bed partners estimate the same parameters for the patients (questions 5 to 8, patient's component) and estimate the change of his/her personal relationship with the patient (question 9) since the beginning of the CPAP treatment. Hence, the question number 9 was considered both as a main component of the score and as a single question. The score of each item ranges from –1 (worsening) to 3 (marked improvement). Dublin bed partner's

assessment was performed only at follow-up. The calculation of the corresponding theoretical effect sizes was possible because the scores represent a change in the measure and were divided by the overall SD of these changes. Thus, this is slightly different from the other effect size calculations, but unlikely to introduce any significant bias.

CPAP and Assessment of Sleepiness

After enrolment, patients were randomly assigned to either therapeutic (real) or subtherapeutic (sham) CPAP, and then underwent a night of CPAP titration, during which respiratory polysomnography was repeated and CPAP was used. For patients assigned to real CPAP, the therapeutic pressure was determined from overnight use of the Sullivan Autoset-T auto-adjusting (ResMed, Abingdon, UK) CPAP machine,[27] from which mask pressure was recorded and synchronized with the sleep study signals. The record was reviewed the next morning, and the optimum pressure to prevent sleep apnea was confirmed by a sleep technician. A conventional CPAP with fixed pressure was then given to patients to continue treatment. Patients assigned to sham CPAP used a machine that delivered <1 cm H_2O pressure as previously described,[4,11] being insufficient to hold the pharynx open.

Patients and investigators remained blinded to real or sham CPAP assignment. Patients were told that we were comparing 2 CPAP pressures to find out which was more effective in controlling their symptoms, and that one might be more effective than the other. Since they had never experienced CPAP before, there was no reason for patients to realize that the lower pressure might be subtherapeutic. The sleep nurse, who randomly assigned patients to the 2 groups, maintained the machines and assisted the patients, was not involved in outcome assessments.

Subjective sleepiness (ESS), objective sleepiness (Osler test), and self-reported health status were assessed before randomization as indicated.

Follow-Up

Four weeks after baseline assessments, the patients were readmitted for repeated measurements of ESS, Osler test, and self-reported health status. Hour meters on the CPAP machines were downloaded to calculate mean nightly use. At the end of the trial, CPAP pressure was retitrated in every patient to establish subsequent long-term therapeutic pressure.

Data Analysis

All statistical analyses were performed with SPSS software (version 15.0.0). Baseline characteristics, subjective and objective sleepiness data, and self-reported health status data are expressed as mean and standard deviation. The Student *t* tests were used for comparison of baseline characteristics between the real and sham group and for comparison of the self-reported health status data measured at baseline and after 4 weeks between and within the real and sham group on an intention-to-treat basis, with no change assumed when follow-up data were missing. When data were analyzed as per protocol, all statistically significant differences persisted; therefore data from these analyses are not shown. The same procedure was followed for the comparison of effect sizes between and within groups. $P < 0.05$ was considered significant.

In order to identify the most sensitive and specific predictors of a real CPAP response, binary logistic regression models were tested using a forward conditional method with entry value set to 0.05 and removal value set to 0.1; assignment to either CPAP or placebo group was the dependent variable. First, we explored the power to predict in which group the patient was, using the summary components/scores from the different health outcome measures, i.e., sleepiness (ESS), QOL (physical and mental component summary from the SF-36/SF-12), and composite of sleepiness and QOL (total score from the SAQLI). No summary component was available for the bed partner's Dublin. Second, the main components from the different questionnaires (SF-36/SF-12, SAQLI, bed partner's Dublin) were tested with the other main components of the same questionnaire and then in a combined model. Third, the single questions within the ESS and within the bed partner's Dublin (single questions from SF-36/SF-12 and from SAQLI were not considered as outcome measures) were tested in a combined model to detect the best single question predicting which group the patients were in.

