

Relationship Between Hours of CPAP Use and Achieving Normal Levels of Sleepiness and Daily Functioning

Terri E. Weaver, PhD, RN, FAAN¹⁻³; Greg Maislin, MS, MA¹⁻³; David F. Dinges, PhD^{3,4}; Thomas Bloxham, MD⁵; Charles F. P. George, MD⁶; Harly Greenberg, MD⁷; Gihan Kader, MD⁸; Mark Mahowald, MD⁹; Joel Younger, MD¹⁰; Allan I. Pack, MD, PhD^{2,3}

¹Biobehavioral and Health Sciences Division, School of Nursing, University of Pennsylvania, Philadelphia, PA; ²Division of Sleep Medicine, Department of Medicine, School of Medicine, University of Pennsylvania, Philadelphia, PA; ³Center for Sleep and Respiratory Neurobiology, School of Medicine, University of Pennsylvania, Philadelphia, PA; ⁴Division of Sleep and Chronobiology, Department of Psychiatry, School of Medicine, University of Pennsylvania, Philadelphia, PA; ⁵Sleep Medicine Center of Kansas, Wichita, KS ⁶University of Western Ontario, London, Ontario, Canada; ⁷North Shore/Long Island Jewish Health System, New Hyde Park, NY; ⁸St. Luke's Hospital, Chesterfield, MO; ⁹Hennepin County Medical Center, Minneapolis, MN; ¹⁰Bon Secours Holy Family Regional Health System, Altoona, PA

Study Objectives: Evidence suggests that, to maintain treatment effects, nasal continuous positive airway pressure (CPAP) therapy for obstructive sleep apnea (OSA) needs to be used every night. What remains unknown is the nightly duration of use required to normalize functioning. This study, employing probit analyses and piecewise regression to estimate dose-response functions, estimated likelihoods of return to normal levels of sleepiness and daily functioning relative to nightly duration of CPAP.

Design: Multicenter, quasi-experimental study.

Setting: Seven sleep centers in the United States and Canada.

Participants: Patients with severe OSA (total cohort n = 149; the numbers of included participants from 85 - 120, depending on outcome analyzed.)

Interventions: CPAP.

Measurements and Results: Before treatment and again after 3 months of therapy, participants completed a day of testing that included measures of objective and subjective daytime sleepiness and functional status. There were significant differences in mean nightly CPAP duration between treatment responders and nonresponders across outcomes. Thresholds above

which further improvements were less likely relative to nightly duration of CPAP were identified for Epworth Sleepiness Scale score (4 hours), Multiple Sleep Latency Test (6 hours), and Functional Outcomes associated with Sleepiness Questionnaire (7.5 hours). A linear dose-response relationship ($P < 0.01$) between increased use and achieving normal levels was shown for objective and subjective daytime sleepiness, but only up to 7 hours use for functional status.

Conclusions: Our analyses suggest that a greater percentage of patients will achieve normal functioning with longer nightly CPAP durations, but what constitutes adequate use varies between different outcomes.

Keywords: Obstructive sleep apnea, dose response, nightly duration, daytime sleepiness, quality of life, CPAP, adherence, daily functioning, alertness

Citation: Weaver TE; Maislin G; Dinges DF et al. Relationship between hours of CPAP use and achieving normal levels of sleepiness and daily functioning. *SLEEP* 2007;30(6):711-719.

Disclosure Statement

This study was an investigator initiated study supported primarily with funds from NHLBI. Additional funds were procured to carry out the study from: Respirationics, Nellcore Puritan Bennett, DeVilbiss Health Care Inc., and Healthdyne Technologies. The industries supporting this study had no input into the conduct or aims of the study and were not given the opportunity to provide input or review the manuscript. Dr. Weaver has received research support from Protech and Respirationics; has served as a consultant for Sanofi-Aventis; participated in a continuing education program sponsored by a grant from Cephalon; and has received license fees for the Functional Outcomes of Sleep Questionnaire from Jazz Pharmaceuticals, Sanofi-Aventis, Organon NV, Sleep Solutions, Aspire Medical, and Influent Medical. Dr. Dinges has received research support from Cephalon and Mars Masterfoods; has participated in speaking engagements for Cephalon and Takeda; and has consulted for Cephalon, Mars Masterfoods, Pfizer, and Merck. Dr. George is on the medical advisory board for SleepTech LLC and has received research support from Orphan Medical Inc., and Xyrem. Dr. Mahowald has received research support from Sanofi-Aventis, Advanced Medical Electronics, Merck, Kyowa Pharmaceutical, and Schwarz Biosciences. Dr. Pack has consulted for Cypress Biosciences and has participated in speaking engagements for Merck and Respirationics.

Submitted for publication October 2006

Accepted for publication February 2007

Address correspondence to: Terri E. Weaver, PhD, RN, FAAN, University of Pennsylvania School of Nursing, Claire M. Fagin Hall, 418 Curie Blvd., Philadelphia, PA 19104-6096; Tel: (215) 898-2992; Fax: (215) 573-7492; E-mail: tew@nursing.upenn.edu

CONTINUOUS POSITIVE AIRWAY PRESSURE (CPAP) CURRENTLY IS CONSIDERED THE MOST EFFECTIVE TREATMENT FOR OBSTRUCTIVE SLEEP APNEA (OSA). THIS device provides a pneumatic splint to prevent nocturnal airway collapse. In randomized clinical trials that included use of sham CPAP as a placebo, treatment with this device has been shown to produce improvements in symptoms, quality of life, sleepiness, neuropsychological performance, and hypertension.¹ It has also been established that optimal effectiveness depends on consistent use. Skipping even 1 night of treatment reverses improvements in daytime sleepiness, response performance, and the physiologic measure of disease severity, the apnea-hypopnea index (AHI).²⁻⁴

Although these benefits have been shown in randomized clinical trials, there is considerable variation in use of CPAP by patients in routine clinical practice,⁵⁻⁸ with approximately half using it consistently every night on average 6 hours per night and the other half skipping from 1 to 7 nights per week using it on average 3.5 hours per night.⁵ It is not known, however, what the impact of differential use of CPAP is on effectiveness of therapy in routine clinical practice. It may be that individuals with shorter hours of use are, in general, less well treated with respect to improvements in outcomes such as sleepiness, than are those with longer durations of use. Alternatively, there might be individual variation in need for CPAP so that, even though some individuals have shorter hours of CPAP use, they are effectively treated with respect to sleepiness. This individual variation might be mediated by biologic mechanisms similar to those of the recently described

individual variation in response to sleep deprivation.⁹

To address this question, we conducted a multisite effectiveness study in which we evaluated clinical status outcomes before and after 3 months of routine clinical care that included measurement of CPAP adherence during the entire follow-up period. This multisite international study capitalized on patient heterogeneity with regard to average nightly use of CPAP, in order to estimate relationships between the likelihood of achieving a “normal” value on salient clinical measures of sleepiness and functional status and the “dose” of CPAP treatment received (i.e., hours of CPAP use per night). Because there is no consensus regarding which outcomes are primary with regard to the assessment of CPAP treatment response, we chose those deemed most applicable to the clinical management of OSA—subjective sleepiness (Epworth Sleepiness Scale [ESS]¹⁰), physiologic sleepiness (Multiple Sleep Latency Test [MSLT]¹¹), and disease-specific functional status (Functional Outcomes of Sleep Questionnaire [FOSQ]¹²). By evaluating the nightly duration of CPAP relative to obtaining normal values in those impaired on each of these measures, we provide data that the clinician can employ in determining optimal treatment response within the context of treatment goals, e.g., to reduce daytime sleepiness, improve daily functioning, or both.

MATERIALS AND METHODS

Participants

Participants attended 1 of 7 sleep disorders centers in the United States and Canada. Selection criteria included diagnosis of OSA, age 21 to 60 years; AHI of 15 or greater, and a candidate for CPAP. Participants were excluded if they had a history of coexisting pulmonary disease, congestive heart failure, cerebrovascular accident, or psychiatric illness; used sedative-hypnotic medication; had another sleep disorder; or were older than 60 years because of common age-related changes that could affect response to treatment such as central apneas. The study was approved by the Institutional Review Board at each institution.