Table 1—Characteristics of Patients with Sham and Real CPAP

	"Sham" CPAP n = 51	"Real" CPAP n = 51	P value
Age (years)	48.7 ± 10.6	48.1 ± 9.5	0.76
BMI (kg/m^2)	34.5 ± 5.0	35.8 ± 7.3	0.30
Neck circumference (cm)	44.6 ± 3.3	45.1 ± 4.0	0.54
Waist/hip circumference ratio	1.01 ± 0.06	1.02 ± 0.06	0.93
ESS at baseline	15.2 ± 4.0	15.8 ± 4.0	0.48
Osler at baseline (minutes)	17.3 ± 13.1	18.1 ± 13.1	0.75
Oxygen saturation dips >4% (per hour of sleep)	42.7 ± 21.6	41.9 ± 25.4	0.87
CPAP compliance (h/night)	3.9 ± 2.5	4.7 ± 2.1	0.08
Retitration CPAP pressure following study (cm H_2O)	10.1 ± 1.6	10.0 ± 1.9	0.72

Values are means and standard deviations. BMI = body mass index; ESS = Epworth Sleepiness Scale score; CPAP = continuous positive airway pressure. There were no significant differences between the 2 groups for any of the measures.

RESULTS

Trial Profile and Patient Characteristics

The trial profile is shown in figure 1. A total of 102 patients aged 48.1 ± 10.1 years were randomized, 51 to real CPAP and 51 to sham CPAP. The 2 groups did not differ in baseline characteristics, including demographics (age, body mass index, neck circumference, and waist-hip circumference ratio), sleepiness data, number of oxygen saturation dips per hour, and CPAP compliance (Table 1).

Table 2—Subjective (Total Epworth Sleepiness Score and Single Questions) and Objective Sleepiness with Corresponding Effect Sizes, Before/After Real and Sham CPAP

	Before (Baseline)	After (Follow-up)	Effect size	P before/ after	Difference in effect sizes*	P for effect sizes*
"Sham" CPAP						
ESS, total score	15.2 ± 4.0	11.9 ± 5.9	**0.82**	**0.001**		
Q1 (sitting)	2.3 ± 0.9	1.9 ± 1.1	0.48	**0.009**		
Q2 (watching TV)	2.6 ± 0.7	1.9 ± 1.0	**0.99**	**<0.0001**		
Q3 (public place)	1.8 ± 1.0	1.3 ± 1.0	**0.52**	**0.003**		
Q4 (car passenger)	2.2 ± 1.0	1.7 ± 1.1	0.43	**0.001**		
Q5 (lying down)	2.7 ± 0.6	2.4 ± 0.9	**0.63**	**0.003**		
Q6 (talking)	0.9 ± 0.8	0.7 ± 0.9	0.23	**0.040**		
Q7 (after lunch)	2.0 ± 0.8	1.6 ± 1.0	0.45	**0.013**		
Q8 (traffic light)	0.6 ± 0.8	0.4 ± 0.7	0.41	**0.003**		
Osler test (min)	17.3 ± 13.1	18.3 ± 14.3	0.08	0.58		
CPAP use (h/night)		3.9 ± 2.5				
'Real" CPAP						
ESS, total score	15.8 ± 4.0	6.8 ± 5.1	**2.15**	**<0.0001**	**1.33**	**<0.0001**
Q1 (sitting)	2.4 ± 0.8	1.0 ± 0.9	**1.59**	**<0.0001**	**1.11**	**<0.0001**
Q2 (watching TV)	2.5 ± 0.7	1.1 ± 0.9	**2.02**	**<0.0001**	**1.03**	**<0.0001**
Q3 (public place)	2.0 ± 0.8	0.7 ± 0.8	**1.39**	**<0.0001**	**0.87**	**<0.0001**
Q4 (car passenger)	2.2 ± 1.0	0.9 ± 1.0	**1.23**	**<0.0001**	**0.79**	**<0.0001**
Q5 (lying down)	2.7 ± 0.6	1.7 ± 1.0	**1.64**	**<0.0001**	**1.01**	**0.002**
Q6 (talking)	1.2 ± 0.9	0.3 ± 0.6	**1.07**	**<0.0001**	**0.84**	**<0.0001**
Q7 (after lunch)	2.1 ± 0.8	0.9 ± 1.0	**1.41**	**<0.0001**	**0.96**	**<0.0001**
Q8 (driving a car)	0.7 ± 0.8	0.2 ± 0.5	**0.76**	**<0.0001**	0.35	0.09
Osler test (min)	18.1 ± 13.1	26.8 ± 12.9	**0.67**	**<0.0001**	**0.59**	**0.007**
CPAP use (hours/night)		4.7 ± 2.1				