Polysomnography

Because this was an effectiveness and not a treatment-efficacy study, methods used in routine clinical practice were employed. Patients participating in the study underwent either an attended in-laboratory, full-night, diagnostic polysomnogram followed by a titration polysomnogram to determine the therapeutic level of CPAP pressure or a split-night study ($n = 38$ out of $n = 149$) during which the diagnosis of sleep apnea was made in the first half of the night followed by a determination of effective CPAP pressure during the second half of the night. Each site conducted the polysomnogram in accordance with their routine procedures and guidelines established by the American Academy of Sleep Medicine¹³ and scoring in accordance with the standard methods of Rechtschaffen and Kales¹⁴ using electroencephalogram, electrooculogram, electromyogram, and electrocardiogram. Airflow was recorded using nasal thermistors, thoracic abdominal respiratory movements were documented by respiratory inductance plethysmography, and arterial oxygen saturation was measured continuously using a finger oximeter. The study was initiated prior to the demonstration of the superiority of nasal pressure to assess respiratory events.¹⁵ The AHI was computed from the number of

obstructive apneas and hypopneas ($> 50\%$ reduction in airflow) according to standard criteria.¹⁶

Adherence Monitor

Participants received the CPAP device and mask prescribed for them by the physician managing their sleep apnea. Adherence to CPAP treatment was documented using an overt monitor attached to the CPAP machine.⁶ Using an external monitor allowed the physician and participant choice of devices as well as the utilization of software to monitor adherence that we have developed and has known validity and reliability.⁶ This monitor employs a microprocessor that utilizes an algorithm for the detection of mask-on pressure. The CPAP device electrical cord was plugged into the external monitor so that, when the monitor was turned on, the CPAP device was turned on simultaneously. Drops in therapeutic pressure greater than 5 cm and lasting longer than 10 seconds were documented as a mask-off event. The monitor logged the time and date of 4 events—machine on/off and mask on/off—for each 24-hour period. Thus, the duration of therapeutic pressure delivery per day and number of missed days of use could be determined. Only durations of 20 minutes or longer per daily episode were considered CPAP “use,” and adherence was calculated as the mean of the hours of daily use (including 0 hrs of use) over the entire follow-up period up to the day before posttreatment testing. Participants were informed in advance that their CPAP use was being monitored.

Assessments

Demographic data, medical history, and sleep-habit information on each subject were obtained using standard instruments.

Subjective daytime sleepiness was measured using the ESS.¹⁰ The ESS is a well validated and reliable instrument that has been employed extensively in sleep apnea research to measure a patient’s perceived likelihood of falling asleep in a subset of daytime activities. A score is produced by summing the responses for each item, with scores ranging from 0 to 24. Scores greater than 10 have a sensitivity of 93.5% and specificity of 100% to distinguish pathologic from normal daytime sleepiness.¹⁰ This, therefore, was the cutpoint used to define normal subjective sleepiness.

Objective daytime sleepiness was measured employing the research application of the MSLT.¹¹ The MSLT measures, polysomnographically, the time required for the participant to fall asleep (latency time) in a 20-minute period, when attempting to do so. The test was administered at 2-hour intervals during the test day starting at 10:00 AM for a total of 4 test bouts, with the reported value calculated as the average latency across the test bouts. Sleep latency of less than 5 minutes is the conventional value for defining pathologic sleepiness; while values of 10 minutes or longer are considered normal.¹¹ Values at least equal to 5 but less than 10 minutes are considered in the “grey” zone.^{10,11,17} Because there were too few participants in our study with an MSLT value less than 5 minutes ($n = 59$), we chose a criterion of 7.5 minutes, which is midway between these 2 values (i.e., we considered abnormal to be < 7.5 minutes versus normal ≥ 7.5 minutes). This provides a sample size for this outcome of 85 participants who had an MSLT of less than 7.5 minutes before therapy.

Functional status, a component of quality of life, was measured using the FOSQ. The FOSQ is a 30-item Likert-style question-

naire with 5 domains that examine the impact of being sleepy or tired on the conduct of daily activities.¹² It has established validity and reliability. A Total score is generated from the mean item scores for each subscale and ranges from 5 to 20. A Total score value of less than 17.9 was chosen as the cutpoint for abnormal scores on the FOSQ because this was the mean of a sample ($n = 20$) of normal individuals free of any sleep disorders verified through polysomnography (unpublished data) and because the mean Total score of the participants in this study ($N = 147$) at baseline was 14.1 (3.7), making 17.9 approximately 1 standard deviation above the mean. Thus, using 17.9 as the cutpoint excludes participants with values above the mean among normal individuals and above the mean plus 1 standard deviation among participants with moderate to severe sleep apnea.

Procedures

Prior to their diagnostic or split-night polysomnogram and initiation of CPAP, after providing informed written consent, participants completed a day of testing using a predetermined schedule of test administration. Testing began with the administration of the demographics questionnaire, the ESS, the FOSQ, and the MSLT, in addition to a test battery (results not reported here) that included the Pittsburgh Sleep Quality Index, Profile of Mood States, and Sickness Impact Profile Scale, followed by assessments of neurobehavioral performance (Psychomotor Vigilance Task, Digit Symbol Substitution Test, Probed Memory Recall Test) and the Stanford Sleepiness Scale. After the day of testing, participants were provided their CPAP devices with the adherence monitor attached. Following 3 months of treatment, participants returned to the laboratory, where they underwent the same testing protocol that they had experienced prior to treatment. Participants were instructed to obtain their usual duration of sleep and use their CPAP devices as they had the previous nights.

Statistical Analyses

The sample characteristics were evaluated using summary statistics with values given as mean (standard deviation), medians, and ranges, or proportions as appropriate. In the determination of the effective dose necessary to return those with pathologic values to normal functioning, only participants with “abnormal” response values for each outcome prior to treatment were included in the dose-response analyses. As described in Methods, baseline impairment was defined separately for each outcome variables: ESS greater than 10,¹⁰ MSLT less than 7.5 minutes, and FOSQ Total score less than 17.9. Within these subsamples, a participant was defined to have *responded to treatment* if his or her posttreatment value no longer fell into the measure-specific “abnormal” range. The coefficient of variation of average nightly hours of CPAP use ($CV = 41.4\%$ for the entire sample and 46.8%, 51.2%, and 47.8%, respectively for the subsamples used in the ESS, MSLT, and FOSQ analysis) confirmed that the variability in CPAP nightly duration was sufficient to permit analysis of CPAP-use dose response with respect to hours of nightly use.

To examine the relationship between CPAP duration and outcome, a piecewise regression analysis¹⁸ was performed for each of the 3 cohorts with impairments as defined by ESS ($n = 106$), MSLT ($n = 85$), and FOSQ ($n = 120$) using the definitions described above. Posttreatment outcomes were assessed as continu-

ous variables. The objective of these analyses was to investigate whether there were threshold values for minimum mean CPAP-use values beyond which greater improvements in sleepiness and function should be expected. This analysis estimates separate dose-response relationships of participants with mean CPAP use below a specified threshold (first segment) and above a specified threshold (second segment). Since it is unreasonable to expect discontinuities in dose-response curves, a “join” model was estimated in which the final expected response value in first segment is forced to be equal to the first expected response value in the second segment. The slopes are allowed to vary between segments. For each outcome variable, join points from 1 to 7.5 hours of mean CPAP use in 0.5-hour increments were evaluated, and the first join point that maximized explained variance was selected for each clinical outcome. Complementary probit analyses¹⁹ were also undertaken to test for the presence of a linear dose response in terms of clinical outcomes defined as the elimination of baseline impairment. Analyses were again done using each outcome variable separately and with the different subsamples of participants that met impairment criteria for each pretreatment outcome. These analyses provided estimates of the sensitivity of increases in the likelihood of returning to normal functioning as CPAP mean nightly duration increased. F-tests¹⁸ were used to determine if the join-point dose-response model was significantly better than a simple linear dose model.

RESULTS

Participant Characteristics

Prior to their baseline polysomnogram, 300 participants were recruited and provided written consent. Of those 300, 124 were eliminated because they did not meet the study inclusion criteria (e.g., an $AHI < 15$, $n = 21$), 10 did not return for follow-up testing, and CPAP adherence data for 17 participants were lost due to technical reasons (monitor transformer breakage, monitor battery failure, corruption of electronic data during data transfer) resulting in 149 study participants (130 men, 19 women). Only participants having complete pretreatment and posttreatment data and also abnormal scores at baseline comprised the sample employed for each analysis, ranging from 85 to 120 participants for the different outcomes. Thus, the sample sizes were 106 for the analysis of response to treatment using the ESS, 85 for the MSLT, and 120 for the FOSQ. Drawing from the data of the 149 participants, 137 had complete baseline data on all 3 outcome measures. Of these, 69 (50.4%) were impaired pretreatment on all 3 variables, and 5 (5.8%) were not impaired on any of the 3 measures after 3 months of therapy. Of the 128 participants with complete posttreatment data, only 14 participants (10.9%) remained impaired across all 3 measures, and 41 (32.0%) did not have values indicating impairment on any of the 3 outcome variables.

The demographic characteristics for these participants and their degree of sleep apnea are summarized in Table 1. Additionally, 87.2% were white, 8.8% were African American, 79.2% were married, 92.4% completed high school, and 33.8% completed at least 4 years of college. Thus, the sample studied was predominantly white, male, middle-aged, and quite obese and, in general, had severe sleep apnea. Six participants were included in the cohort ($n = 149$), whose age values were deviations from the protocol but were not felt to be sufficiently clinically important

individual variation in response to sleep deprivation.⁹

To address this question, we conducted a multisite effectiveness study in which we evaluated clinical status outcomes before and after 3 months of routine clinical care that included measurement of CPAP adherence during the entire follow-up period. This multisite international study capitalized on patient heterogeneity with regard to average nightly use of CPAP, in order to estimate relationships between the likelihood of achieving a “normal” value on salient clinical measures of sleepiness and functional status and the “dose” of CPAP treatment received (i.e., hours of CPAP use per night). Because there is no consensus regarding which outcomes are primary with regard to the assessment of CPAP treatment response, we chose those deemed most applicable to the clinical management of OSA—subjective sleepiness (Epworth Sleepiness Scale [ESS]¹⁰), physiologic sleepiness (Multiple Sleep Latency Test [MSLT]¹¹), and disease-specific functional status (Functional Outcomes of Sleep Questionnaire [FOSQ]¹²). By evaluating the nightly duration of CPAP relative to obtaining normal values in those impaired on each of these measures, we provide data that the clinician can employ in determining optimal treatment response within the context of treatment goals, e.g., to reduce daytime sleepiness, improve daily functioning, or both.

MATERIALS AND METHODS

Participants

Participants attended 1 of 7 sleep disorders centers in the United States and Canada. Selection criteria included diagnosis of OSA, age 21 to 60 years; AHI of 15 or greater, and a candidate for CPAP. Participants were excluded if they had a history of coexisting pulmonary disease, congestive heart failure, cerebrovascular accident, or psychiatric illness; used sedative-hypnotic medication; had another sleep disorder; or were older than 60 years because of common age-related changes that could affect response to treatment such as central apneas. The study was approved by the Institutional Review Board at each institution.

Polysomnography

Because this was an effectiveness and not a treatment-efficacy study, methods used in routine clinical practice were employed. Patients participating in the study underwent either an attended in-laboratory, full-night, diagnostic polysomnogram followed by a titration polysomnogram to determine the therapeutic level of CPAP pressure or a split-night study ($n = 38$ out of $n = 149$) during which the diagnosis of sleep apnea was made in the first half of the night followed by a determination of effective CPAP pressure during the second half of the night. Each site conducted the polysomnogram in accordance with their routine procedures and guidelines established by the American Academy of Sleep Medicine¹³ and scoring in accordance with the standard methods of Rechtschaffen and Kales¹⁴ using electroencephalogram, electrooculogram, electromyogram, and electrocardiogram. Airflow was recorded using nasal thermistors, thoracic abdominal respiratory movements were documented by respiratory inductance plethysmography, and arterial oxygen saturation was measured continuously using a finger oximeter. The study was initiated prior to the demonstration of the superiority of nasal pressure to assess respiratory events.¹⁵ The AHI was computed from the number of

obstructive apneas and hypopneas ($> 50\%$ reduction in airflow) according to standard criteria.¹⁶

Adherence Monitor

Participants received the CPAP device and mask prescribed for them by the physician managing their sleep apnea. Adherence to CPAP treatment was documented using an overt monitor attached to the CPAP machine.⁶ Using an external monitor allowed the physician and participant choice of devices as well as the utilization of software to monitor adherence that we have developed and has known validity and reliability.⁶ This monitor employs a microprocessor that utilizes an algorithm for the detection of mask-on pressure. The CPAP device electrical cord was plugged into the external monitor so that, when the monitor was turned on, the CPAP device was turned on simultaneously. Drops in therapeutic pressure greater than 5 cm and lasting longer than 10 seconds were documented as a mask-off event. The monitor logged the time and date of 4 events—machine on/off and mask on/off—for each 24-hour period. Thus, the duration of therapeutic pressure delivery per day and number of missed days of use could be determined. Only durations of 20 minutes or longer per daily episode were considered CPAP “use,” and adherence was calculated as the mean of the hours of daily use (including 0 hrs of use) over the entire follow-up period up to the day before posttreatment testing. Participants were informed in advance that their CPAP use was being monitored.

Assessments

Demographic data, medical history, and sleep-habit information on each subject were obtained using standard instruments.

Subjective daytime sleepiness was measured using the ESS.¹⁰ The ESS is a well validated and reliable instrument that has been employed extensively in sleep apnea research to measure a patient’s perceived likelihood of falling asleep in a subset of daytime activities. A score is produced by summing the responses for each item, with scores ranging from 0 to 24. Scores greater than 10 have a sensitivity of 93.5% and specificity of 100% to distinguish pathologic from normal daytime sleepiness.¹⁰ This, therefore, was the cutpoint used to define normal subjective sleepiness.

Objective daytime sleepiness was measured employing the research application of the MSLT.¹¹ The MSLT measures, polysomnographically, the time required for the participant to fall asleep (latency time) in a 20-minute period, when attempting to do so. The test was administered at 2-hour intervals during the test day starting at 10:00 AM for a total of 4 test bouts, with the reported value calculated as the average latency across the test bouts. Sleep latency of less than 5 minutes is the conventional value for defining pathologic sleepiness; while values of 10 minutes or longer are considered normal.¹¹ Values at least equal to 5 but less than 10 minutes are considered in the “grey” zone.^{10,11,17} Because there were too few participants in our study with an MSLT value less than 5 minutes ($n = 59$), we chose a criterion of 7.5 minutes, which is midway between these 2 values (i.e., we considered abnormal to be < 7.5 minutes versus normal ≥ 7.5 minutes). This provides a sample size for this outcome of 85 participants who had an MSLT of less than 7.5 minutes before therapy.

Functional status, a component of quality of life, was measured using the FOSQ. The FOSQ is a 30-item Likert-style question-

naire with 5 domains that examine the impact of being sleepy or tired on the conduct of daily activities.¹² It has established validity and reliability. A Total score is generated from the mean item scores for each subscale and ranges from 5 to 20. A Total score value of less than 17.9 was chosen as the cutpoint for abnormal scores on the FOSQ because this was the mean of a sample ($n = 20$) of normal individuals free of any sleep disorders verified through polysomnography (unpublished data) and because the mean Total score of the participants in this study ($N = 147$) at baseline was 14.1 (3.7), making 17.9 approximately 1 standard deviation above the mean. Thus, using 17.9 as the cutpoint excludes participants with values above the mean among normal individuals and above the mean plus 1 standard deviation among participants with moderate to severe sleep apnea.

Procedures

Prior to their diagnostic or split-night polysomnogram and initiation of CPAP, after providing informed written consent, participants completed a day of testing using a predetermined schedule of test administration. Testing began with the administration of the demographics questionnaire, the ESS, the FOSQ, and the MSLT, in addition to a test battery (results not reported here) that included the Pittsburgh Sleep Quality Index, Profile of Mood States, and Sickness Impact Profile Scale, followed by assessments of neurobehavioral performance (Psychomotor Vigilance Task, Digit Symbol Substitution Test, Probed Memory Recall Test) and the Stanford Sleepiness Scale. After the day of testing, participants were provided their CPAP devices with the adherence monitor attached. Following 3 months of treatment, participants returned to the laboratory, where they underwent the same testing protocol that they had experienced prior to treatment. Participants were instructed to obtain their usual duration of sleep and use their CPAP devices as they had the previous nights.