Values are means and standard deviations. ESS = Epworth Sleepiness Scale score; Q = ESS, single questions; CPAP = continuous positive airway pressure. *P values calculated for comparison between real and sham CPAP group. P, Bold = statistically significant. Effect size, Bold = medium-large effect size (>0.5). Effect size difference, Bold = medium-large difference (>0.5).

Effect Sizes

Excessive Daytime Sleepiness

Table 2 shows the subjective and objective sleepiness before and after real and sham CPAP. Real CPAP significantly reduced subjective sleepiness measured by the ESS ($P < 0.0001$), and improved objective sleepiness measured by Osler test ($P < 0.0001$). With sham CPAP, ESS improved less than real CPAP, but the improvement remained significant ($P < 0.0001$). In contrast, no improvement was observed in objective sleepiness with sham CPAP. The difference in effect sizes between sham and real CPAP was significant for both ESS and Osler, but larger for ESS. The large effect size in the sham group for ESS (0.82) combined with the very low effect size for Osler (0.08) in this group suggests a considerable placebo effect.

Single Questions from the Epworth Sleepiness Score

The 8 scores of the single ESS question, as well as the corresponding effect sizes, are reported in Table 2. All scores improved significantly after treatment in both sham and real CPAP groups. The improvement in all single scores was significantly higher ($P < 0.0001$) in the real group, except for the question number 8 (dozing at traffic light), in which no significant difference could be observed between the 2 groups. Questions 1 (dozing while sitting), 2 (dozing while watching TV), and 5 (dozing while lying down) showed the highest difference (>1.00) in effect sizes between real and sham CPAP. In none of the single questions was the difference in effect sizes larger than in the total ESS.

Self-Reported Health Status

Table 3 shows the results of the SF-36 and SF-12 questionnaires with the corresponding effect sizes. Most scores of the main components (except for the domains of physical functioning and physical role) and the 2 component summary scores of both SF-36 and SF-12 significantly improved with real CPAP. In contrast, with sham CPAP, significant improvement was observed only in the components of reported health transition, physical role, and energy/vitality. The highest differences in effect size between the real and the sham group were found in energy/vitality, reported health transition, mental health, and in the mental component summary.

Table 4 shows the results of SAQLI with the corresponding effect sizes. The scores of all SAQLI components significantly improved with real CPAP, except for component D (symptoms of OSA, i.e., fatigue, subjective sleepiness, dry mouth, waking up often, difficulty returning to sleep, morning headache). The questionnaire design means that in component D, different symptoms may be selected or nominated by the patient on different occasions; therefore a strict "before" and "after" comparison is not possible, with the after scores possibly including

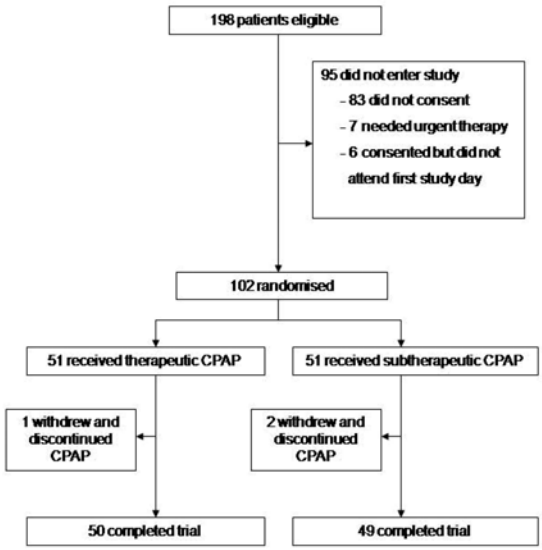

Figure 1—Trial Profile

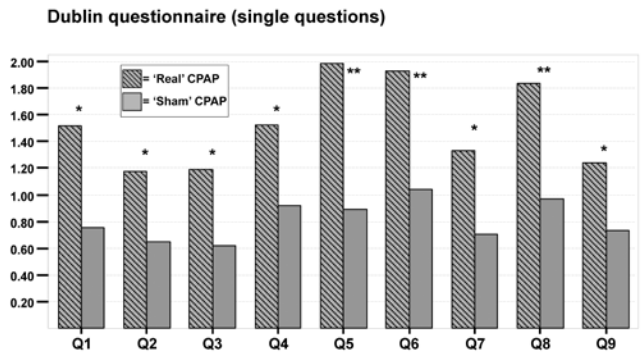

Figure 2—Effect sizes of the single questions from the bed partner's Dublin questionnaire in patients with real and sham CPAP. *P values < 0.05, calculated for comparison of effect sizes between real and sham CPAP group; **P values < 0.001, calculated for comparison of effect sizes between real and sham CPAP group.

treatment related side effects. Medium-large effect sizes were observed in all components of SAQLI, except for component D (nominated symptoms of OSA), whereas component B (social interactions) showed the largest effect size. Interestingly, component E (measuring potential CPAP side effects) was worse (nonsignificant) in the real group, supporting the suggestion that component D was worse for similar reasons.

Table 4 displays the scores of the main components from the bed partner's Dublin questionnaire with their respective effect sizes. Medium-large effect sizes were observed in both groups for all 3 main components (partner's component, patient's component, and personal relationship), and were significantly higher in the real group. Regarding the single items, questions number 5 (partner assessed patient's sleep quality), 6 (partner assessed patient's daytime alertness), and 8 (partner assessed patient's quality of life) showed the largest differences in effect sizes (Figure 2) between real and sham group.

Predictive Power of Outcome Measures (Binary Logistic Regression Models)

Considering all the summary components/scores from the different health outcome measures (ESS, SF-36/SF-12, SAQLI), ESS was the best (most sensitive and specific) in predicting CPAP response ($P < 0.0001$), explaining 21.2% of variance in the model ($r^2 = 0.21$).

In the analysis of the main components from the different questionnaires (SF-36, SAQLI, bed partner's Dublin), energy/vitality from SF-36 ($P = 0.001$, $r^2 = 0.11$), social interactions (B) from SAQLI ($P < 0.0001$, $r^2 = 0.23$), and partner assessed patient's component from the Dublin questionnaire ($P < 0.0001$, $r^2 = 0.27$) carried the highest predictive power in each respective questionnaire. From all 3 questionnaires, social interactions (B) from SAQLI best predicted real CPAP response.

In the single question analysis, question number 5 from the Dublin questionnaire (partner assessed patient's sleep quality) and question number 6 from the ESS (dozing while talking) were the best single item predictors of CPAP response, explaining 43.4% of variance in a combined model ($P = 0.001$, $r^2 = 0.43$).

DISCUSSION

The aim of our study was to systematically investigate different outcome measures of health related QOL in patients with therapeutic (real) versus subtherapeutic (sham) CPAP, to identify the best predictors of real response to CPAP compared to nonspecific placebo effects, and to improve the current understanding about the impact of CPAP on health status.

To the best of our knowledge, this is the most comprehensive work with randomized double-blind controlled trial data on measures of QOL and health status in patients with obstructive sleep apnea undergoing CPAP treatment. In particular, data from the bed partner's assessment of health status have not been reported previously in the same context.

We found that (1) real CPAP was superior to sham CPAP in almost all tested outcome measures of health status; ESS, the bed partner's assessed patient's component from the Dublin questionnaire, and the social interactions from SAQLI (main component B) showed the largest differences in effect sizes between real and sham CPAP; the placebo effect was generally larger than expected in the sham group. (2) Among the different scores/questionnaires, ESS was the best in predicting real CPAP response; energy/vitality from SF-36, social interactions (B) from SAQLI, and the bed partner assessed patient's component from the Dublin questionnaire carried the highest predictive power from each questionnaire respectively; question number 5 from the Dublin questionnaire (partner assessed patient's sleep quality) and question 6 from the ESS (dozing while talking) were the best single item predictors of real CPAP response.