Statistical Analyses

The sample characteristics were evaluated using summary statistics with values given as mean (standard deviation), medians, and ranges, or proportions as appropriate. In the determination of the effective dose necessary to return those with pathologic values to normal functioning, only participants with “abnormal” response values for each outcome prior to treatment were included in the dose-response analyses. As described in Methods, baseline impairment was defined separately for each outcome variables: ESS greater than 10,¹⁰ MSLT less than 7.5 minutes, and FOSQ Total score less than 17.9. Within these subsamples, a participant was defined to have *responded to treatment* if his or her posttreatment value no longer fell into the measure-specific “abnormal” range. The coefficient of variation of average nightly hours of CPAP use ($CV = 41.4\%$ for the entire sample and 46.8%, 51.2%, and 47.8%, respectively for the subsamples used in the ESS, MSLT, and FOSQ analysis) confirmed that the variability in CPAP nightly duration was sufficient to permit analysis of CPAP-use dose response with respect to hours of nightly use.

To examine the relationship between CPAP duration and outcome, a piecewise regression analysis¹⁸ was performed for each of the 3 cohorts with impairments as defined by ESS ($n = 106$), MSLT ($n = 85$), and FOSQ ($n = 120$) using the definitions described above. Posttreatment outcomes were assessed as continu-

ous variables. The objective of these analyses was to investigate whether there were threshold values for minimum mean CPAP-use values beyond which greater improvements in sleepiness and function should be expected. This analysis estimates separate dose-response relationships of participants with mean CPAP use below a specified threshold (first segment) and above a specified threshold (second segment). Since it is unreasonable to expect discontinuities in dose-response curves, a “join” model was estimated in which the final expected response value in first segment is forced to be equal to the first expected response value in the second segment. The slopes are allowed to vary between segments. For each outcome variable, join points from 1 to 7.5 hours of mean CPAP use in 0.5-hour increments were evaluated, and the first join point that maximized explained variance was selected for each clinical outcome. Complementary probit analyses¹⁹ were also undertaken to test for the presence of a linear dose response in terms of clinical outcomes defined as the elimination of baseline impairment. Analyses were again done using each outcome variable separately and with the different subsamples of participants that met impairment criteria for each pretreatment outcome. These analyses provided estimates of the sensitivity of increases in the likelihood of returning to normal functioning as CPAP mean nightly duration increased. F-tests¹⁸ were used to determine if the join-point dose-response model was significantly better than a simple linear dose model.

RESULTS

Participant Characteristics

Prior to their baseline polysomnogram, 300 participants were recruited and provided written consent. Of those 300, 124 were eliminated because they did not meet the study inclusion criteria (e.g., an $AHI < 15$, $n = 21$), 10 did not return for follow-up testing, and CPAP adherence data for 17 participants were lost due to technical reasons (monitor transformer breakage, monitor battery failure, corruption of electronic data during data transfer) resulting in 149 study participants (130 men, 19 women). Only participants having complete pretreatment and posttreatment data and also abnormal scores at baseline comprised the sample employed for each analysis, ranging from 85 to 120 participants for the different outcomes. Thus, the sample sizes were 106 for the analysis of response to treatment using the ESS, 85 for the MSLT, and 120 for the FOSQ. Drawing from the data of the 149 participants, 137 had complete baseline data on all 3 outcome measures. Of these, 69 (50.4%) were impaired pretreatment on all 3 variables, and 5 (5.8%) were not impaired on any of the 3 measures after 3 months of therapy. Of the 128 participants with complete posttreatment data, only 14 participants (10.9%) remained impaired across all 3 measures, and 41 (32.0%) did not have values indicating impairment on any of the 3 outcome variables.

The demographic characteristics for these participants and their degree of sleep apnea are summarized in Table 1. Additionally, 87.2% were white, 8.8% were African American, 79.2% were married, 92.4% completed high school, and 33.8% completed at least 4 years of college. Thus, the sample studied was predominantly white, male, middle-aged, and quite obese and, in general, had severe sleep apnea. Six participants were included in the cohort ($n = 149$), whose age values were deviations from the protocol but were not felt to be sufficiently clinically important

Table 1—Participant Characteristics at Baseline

Characteristic	All Participants n = 149	ESS ≥ 11 n = 106	MSLT < 7.5 min n = 85	FOSQ Total Score <17.9 n = 120
Men, %	87.3	85.6	84.7	87.5
Age, y	46.8 (8.8)	46.1 (8.4)	44.4 (8.6)	46.2 (8.6)
BMI, kg/m ²	38.0 (8.1)	38.5 (8.4)	40.0 (8.5)	38.1 (8.5)
AHI, no./h	64.1 (29.1)	67.6 (27.9)	69.2 (29.9)	65.4 (29.6)
SaO ₂ nadir during				
NREM	73.5 (17.4)	73.2 (17.5)	71.3 (16.9)	72.8 (14.7)
REM	66.9 (18.5)	66.8 (18.8)	64.2 (18.3)	68.3 (18.0)
MSLT, min	7.2 (5.1)	6.5 (4.8)	3.7 (1.9)	6.9 (5.1)
ESS, score	14.7 (4.8)	16.7 (3.3)	15.7 (4.3)	16.0 (4.2)
FOSQ, Total score	14.7 (3.1)	14.0 (2.7)	14.2 (2.9)	13.8 (2.6)
PSQI, Global score	8.4 (3.7)	8.9 (3.4)	8.8 (3.6)	9.0 (3.7)
Nightly CPAP duration, h	4.7 (2.1)	4.7 (2.2)	4.3 (2.2)	4.6 (2.2)
Subjects who used CPAP for given number of hours per night, %				
≥ 6	45.1 (31.5)	44.6 (32)	38.8 (30.9)	43.5 (30.9)
≥ 5	57.8 (32.1)	57.1 (33.1)	51.8 (32.8)	56.1 (32)
≥ 4	66.1 (31.2)	65.1 (32.2)	60.7 (32.9)	64.4 (31.4)
≥ 3	71.6 (29.6)	70.5 (30.6)	63.9 (31.5)	70 (30.1)
Any use	79.5 (25.9)	78.6 (26.9)	76.5 (27.9)	77.8 (26.6)

Data are presented as mean ± SD unless otherwise indicated. ESS, refers to Epworth Sleepiness Scale; MSLT, Multiple Sleep Latency Test; FOSQ, Functional Outcomes of Sleep Questionnaire; BMI, body mass index; AHI, apnea-hypopnea index; REM, rapid eye movement; NREM, non-rapid eye movement; PSQI; Pittsburgh Sleep Quality Index; CPAP continuous positive airway pressure;

to warrant exclusion from the analyses. One participant was less than 21 years of age (age 19) and included in the analyses for all 3 outcomes. The age of 6 participants was greater than the study criteria of being younger than 60 years. Of these, 2 participants' data were not included in any analyses because their values were normal at baseline or the data were incomplete. Among the remaining 4 participants, all were younger than 66 years of age; the data of 1 were included in all the analyses, and the other 3 had abnormal values on at least 1 variable and their data were included in that analysis.

Relationship Among Outcome Measures

As expected, there was a strong relationship between the 2 subjective measures at baseline and after treatment (ESS Score and FOSQ: baseline $r = -0.61$, $P < .0001$; posttreatment ($r = -0.64$, $P < 0.0001$). The relationship was less robust between the MSLT, an objective measure, and the 2 subjective measures (ESS: baseline $r = 0.31$, $P = 0.0002$; posttreatment $r = 0.23$, $P = 0.008$; FOSQ: baseline $r = 0.17$, $P < 0.05$; posttreatment $r = 0.11$, $P = 0.21$).

CPAP Adherence

The distribution of mean CPAP hours per night is summarized in Table 1. Mean (SD) CPAP use was 4.7 (2.1) hours, with a range from 0 to 8.1 hours. Median use was 5.3 hours. As shown in Table 1, there were similar patterns of CPAP use in the subsamples for each outcome variable, with the MSLT subgroup (value < 7.5 minutes) having the lowest proportion of participants using CPAP for each adherence threshold. The Pearson correlation between mean duration of nightly CPAP use and proportion of nights when CPAP was used was $r = 0.89$ ($P < 0.0001$).

In this study, there were participants who did not have excessive sleepiness even at the time of initial assessment. That this

occurs is well known and has led to the concept of the sleep apnea syndrome.¹⁶ The CPAP use in those whose values for each of our 3 outcome variables were normal or abnormal at baseline is shown in Table 2. There was a statistically significant difference in CPAP use ($P = 0.009$) comparing those with baseline values on the MSLT of less than 7.5 minutes and those with values of 7.5 minutes or greater. This was not the case for the comparison of those with normal or abnormal values on the ESS ($P = 0.44$) or the FOSQ ($P = 0.14$).