Effect Sizes of Outcome Measures

Excessive daytime sleepiness, the main symptom of obstructive sleep apnea, improved with real CPAP, consistent with findings of previous studies.[4,6] The observed difference of 1.33 in effect size between real and sham was very large. However, the

Table 3—Self-Reported Health Status Scores from the SF-36 and SF-12 Questionnaires and Corresponding Effect Sizes, Before and After Real and Sham CPAP

	Before (Baseline)	After (Follow-up)	Effect size	P before/ after	Difference in effect sizes*	P for effect sizes*
"Sham" CPAP						
SF-36						
General health perception	62.0 ± 21.5	64.3 ± 23.9	0.09	0.41		
Reported health transition	44.6 ± 21.4	52.7 ± 26.2	**0.56**	**0.020**		
Physical functioning	80.1 ± 18.5	82.1 ± 16.4	0.11	0.22		
Physical role	63.7 ± 41.3	51.1 ± 28.5	-0.27	**0.016**		
Mental role	63.7 ± 39.7	77.5 ± 35.9	0.06	0.37		
Social functioning	71.2 ± 30.8	77.3 ± 28.7	0.19	0.19		
Bodily pain	83.4 ± 23.6	83.7 ± 20.3	0.03	0.77		
Energy and vitality	38.5 ± 23.4	52.6 ± 26.7	**0.58**	**<0.0001**		
Mental health	72.1 ± 17.7	72.0 ± 22.1	-0.05	0.70		
Physical CS	69.4 ± 21.5	70.0 ± 18.8	0.04	0.68		
Mental CS	64.8 ± 21.2	68.6 ± 22.7	0.16	0.17		
SF-12						
Physical CS	66.2 ± 20.8	69.8 ± 20.1	0.18	0.09		
Mental CS	66.8 ± 21.1	70.6 ± 22.6	0.19	0.13		
"Real" CPAP						
SF-36						
General health perception	55.0 ± 22.7	62.8 ± 23.4	0.34	**0.003**	0.24	0.12
Reported health transition	46.1 ± 24.7	65.5 ± 23.1	**1.21**	**<0.0001**	**0.65**	**0.001**
Physical functioning	72.5 ± 21.5	72.8 ± 22.5	0.01	0.90	–0.10	0.38
Physical role	54.4 ± 39.9	63.0 ± 22.7	0.18	0.16	0.45	**0.008**
Mental role	69.3 ± 43.1	90.7 ± 26.1	0.48	**<0.0001**	0.41	**0.021**
Social functioning	74.3 ± 28.1	84.0 ± 23.0	0.30	**0.030**	0.10	0.60
Bodily pain	66.2 ± 29.3	73.1 ± 30.8	0.22	0.10	0.19	0.27
Energy and vitality	36.1 ± 21.1	64.7 ± 20.4	**1.26**	**<0.0001**	**0.68**	**0.001**
Mental health	72.6 ± 16.5	81.0 ± 16.1	0.46	**<0.0001**	**0.51**	**0.003**
Physical CS	62.0 ± 20.0	70.8 ± 18.5	0.39	**<0.0001**	0.35	**0.010**
Mental CS	62.2 ± 20.2	76.8 ± 16.2	**0.66**	**<0.0001**	**0.51**	**0.002**
SF-12						
Physical CS	58.8 ± 20.6	72.4 ± 18.8	**0.64**	**<0.0001**	0.46	**0.002**
Mental CS	63.5 ± 19.9	77.9 ± 16.5	**0.70**	**<0.0001**	**0.51**	**0.001**

Values are means and standard deviations. CS = component summary. *P values calculated for comparison between real and sham"CPAP group. P, Bold = statistically significant. Effect size, Bold = medium-large effect size (>0.5). Effect size difference, Bold = medium-large difference (>0.5).

large effect size of 0.82 in the sham group, despite no improvement in objective sleepiness, indicates a larger placebo effect than one would expect. This may be explained in different ways. First, patients suffering with chronic and disabling disease such as obstructive sleep apnea are likely to have high expectations of their treatment and are told that sleepiness would improve. Second, subjective daytime sleepiness measured by the ESS is likely to include components of other mood dimensions, such as depression or anxiety, which would be expected to respond to sympathetic professional interactions between staff and patients. Third, subjective sleepiness may be partially improved or reversed by just "feeling better" after having received a new treatment, even without any change in sleep fragmentation and objective sleepiness. Assuming that sham CPAP does improve some aspects of real OSA symptoms via a psychological effect, this would have led to a smaller difference between real and sham treatment, and thus to an underestimation of the true CPAP response. Hence, the differences between real and sham CPAP represent the "worst case scenario," and the real CPAP response might even be higher than quoted here. In contrast, objective sleepiness significantly only improved with real CPAP. This confirms that true objective sleepiness is not affected by a nonspecific improvement in well-being ("feeling better").

The analysis of the single ESS questions revealed a significant effect of real CPAP on all items. Question 1 (dozing while sitting), 2 (dozing while watching TV), and 5 (dozing while lying down) carried the largest differences in effect sizes, indicating that these 3 items—along with the total ESS score—are particularly responsive to real CPAP treatment. Question number 8 (dozing at traffic light) carried the lowest effect size (difference in effect size within and between real and sham group), possibly because patients tend to underestimate their own sleepiness in this situation or fear consequences regarding their driving license.

Real CPAP improved in most of the SF-36 domain scores and in the SF-12 component summary scores. The largest differences in effect sizes were observed in the components of energy/vitality, reported health transition, mental health, and

Table 4—Self-Reported Health Status Scores from the SAQLI and from the Main Components of the Bed Partner's Dublin Questionnaire with Corresponding Effect Sizes, Before and After Real and Sham CPAP

	Before (Baseline)	After (Follow-up)	Effect size	P before/ after	Difference in effect sizes*	P for effect sizes*
"Sham" CPAP						
SAQLI						
A (Daily routine)	3.9 ± 1.4	4.8 ± 1.6	**0.72**	**<0.0001**		
B (Social interactions)	4.5 ± 1.4	5.0 ± 1.5	0.43	**0.001**		
C (Emotional functioning)	4.3 ± 1.4	5.0 ± 1.5	0.44	**<0.0001**		
D (Symptoms of OSA)	2.5 ± 1.1	2.8 ± 1.4	0.27	0.18		
E (Treatment related symptoms)		2.3 ± 1.7				
Total score	3.8 ± 1.1	3.8 ± 1.6	0.08	0.65		
Bed partner's Dublin						
Partner's component (Q1-Q4)		0.9 ± 1.0	**0.84**			
Patient's component (Q5-Q8)		1.1 ± 1.2	**0.94**			
Personal relationship (Q9)		0.9 ± 1.1	**0.74**			
"Real" CPAP						
SAQLI						
A (Daily routine)	3.7 ± 1.2	5.5 ± 1.0	**1.38**	**<0.0001**	**0.66**	**0.003**
B (Social interactions)	4.0 ± 1.3	5.8 ± 1.2	**1.34**	**<0.0001**	**0.92**	**<0.0001**
C (Emotional functioning)	4.0 ± 1.3	5.4 ± 1.0	**1.02**	**<0.0001**	**0.58**	**<0.0001**
D (Symptoms of OSA)	2.4 ± 1.0	2.2 ± 1.6	-0.19	0.44	-0.46	0.15
E (Treatment related symptoms)		2.1 ± 1.5				
Total score	3.5 ± 1.0	4.4 ± 1.1	**0.84**	**<0.0001**	**0.76**	**0.001**
Bed partner's Dublin						
Partner's component (Q1-Q4)		1.6 ± 1.0	**1.55**		**0.71**	**0.005**
Patient's component (Q5-Q8)		2.2 ± 0.8	**1.92**		**0.98**	**<0.0001**
Personal relationship (Q9)		1.5 ± 1.2	**1.24**		**0.50**	**0.041**