CPAP Dose Response

To examine the relationship between hours of CPAP use and response to therapy for each outcome, we restricted attention to the

Table 2—hours of CPAP Use Per Night for Participants with Values Above or Below the Threshold Defining Normalcy at Baseline for Each of the 3 Outcome Variables

Variable at baseline	CPAP Use, h/night		
	Mean (SD)	Range	P Value
ESS, score			0.44
> 10 (n = 106)	4.7 (2.2)	0 - 8.1	
≤ 10 (n = 31)	5.1 (1.7)	0 - 7.7	
MSLT, min			0.009
< 7.5 (n = 85)	4.3 (2.2)	0 - 7.7	
≥ 7.5 (n = 51)	5.3 (1.9)	0.5 - 8.1	
FOSQ, Total score			0.14
< 17.9 (n = 120)	4.6 (2.2)	0 - 8.1	
≥ 17.9 (n = 27)	5.3 (1.8)	1.1 - 8.1	

CPAP refers to continuous positive airway pressure; ESS, Epworth Sleepiness Scale; MSLT, Multiple Sleep Latency Test; FOSQ, Functional Outcomes of Sleep Questionnaire.

subset of participants who had abnormal values for that particular outcome at baseline. This avoids the issue of “ceiling effect” obscuring this relationship by including individuals who had normal values at baseline. In these subsets, mean baseline values did not significantly differ between these individuals where the value of the outcome became normal after therapy, i.e., those in whom this did not occur—responders and nonresponders. This was so for all 3 outcome variables (ESS score: responders = 16.6 (3.1), nonresponders = 17.0 (3.5), $P = 0.60$; MSLT: 4.2 (1.8), 3.5 (1.9), respectively, $P = 0.13$; FOSQ: 14.1 (2.5), 13.3 (2.6), respectively, $P = 0.08$).

Subjective Daytime Sleepiness

Among 137 participants with ESS assessments before and after treatment, 106 (77.4%) had a pretreatment ESS Score greater than 10, i.e., were excessively sleepy by self-report. Of these 106 participants, 70 (66%) had ESS score of 10 or less after treatment, i.e., were normal based on current criteria.¹⁰ The mean (SD) hours of CPAP use among those participants whose posttreatment ESS score was 10 or less was 5.1 (2.1) hours per night, whereas the mean hours (SD) of CPAP use among those who remained excessively sleepy was 4.0 (2.3) hours per night ($P = 0.02$). Thus, among participants who reported excessive sleepiness at baseline, as measured by the ESS, those whose sleepiness resolved after treatment used CPAP, on average, about 1 hour more per night than those whose subjective sleepiness did not resolve.

The likelihood of response (ESS score ≤ 10) was not dependent on how close to the threshold (i.e., > 10) participants were prior to treatment. Among the 106 participants with pretreatment ESS scores above 10, pretreatment mean ESS score did not significantly differ ($P = 0.60$) between the 70 participants whose subjective sleepiness resolved (mean = 16.6, SD = 3.1) and the 36 participants whose subjective sleepiness did not resolve (mean = 17.0, SD = 3.5).

Table 3 (row 1) summarizes the percentages of participants among those who were excessively sleepy at baseline and who subsequently had a normal ESS score (≤ 10) after treatment in different categories of hours of CPAP use. The percentage of participants who normalized their self-report sleepiness on therapy rose with increased CPAP use, being 41.2% in those using CPAP on average of 0 to 2 hours per night and 92.9% in those using CPAP on average of more than 7 hours per night.

We performed piecewise regression analyses with a single knot in order to determine if there was a specific threshold CPAP use value, above which further increases in CPAP use were not expected to result in further substantial improvements. These analyses complement the “responder” analyses summarized in Table 3

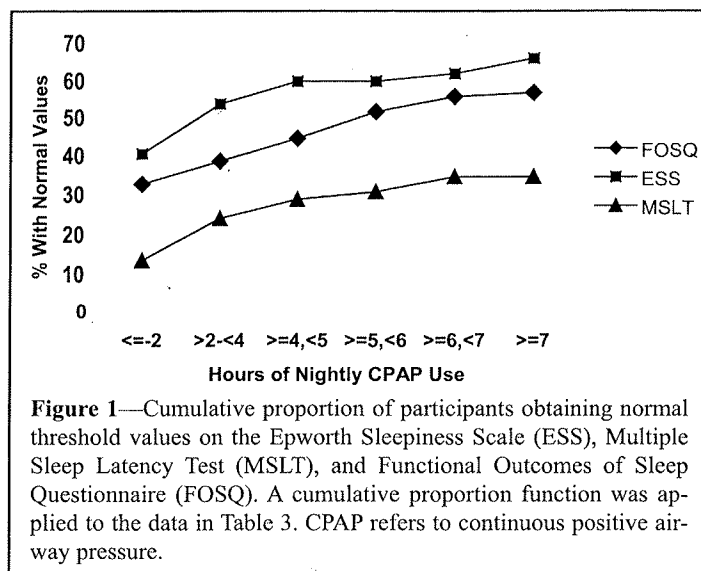


Figure 1—Cumulative proportion of participants obtaining normal threshold values on the Epworth Sleepiness Scale (ESS), Multiple Sleep Latency Test (MSLT), and Functional Outcomes of Sleep Questionnaire (FOSQ). A cumulative proportion function was applied to the data in Table 3. CPAP refers to continuous positive airway pressure.

and Figure 1. Among participants with impairments, the optimal knot for the ESS score was located at 4.0 hours. However, the null hypothesis of linear dose response could not be rejected [$F_{3,102} = 0.18$, $P = 0.91$], indicating greater benefit with increased nightly duration. The estimated regression equation for the posttreatment ESS linear dose response was 13.41 (SE = 0.49) - 0.90 (SE = 0.22) \times mean hours of use. The slope was statistically significant ($P < 0.0001$). The root mean square error and R^2 were 5.01 and 0.14, respectively. The probability of having a posttreatment ESS score of 10 or less was 0.42 even with no CPAP use (the response rate with 0 dose is sometimes referred to as the *natural response rate* in probit analyses). However, even in those with an average of 8 hours of use of CPAP per night, about 20% of these participants remained excessively sleepy by self-report.

Objective Daytime Sleepiness

Similar results were obtained for the assessment of objective sleepiness using the MSLT, albeit with some important differences. Among 136 participants with MSLT assessments before and after treatment, 85 (62.5%) had excessive daytime sleepiness as determined by an MSLT value of less than 7.5 minutes. Of these 85 participants, 30 (35.3%) had values of 7.5 minutes or longer following treatment, which was a much smaller percentage of participants “normalized” in terms of the MSLT, compared with the percentage for normalized with respect to ESS. The mean (SD) hours of CPAP use among those whose posttreatment MSLT values were at least 7.5 minutes was 5.1 (1.7) hours per night,

Table 3—Percentage of Patients with Abnormal Pretreatment Values Who Achieved Normal Values After 3 Months of Treatment

Measure	Mean CPAP Hours Per Night					
	≤ 2	$>2 - <4$	$\geq 4 - <5$	$\geq 5 - <6$	$\geq 6 - <7$	≥ 7
ESS, score	41.2 (7/17)	68.8 (11/16)	73.3 (11/15)	59.1 (13/22)	68.2 (15/22)	92.9 (13/14)
MSLT, min	12.5 (2/16)	35.3 (6/17)	41.7 (5/12)	35.3 (6/17)	53.3 (8/15)	37.5 (3/8)
FOSQ, Total score	33.3 (6/18)	43.5 (10/23)	58.8 (10/17)	68.0 (17/25)	72.7 (16/22)	60.0 (9/15)

Values are presented as percentages with parenthetical numbers describing the number of patients achieving normal values with treatment/total number of patients with abnormal values before treatment. CPAP refers to continuous positive airway pressure; ESS, Epworth Sleepiness Scale; MSLT, Multiple Sleep Latency Test; FOSQ, Functional Outcomes of Sleep Questionnaire.

compared with 3.9 (2.4) hours per night in those whose post-treatment MSLT values remained less than 7.5 minutes (*t*-test for groups with unequal variance $P = 0.01$). Thus, among participants presenting with objectively determined excessive sleepiness, participants whose sleepiness “resolved” (i.e., had an MSLT ≥ 7.5 minutes after CPAP treatment), used CPAP, on average, about 1.2 hours longer per night than those whose objective sleepiness did not “resolve” (i.e., had posttreatment MSLT values of < 7.5 minutes).