Values are means and standard deviations. *P values calculated for comparison between real and sham CPAP group. P, Bold = statistically significant. Effect size, Bold = medium-large effect size 8 (>0.5). Effect size difference, Bold = medium-large difference (>0.5).

in the mental component summary. The impact of real CPAP on SF-36 and SF-12 outcome measures is similar to that reported in previous treatment trials,[4-6] and larger than reported in patients with mild OSA.[7,28] This means that our data, which relate to patients with moderate-severe OSA, might not reflect responses seen in patients with mild disease.

The higher placebo component in the domain of energy/vitality and reported health transition confirm that improvement of distinct domains of well-being may be achieved with a nonspecific treatment, independent of any improvement in sleep fragmentation and objective sleepiness. In contrast, the significant difference in effect sizes observed in the mental components of SF-36/SF-12, with little placebo response, suggests clear differences across the dimensions in their response to real versus sham treatment. These results fit well with the medium-large difference in effect sizes observed in the components of daily routine (A), social interactions (B), and emotional functioning (C) from the SAQLI.

The results of the Dublin bed partner's assessment deserve a separate mention. We found both in the main domain scores and in the single questions, large differences in effect sizes between the sham and real group. This observation suggests that bed partners of patients with OSA experience important improvements in their own condition, perceive improvements in the patient's condition and their relationship following both real and sham CPAP treatment, but this improvement is higher with real CPAP. This confirms the large placebo component already observed in the other questionnaires, suggesting that partner's expectations of improvement in their own and patient's QOL might be even higher than the patient's own expectations.

Predictors of Real CPAP Response

As additional information to the effect sizes, which simply reflect the magnitude of the response to real or sham treatment, binary logistic regression was performed to identify the best and independent (most sensitive and specific) items distinguishing real CPAP responses from nonspecific placebo effects.

Of all the summary components/scores, ESS was the best in predicting real CPAP response. This means that subjective sleepiness remains the main symptom improving with CPAP treatment.

In the combined analysis of the best single questions, question number 5 from the Dublin questionnaire (partner assessed patient's sleep quality) and question 6 from the ESS (dozing while talking) were the best single item predictors of real CPAP response, indicating their high sensitivity and specificity, respectively. These highly predictive single questions might potentially be useful as screening questions in studies looking at large numbers of patients or individuals.

In summary, ESS remains an essential instrument in evaluating real response to CPAP, but clearly contains components responding to the placebo effect. Our study further supports the use of self-reported QOL measures as a useful tool for future randomized controlled trials, contributing to the knowledge about the specific impact of CPAP treatment on the different

components of health status. Finally, the results support the role of the patients and partners in the evaluation of the health benefits of CPAP.

ABBREVIATIONS

Q1, partner's sleep quality; Q2, partner's daytime alertness; Q3, partner's mood; Q4, partner's quality of life; Q5, partner assessed patient's sleep quality; Q6, partner assessed patient's daytime alertness; Q7, partner assessed patient's mood, Q8, partner assessed patient's quality of life, Q9, personal relationship.

DISCLOSURE STATEMENT

This was not an industry supported study. Dr. Davies has received the use of equipment from ResMed for research. Dr. Stradling has received the use of equipment from ResMed for research and has participated in a recent speaking engagement for a company that produces an asthma drug. The other authors have indicated no financial conflicts of interest.