Again, the pretreatment “closeness” to the threshold of 7.5 minutes did not differ significantly between those whose objective sleepiness resolved or did not resolve with CPAP treatment. The mean (SD) pretreatment MSLT values were 4.1 (1.8) minutes in the group whose posttreatment MSLT was 7.5 minutes or longer ($n = 30$) and 3.5 (1.9) minutes for the MSLT in those whose posttreatment MSLT remained less than 7.5 minutes ($n = 55$) ($P = 0.13$).

Table 3 summarizes the percentages of participants whose post-treatment MSLT went to longer than 7.5 minutes with different amounts of CPAP use when it had been 7.5 minutes or less before treatment. Among 16 participants with less than 2 hours use per night, only 2 (12.5%) achieved an MSLT value of longer than 7.5 minutes after treatment. In contrast, among 40 participants with 5 or more hours of use per night, 17 (43%) achieved values of 7.5 minutes or longer. The optimal knot for the MSLT was located at 6.0 hours. This is also reflected in the “responder” analysis illustrated in Figure 1. However, the null hypothesis of linear dose response could not be rejected ($F_{3,81} = 0.78$, $P = 0.51$). The expected posttreatment MSLT linear dose response was estimated to be 3.75 (SE = 1.08) + 0.64 (SE = 0.22) \times mean hours of use. The slope was statistically significant ($P = 0.005$). The root mean square error and R^2 were 4.49 and 0.09, respectively. The natural response rate, i.e., the proportion of individuals whose posttreatment MSLT rose to longer than 7.5 minutes with no CPAP use, was 0.14.

Functional Status

The final outcome we assessed was functional status as measured by the FOSQ.¹² The relationship between hours of CPAP and improvement in this outcome was similar to the overall characteristics of response observed with the other self-report measure (the ESS). Among 147 participants with FOSQ assessments before and after treatment, 120 (81.6%) had sleep-related disease-specific functional impairment, as determined by a FOSQ Total score less than 17.9. Of these 120 participants, 68 (56.7%) had FOSQ Total scores of 17.9 or greater following treatment. Responders used their CPAP almost 1 hour longer (5.1 [1.9] hours) than nonresponders (4.1 [2.3] hours) ($P = 0.01$).

The mean (SD) pretreatment value of the FOSQ Total score for those whose values for the score was 17.9 or greater after treatment, i.e., responders, was 14.1 (2.6), compared with those who did not normalize (mean [SD] = 13.3 [2.6]). This difference was not statistically significant ($P = 0.08$).

Table 3, row 3, summarizes the percentages of participants who went from baseline functional impairment to a normal FOSQ Total score with different hours of CPAP use. Among 18 participants with less than 2 hours of use per night, 6 (33.3%) achieved a normal score of 17.9 or greater after treatment. In contrast, among 62 participants with at least 5 hours use per night, 42 (67.7%)

achieved a score of 17.9 or above. For those whose functional status was impaired prior to treatment, the optimal knot was located at 7.5 hours of mean CPAP use. However, given the small number of participants with mean CPAP use above 7.5 hours ($n = 8$) and since the plots of predicted values for models employing knots of 7 or 7.5 hours were very similar, we interpreted the 7-hour solution. Using 7 hours, the null hypothesis of linear dose response was rejected [$F_{3,116} = 4.04$, $P = 0.009$]. Therefore, the linear dose-response relationships differed above and below 7.0 hours. Below 7 hours, expected posttreatment FOSQ Total score value linear dose response was estimated to be 17.5 (SE = 0.25) + 0.33 (SE = 0.12) \times mean hours of use. The slope was statistically significant ($P = 0.008$). In contrast, above 7 hours, the significance of the slope was $P = 0.06$, and its parameter estimate was negative, suggesting no further functional gains after 7 hours of mean use. Thus, because $17.5 + 0.33 \times 7 = 19.8$, this value is the predicted value for more than 7 hours of use. The overall segmented dose-response curve had a root mean square of 2.56 and an R^2 equal to 0.077. Similar to the findings regarding subjective sleepiness, the estimated probability of response with no CPAP use was relatively large (approximately 0.33).

We explored whether initial severity of sleep apnea, as measured by the AHI, or change in weight were factors in determining the relationship between hours of use and the probability of normalizing each of the 3 outcomes we assessed, especially in those with little or no CPAP use. We did not find evidence that severity of disease played a role in determining this relationship (data not shown). There were too few participants who experienced weight loss to determine the relationship between hours of use and achieving a normal score for each of the 3 outcomes. However, of those who used CPAP for 2 hours or less per night and achieved normal values, only 1 of 7 participants with normal ESS scores reported weight loss, none with normal MSLT values, and 2 of 6 with normal FOSQ Total scores. We also considered other putative variables, baseline value, age, baseline body mass index, sex, and education that may influence the estimation of the dose response slope. Compared with the original model, the addition of these variables produced only a 15% decrease in the point estimate in the probit slope coefficient for normalization of the ESS (point estimate 0.1365 - 0.1161), which is marginal with respect to reflecting clinically significant confounding. Adding these variables to the model for the MSLT produced little movement in the point estimate (0.1619 - 0.1652). This demonstrated that these variables were not confounding factors when estimating the dose response of normalization of the MSLT to varying CPAP use. Finally, for the Total score of the FOSQ, the point estimate was increased from 0.1366 to 0.1891 when the selected variables were added, suggesting that the original model point estimate may have slightly underestimated the dose-response slope.

DISCUSSION

In this multisite effectiveness study conducted as part of routine clinical care, we examined the relationship between hours of CPAP use and effectiveness of therapy relative to 3 widely used measures. We identified thresholds above which further improvements in sleepiness and function were less likely relative to the nightly duration of CPAP use and the presence of a dose-response relationship. Our findings indicate that a greater percentage of participants impaired prior to the initiation of therapy will

achieve decreased objective and subjective daytime sleepiness and enhanced functioning to normal levels with longer nightly CPAP durations. However, what constitutes adequate use varies between outcomes, and, although continuous gains are achieved for objective and subjective daytime sleepiness with greater use, it appears no further benefit in functional status is realized with use beyond 7 hours. Improvements in these outcomes following no CPAP use (14% - 30% of the sample) in those impaired before treatment were more likely when outcomes were measured with subjective rather than objective measures.

Our results that better outcomes occur, for the most part, with more hours of nightly CPAP use are consistent with those from controlled trials.^{20,21} Stradling and Davies reported that CPAP use of more than 5 hours per night resulted in scores within the normal range for the ESS, whereas this linear relationship did not exist for those participants treated with subtherapeutic therapy (0.5 - 1.0 cm H₂O).²¹ Moreover, in a retrospective study, Campos-Rodrigues and colleagues showed significant differences in 5-year survival rates between those with mean CPAP use of less than 1 hour per day and those using CPAP 1 to 6 and more than 6 hours per day.²² Increased use was associated with longer survival. In a recent report, Zimmerman and associates found that memory-impaired participants with sleep apnea were 8 times more likely to test within the normal range with an average of 6 hours per night of CPAP use, compared with those who used it 2 hours or less.²³ Also congruent with our results are experiments in healthy adults showing robust differences in daytime sleepiness between individuals with 4 and 6 hours and 6 and 8 hours of sleep duration, suggesting that the more time in bed up to 8 hours results in decreased daytime sleepiness.^{24,25}

The mean nightly duration of CPAP use in our study is typical of previous studies of CPAP use.²⁶ For each of our 3 outcomes, those who normalized on therapy used CPAP on average 1.1 (95% confidence interval [CI] 0.2 - 2.0), 1.1 (95% CI 0.2 - 2.1), and 1.0 (95% CI 0.2 - 1.8) hours per night more than those who did not for the ESS, MSLT, and FOSQ, respectively. All of these differences were significant. Group difference standardized effect sizes, or magnitude of change, for mean CPAP use comparing responders to nonresponders were 0.49, 0.54, and 0.47, respectively, for these 3 outcomes, indicating moderate and clinically meaningful change.