REFERENCES

1. Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. N Engl J Med 1993;328:1230-5.
2. Young T, Peppard PE, Gottlieb DJ. Epidemiology of obstructive sleep apnea: a population health perspective. Am J Respir Crit Care Med 2002;165:1217-39.
3. Lacasse Y, Godbout C, Series F. Health-related quality of life in obstructive sleep apnoea. Eur Respir J 2002;19:499-503.
4. Jenkinson C, Davies RJ, 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;353:2100-5.
5. 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.
6. Jenkinson C, Davies RJ, Mullins R, Stradling JR. Long-term benefits in self-reported health status of nasal continuous positive airway pressure therapy for obstructive sleep apnoea. The Quarterly journal of medicine 2001;94:95-9.
7. Engleman HM, Kingshott RN, Wraith PK, Mackay TW, Deary IJ, Douglas NJ. Randomized placebo-controlled crossover trial of continuous positive airway pressure for mild sleep apnea/hypopnea syndrome. Am J Respir Crit Care Med 1999;159:461-7.
8. Engleman HM, Martin SE, Kingshott RN, Mackay TW, Deary IJ, Douglas NJ. Randomised placebo controlled trial of daytime function after continuous positive airway pressure (CPAP) therapy for the sleep apnoea/hypopnoea syndrome. Thorax 1998;53:341-5.
9. Giles TL, Lasserson TJ, Smith BH, White J, Wright J, Cates CJ. Continuous positive airways pressure for obstructive sleep apnoea in adults. Cochrane Database Syst Rev (Online) 2006.
10. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 1991;14:540-5.
11. Pepperell JC, Ramdassingh-Dow S, Crosthwaite N, et al. Ambulatory blood pressure after therapeutic and subtherapeutic nasal continuous positive airway pressure for obstructive sleep apnoea: a randomised parallel trial. Lancet 2002;359:204-10.
12. Pitson DJ, Stradling JR. Autonomic markers of arousal during sleep in patients undergoing investigation for obstructive sleep apnoea, their relationship to EEG arousals, respiratory events and subjective sleepiness. J Sleep Res 1998;7:53-60.
13. Argod J, Pepin JL, Levy P. Differentiating obstructive and central sleep respiratory events through pulse transit time. Am J Respir Crit Care Med 1998;158:1778-83.
14. Bennett LS, Stradling JR, Davies RJO. A behavioural test to assess daytime sleepiness in obstructive sleep apnoea. J Sleep Res 1997;6:142-45.
15. Tarlov AR, Ware JE Jr, Greenfield S, Nelson EC, Perrin E, Zubkoff M. The Medical Outcomes Study: an application of methods for monitoring the results of medical care. JAMA 1989;262:925-30.
16. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473-83.
17. Wright L, Harwood D, Coulter A. Health and lifestyles in the Oxford Region. Oxford: Health Services Research Unit, University of Oxford, 1992.
18. Jenkinson C, Stradling JR, 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.
19. Ware JE, Kosinski M, Keller SD. SF-12: an even shorter health survey. Med Outcomes Trust Bull 1996;4:2.
20. Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996;34:220-33.
21. Jenkinson C, Layte R. Development and testing of the UK SF-12 (short form health survey). J Health Serv Res Policy 1997;2:14-8.
22. Jenkinson C, Layte R, Jenkinson D, et al. A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? J Public Health Med 1997;19:179-86.
23. Flemons WW, Reimer MA. Development of a disease-specific health-related quality of life questionnaire for sleep apnea. Am J Respir Crit Care Med 1998;158:494-503.
24. Flemons WW, Reimer MA. Measurement properties of the Calgary sleep apnea quality of life index. Am J Respir Crit Care Med 2002;165:159-64.
25. Kiely JL, McNicholas WT. Bed partners› assessment of nasal continuous positive airway pressure therapy in obstructive sleep apnea. Chest 1997;111:1261-5.
26. Doherty LS, Kiely JL, Lawless G, McNicholas WT. Impact of nasal continuous positive airway pressure therapy on the quality of life of bed partners of patients with obstructive sleep apnea syndrome. Chest 2003;124:2209-14.
27. Stradling JR, Barbour C, Pitson DJ, Davies RJ. Automatic nasal continuous positive airway pressure titration in the laboratory: patient outcomes. Thorax 1997;52:72-5.
28. Engleman HM, Martin SE, Deary IJ, Douglas NJ. Effect of CPAP therapy on daytime function in patients with mild sleep apnoea/hypopnoea syndrome. Thorax 1997;52:114-9.