However, it's noteworthy that there was a substantial minority of participants who were effectively treated with shorter durations of CPAP use. The proportion of participants who benefited with less than 2 hours use of CPAP was 41% and 33% for the ESS and FOSQ respectively, compared with 12.5% for the MSLT. Moreover, even among those using CPAP for more than 7 hours per night, there was a significant proportion that remained excessively sleepy despite what appears to be optimal therapy, i.e., that had residual sleepiness. The bases for such incongruities between CPAP use and achieving normal levels of sleepiness and function at both ends of the spectrum of adherence remain to be explained. We explored several possible explanations for this phenomenon. Foremost, was the proximity of the baseline pretreatment value to the selected threshold defining impairment. If the pretreatment measure was close to the cutpoint defining normal, a relatively minor change could result in the posttreatment value being normal. However, we did not find this to be a factor in determining who responded or did not respond to CPAP therapy. Pretreatment values for all our outcomes were similar, and not significantly dif-

ferent, between those who normalized their value for each outcome variable, as compared with those who did not. It is likely that factors that can influence sleepiness, but that are unrelated to sleep apnea severity, such as caffeine intake, may play a role in variability in sleepiness-related outcomes.

Even in the group that used CPAP for a mean nightly duration of 7 or more hours, the percentage that achieved normal values was less than 100% for all outcomes. Why residual sleepiness, both subjective and objective, remains despite ideal CPAP use is unclear. Our study provides an estimate of the prevalence of residual sleepiness in such optimally treated participants. As displayed in Table 3, 8 of 36 participants (22.2%) who displayed subjective pretreatment daytime sleepiness and 12 of 23 (52.2%) who manifested objective pretreatment daytime sleepiness still had sleepiness after using CPAP for 6 or more hours per night. That is, a fifth to a half of the participants we studied had evidence of residual sleepiness, depending on which outcome was assessed, despite what is reasonably regarded as optimal therapeutic use of CPAP. The presence of residual daytime sleepiness after CPAP treatment has been reported previously.^{2,27-29}

Possible reasons for the observed residual daytime sleepiness that have been proposed have included decreased sleep duration,²⁷ the number of nights CPAP was used,^{30,31} the presence of arousals,^{27,31,32} residual respiratory events^{27,31} that are attributable to inadequate pressure³¹ and permanent alteration of sleep-promoting mechanisms or permanent change in the endogenous waking drive.³¹ The presence of excessive daytime sleepiness in objectively monitored CPAP use of more than 7 hours' duration makes the explanation of inadequate sleep duration or CPAP use unlikely, although we recognize that CPAP adherence is not necessarily a reliable measure of the duration of continuity of actual sleep obtained by patients.³¹ Lamphere and colleagues found that there was significant improvement in excessive daytime sleepiness comparing 1 and 14 days of CPAP use but no further improvement after 42 days of use.³⁰ Because our follow-up evaluation of the outcomes following initiation of CPAP therapy was conducted after at least 90 days of treatment, it is also unlikely that inadequate duration of treatment is a factor in explaining the observed residual sleepiness. The contribution of arousals and residual respiratory events cannot, however, be excluded because we did not perform an overnight sleep study at follow-up due to protocol burden. We note that, in the initial titration, CPAP was adjusted to remove not only respiratory events, but also snoring and snoring-related arousals. Residual sleepiness might also be related to the effects of obesity itself. As a group, the participants in this study were quite obese with a mean body mass index of 38.1 kg/m². Obesity, even in the absence of sleep apnea, is associated with excessive sleepiness,^{33,34} a phenomenon that has been attributed to visceral adipocyte production of cytokines, which can be sleep promoting.^{34,35}

Another possible explanation for residual sleepiness comes from studies in which mice are exposed to chronic intermittent hypoxia simulating the hypoxia associated with severe sleep apnea. Mice exposed over a period of weeks to intermittent hypoxia for 10 hours per day during the lights on (inactive) period demonstrate increased sleep amounts and shortened sleep latency even when tested 2 weeks after termination of the intermittent hypoxic exposure.³⁶ This implies permanent effects and evidence of oxidative damage with, for example, carbonylation, which has been shown in a number of wake-active neuronal groups.³⁶⁻³⁸ Thus,

irreversible injury of neurons that promote wakefulness that results from the chronic intermittent hypoxemia in sleep apnea is a possible explanation for residual sleepiness in treated obstructive sleep apnea.

Although there are similarities between the results for subjective and/or objective measures of sleepiness, there are also important differences. As discussed, the probability of a positive response for low use of CPAP is much lower for the MSLT than for the 2 subjective measures. Moreover, there is a much higher proportion of patients who do not normalize their MSLT even with maximal CPAP use than for the ESS and FOSQ. Differences between effects of CPAP on objective and subjective sleepiness have been shown previously.³⁹ Metaanalysis of the results of CPAP treatment trials have shown that CPAP has much larger effects on self-report than on objective sleepiness.³⁹ The basis of the different effects of CPAP use on subjective and objective measures of sleepiness may lie in the role of demand characteristics. Demand characteristics are pervasive in assessment and refer to all the study factors that influence participants' beliefs about what constitutes being a "good" subject, which is usually tied to their perception of pleasing the investigator or not "failing" the study. Because subjective reports (no matter how authoritative the scale) are especially prone to the effects of demand characteristics, it is possible that the association between limited CPAP use and subjective sleepiness was a reflection of participants trying to please the investigator, or at least to avoiding the appearance of "failing" the study. This is much easier to do when rating oneself on a scale or answering questions than when undergoing objective testing. Hence, demand characteristics are a likely parsimonious explanation for why subjective ratings change in clinical studies, even more so than do objective measures of sleepiness.

Certain aspects of our study design are worthy of comment. This study was designed as an effectiveness study to investigate what normally occurs in routine clinical practice. Thus, the study, by design, did not contain a control group. The design was based on the belief, which turned out to be correct, that there would be sufficient variability in hours of CPAP use in this large sample of patients that we could examine the relationship between hours of CPAP use and treatment effectiveness, when judged using end-points measuring different aspects of excessive sleepiness. We performed responder analyses in which we categorized individuals as impaired or not impaired, i.e., dichotomous outcomes, as well as assessed outcomes as continuous variables in piecewise regression analyses. Both the piecewise regression and probit analyses were limited to individuals who were abnormal at baseline, allowing these approaches to provide complementary information among an identical sample of subjects. Impairment and nonimpairment were based on specific thresholds for each of our outcome measures. For the ESS (normal = ≤ 10), there are normative data in the literature.¹⁰ The MSLT is more complex, since there are limited normative data in the literature, and, by convention, pathologic sleepiness is less than 5 minutes, normal is longer than 10 minutes, and a gray zone is between 5 and 10 minutes.^{10,11} We used 7.5 minutes, i.e., a point midway in the gray zone area. Although our specific results are obviously dependent on the threshold of normalcy used, they are not particularly sensitive to this. To show this, we reran our analysis for each outcome with different values of the threshold defining normal, i.e., ESS (8 and 12), MSLT (5 and 10 minutes), and FOSQ Total score (17.8 and 17.0). Although some of these additional analyses were lim-

ited by the number of participants, we found that all of the major results that we report were not particularly sensitive to the choice of threshold. Specifically, we found the following for all analyses (data now shown, but available on request): a relationship between probability of normalization and hours of use of CPAP for all 3 outcome measures; a higher proportion of participants who normalize with limited CPAP use for subjective measures, as compared to the MSLT; a proportion of participants who do not normalize for each outcome measure even when average CPAP use is 7 hours or more per day; the proportion who do not normalize with this high CPAP use is higher for the MSLT than for the other self-report measures.

Another limitation of our study was the lack of objective data regarding sleep duration prior to testing, as this may have affected the treatment response. However, we were more interested in the results of chronic rather than acute CPAP use. Considering total sleep time the night prior to the assessment of outcomes would reflect acute use, limiting the examination of the relationship between pattern of adherence over time and salient outcomes.

In conclusion, this study shows a clear relationship between effectiveness of CPAP therapy and hours of use of CPAP in routine clinical practice, adding to the growing evidence that increased nightly use leads to better clinical outcomes. There are important clinical messages from this study. First, from a population sense, functions for predicted probabilities of normalization show that more CPAP use is associated with greater relief of sleepiness (no matter how it is measured). However, there are patients who normalize on therapy with somewhat limited CPAP use. The actual need for CPAP in terms of reversing sleepiness is likely to be individually determined. We cannot assume that an individual using CPAP only 4 hours per night is inadequately treated for sleepiness outcomes; we also cannot assume that the patient is effectively treated. Thus, it is important to evaluate treatment effectiveness by assessing level of adherence in conjunction with treatment outcomes. There are, in contrast, individuals who remain excessively sleepy despite more than 7 hours of CPAP use per night. Moreover, the determination of a recommended nightly duration of CPAP use is also dependent on which outcome is viewed as the most reflective of clinical improvement.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the contribution to data collection, management, and analysis of the following individuals: site research coordinators: Sandra Halko (University of Western Ontario), Bernadette Krug (Bons Secours Holy Family Regional Health System), Dr. Angie Randazzo (St. Luke's Hospital), Deborah Sturgeon (Sleep Medicine Center of Kansas), Constance Ullevig (Hennepin County Medical Center); and the research team at the University of Pennsylvania School of Nursing: Brian Honbo, Wendy Black Dorn, Susan McCloskey, and Sharon Hurley.

REFERENCES

1. Gay P, Weaver T, Loube D, Iber C. Evaluation of positive airway pressure treatment for sleep-related breathing disorders in adults. *Sleep* 2006;29:381-401.
2. Kribbs NB, Pack AI, Kline LR, et al. Effects of one night without nasal CPAP treatment on sleep and sleepiness in patients with obstructive sleep apnea. *Am Rev Respir Dis* 1993;147:1162-8.

3. Grunstein RR, Stewart DA, Lloyd H, Akinci M, Cheng N, Sullivan CE. Acute withdrawal of nasal CPAP in obstructive sleep apnea does not cause a rise in stress hormones. *Sleep* 1996;19:774-82.
4. Fiz JA, Abad J, Ruiz J, Riera M, Izquierdo J, Morera J. nCPAP treatment interruption in OSA patients. *Respir Med* 1998;92:28-31.
5. Weaver TE, Kribbs NB, Pack AI, et al. Night-to-night variability in CPAP use over the first three months of treatment. *Sleep* 1997;20:278-83.
6. Kribbs NB, Pack AI, Kline LR, et al. Objective measurement of patterns of nasal CPAP use by patients with obstructive sleep apnea. *Am Rev Respir Dis* 1993;147:887-95.
7. Sin DD, Mayers I, Man GC, Ghahary A, Pawluk L. Can continuous positive airway pressure therapy improve the general health status of patients with obstructive sleep apnea?: a clinical effectiveness study. *Chest* 2002;122:1679-85.
8. Rosenthal L, Gerhardstein R, Lumley A, et al. CPAP therapy in patients with mild OSA: implementation and treatment outcome. *Sleep Med* 2000;1:215-20.
9. Van Dongen HP, Maislin G, Dinges DF. Dealing with inter-individual differences in the temporal dynamics of fatigue and performance: importance and techniques. *Aviat Space Environ Med* 2004;75(3 Suppl):A147-54.
10. Johns MW. Sensitivity and specificity of the multiple sleep latency test (MSLT), the maintenance of wakefulness test and the Epworth sleepiness scale: failure of the MSLT as a gold standard. *J Sleep Res* 2000;9:5-11.
11. Carskadon MA, Dement WC, Mitler MM, Roth T, Westbrook PR, Keenan S. Guidelines for the multiple sleep latency test (MSLT): a standard measure of sleepiness. *Sleep* 1986;9:519-24.
12. Weaver TE, Laizner AM, Evans LK, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep* 1997;20:835-43.
13. Practice parameters for the indications for polysomnography and related procedures. Polysomnography Task Force, American Sleep Disorders Association Standards of Practice Committee. *Sleep* 1997;20:406-22.
14. Rechtschaffen A, Kales A. A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects UCLA, Los Angeles: Brain Information Service/Brain Research Institute; 1968.
15. Ayappa I, Norman RG, Krieger AC, Rosen A, O'Malley R L, Rapoport DM. Non-invasive detection of respiratory effort-related arousals (Reras) by a nasal cannula/pressure transducer system. *Sleep* 2000;23:763-71.
16. Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. The Report of an American Academy of Sleep Medicine Task Force. *Sleep* 1999;22:667-89.
17. Carskadon MA, Dement WC. The multiple sleep latency test: what does it measure? *Sleep* 1982;5(Suppl 2):S67-72.
18. Vieth E. Fitting piecewise linear regression functions to biological responses. *J. Appl. Physiol* 1989;67:390-6.
19. Finney M. Probit analysis: A statistical treatment of the sigmoid response curve Cambridge: University Press; 1947.
20. 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.
21. Stradling JR, Davies RJ. Is more NCPAP better? *Sleep* 2000;23 Suppl 4:S150-3.
22. Campos-Rodriguez F, Pena-Grinan N, Reyes-Nunez N, et al. Mortality in obstructive sleep apnea-hypopnea patients treated with positive airway pressure. *Chest* 2005;128:624-33.
23. Zimmerman ME, Arnedt JT, Stanchina M, Millman RP, Aloia MS. Normalization of memory performance and positive airway pressure adherence in memory-impaired patients with obstructive sleep apnea. *Chest* 2006;130:1772-8.
24. Van Dongen HP, Maislin G, Mullington JM, Dinges DF. The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep* 2003;26:117-26.
25. Belenky G, Wesensten NJ, Thorne DR, et al. Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: a sleep dose-response study. *J Sleep Res* 2003;12:1-12.
26. Engleman HM, Wild MR. Improving CPAP use by patients with the sleep apnoea/hypopnoea syndrome (SAHS). *Sleep Med Rev* 2003;7:81-99.
27. Pack AI, Black JE, Schwartz JR, Matheson JK. Modafinil as adjunct therapy for daytime sleepiness in obstructive sleep apnea. *Am J Respir Crit Care Med* 2001;164:1675-81.
28. Bedard MA, Montplaisir J, Malo J, Richer F, Rouleau I. Persistent neuropsychological deficits and vigilance impairment in sleep apnea syndrome after treatment with continuous positive airways pressure (CPAP). *J Clin Exp Neuropsychol* 1993;15:330-41.
29. Sforza E, Krieger J. Daytime sleepiness after long-term continuous positive airway pressure (CPAP) treatment in obstructive sleep apnea syndrome. *J Neurol Sci* 1992;110:21-6.
30. Lamphere J, Roehrs T, Wittig R, Zorick F, Conway WA, Roth T. Recovery of alertness after CPAP in apnea. *Chest* 1989;96:1364-7.
31. Dinges DF, Weaver TE. Effects of modafinil on sustained attention performance and quality of life in OSA patients with residual sleepiness while being treated with nCPAP. *Sleep Med* 2003;4:393-402.
32. Chugh DK, Weaver TE, Dinges DF. Neurobehavioral consequences of arousals. *Sleep* 1996;19(10 Suppl):S198-201.
33. Vgontzas AN, Bixler EO, Tan TL, Kantner D, Martin LF, Kales A. Obesity without sleep apnea is associated with daytime sleepiness. *Arch Intern Med* 1998;158:1333-7.
34. Vgontzas AN, Papanicolaou DA, Bixler EO, et al. Sleep apnea and daytime sleepiness and fatigue: relation to visceral obesity, insulin resistance, and hypercytokinemia. *J Clin Endocrinol Metab* 2000;85:1151-8.
35. Vgontzas AN, Papanicolaou DA, Bixler EO, Kales A, Tyson K, Chrousos GP. Elevation of plasma cytokines in disorders of excessive daytime sleepiness: role of sleep disturbance and obesity. *J Clin Endocrinol Metab* 1997;82:1313-6.
36. Veasey SC, Davis CW, Fenik P, et al. Long-term intermittent hypoxia in mice: protracted hypersomnolence with oxidative injury to sleep-wake brain regions. *Sleep* 2004;27:194-201.
37. Macey PM, Henderson LA, Macey KE, et al. Brain morphology associated with obstructive sleep apnea. *Am J Respir Crit Care Med* 2002;166:1382-7.
38. Macey PM, Macey KE, Henderson LA, et al. Functional magnetic resonance imaging responses to expiratory loading in obstructive sleep apnea. *Respir Physiol Neurobiol* 2003;138:275-90.
39. Patel SR, White DP, Malhotra A, Stanchina ML, Ayas NT. Continuous positive airway pressure therapy for treating sleepiness in a diverse population with obstructive sleep apnea: results of a meta-analysis. *Arch Intern Med* 2003;163:565-71.

