The Association between Obstructive Sleep Apnea and Neurocognitive Performance—The Apnea Positive Pressure Long-term Efficacy Study (APPLES)

Stuart F. Quan, MD[1,2]; Cynthia S. Chan, BS[3]; William C. Dement, MD, PhD[4]; Alan Gevins, DSc[3]; James L. Goodwin, PhD[1]; Daniel J. Gottlieb, MD, MPH[5]; Sylvan Green, MD[10]; Christian Guilleminault, MD[4]; Max Hirshkowitz, PhD[6]; Pamela R. Hyde, MA[4]; Gary G. Kay, PhD[7]; Eileen B. Leary, RPSGT[4]; Deborah A. Nichols, MS[4]; Paula K. Schweitzer, PhD[8]; Richard D. Simon, MD[9]; James K. Walsh, PhD[8]; Clete A. Kushida, MD, PhD[4]

[1]Arizona Respiratory Center, University of Arizona, Tucson, AZ; [2]Division of Sleep Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA; [3]The San Francisco Brain Research Institute and SAM Technology, San Francisco, CA; [4]Stanford University, Stanford, CA; [5]Boston University School of Medicine, Boston, MA; [6]Baylor College of Medicine and VA Medical Center, Houston, TX; [7]Cogscreen, LLC, Washington, DC; [8]St. Luke's Hospital, Chesterfield, MO; [9]St. Mary Medical Center, Walla Walla, WA; [10]College of Public Health, University of Arizona, Tucson, AZ

Study Objectives: To determine associations between obstructive sleep apnea (OSA) and neurocognitive performance in a large cohort of adults.
Study Design: Cross-sectional analyses of polysomnographic and neurocognitive data from 1204 adult participants with a clinical diagnosis of obstructive sleep apnea (OSA) in the Apnea Positive Pressure Long-term Efficacy Study (APPLES), assessed at baseline before randomization to either continuous positive airway pressure (CPAP) or sham CPAP.
Measurements: Sleep and respiratory indices obtained by laboratory polysomnography and several measures of neurocognitive performance.
Results: Weak correlations were found for both the apnea hypopnea index (AHI) and several indices of oxygen desaturation and neurocognitive performance in unadjusted analyses. After adjustment for level of education, ethnicity, and gender, there was no association between the AHI and neurocognitive performance. However, severity of oxygen desaturation was weakly associated with worse neurocognitive performance on some measures of intelligence, attention, and processing speed.
Conclusions: The impact of OSA on neurocognitive performance is small for many individuals with this condition and is most related to the severity of hypoxemia.
Keywords: Obstructive sleep apnea, neurocognition, epidemiology, sleep, oxygen desaturation
Citation: Quan SF; Chan CS; Dement WC; Gevins A; Goodwin JL; Gottlieb DJ; Green S; Guilleminault C; Hirshkowitz M; Hype PR; Kay GG; Leary EB; Nichols DA; Schweitzer PK; Simon RD; Walsh JK; Kushida CA. The association between obstructive sleep apnea and neurocognitive performance—the Apnea Positive Pressure Long-term Efficacy Study (APPLES). *SLEEP* 2011;34(3):303-314.

INTRODUCTION

The obstructive sleep apnea syndrome (OSA) is a potentially life-threatening sleep related breathing disorder estimated in 1993 to affect 4% of men and 2% of women in the United States between the ages of 30 and 60 years.[1] The prevalence appears to be increasing with trends towards higher rates of obesity.[2] Recent data from several large observational cohort studies provide strong evidence that untreated OSA is an independent risk factor for incident hypertension, cardiovascular disease and stroke, as well as higher all-cause mortality.[3-8] Somewhat less attention has been paid to the relationship between OSA and neurocognitive abilities, especially in the domains of attention, memory and executive function.

A number of studies primarily in clinical populations have examined the relationship between OSA and neurocognitive function.[9,10] However, results are inconsistent. Various studies have observed that OSA is associated with impaired cognition.[11-16] Some report executive and psychomotor deficits, attention and memory problems, and impaired vigilance,[14] and motor and perceptual-organization ability deficits.[12] The mechanism responsible for these impairments also may differ according to the neurocognitive domain tested. In some studies, memory deficits are associated with the apnea hypopnea index, whereas frontal lobe-related dysfunction correlated best with the severity of OSA associated hypoxemia.[15,16] In addition, neurocognitive abilities in persons with OSA may be variably affected for a variety of reasons including differential sensitivity to sleep disturbance. A meta-analytic review suggests that OSA has a moderate to severe impact on vigilance, motor coordination, and executive functions, while there is little effect on intelligence, verbal, and visual-perceptual abilities.[17] Whether mild OSA impacts neurocognition also is unclear. Although there are reports of impairment in vigilance and working memory[18] as well as psychomotor function[19] in persons with mild OSA, not all studies have observed similar findings.[20]

Although the preponderance of evidence indicates reduced neurocognitive performance does occur in persons with OSA, there are inconsistencies in the type and degree of impairment, and the underlying mechanism responsible. These discrepancies could be due to the size and type of population studied, sensitivity of the neurocognitive tests employed, and the definitions used to characterize OSA.[21] There have been few large studies performed, and their neurocognitive assessments were limited.[22,23] A large study using a variety of neurocognitive outcomes with enrollment of a study population having OSA

A commentary on this article appears in this issue on page 249.

Submitted for publication May, 2010
Submitted in final revised form July, 2010
Accepted for publication August, 2010
Address correspondence to: Stuart F. Quan, MD, Division of Sleep Medicine, Harvard Medical School, 401 Park Dr., 2nd Floor East, Boston, MA 02215; Tel: (617) 998-8842; Fax: (617) 998-8823; E-mail: squan@arc.arizona.edu

Table 1—Baseline neurocognitive, mood, mental status, and sleepiness data for all participants[4] and stratified by gender

	Total Mean (SD)	Men Mean (SD)	Women Mean (SD)
Intelligence[3]			
WASI Full	112.0 (13.2)	112.8[2] (12.9)	110.7 (13.5)
WASI Verbal	109.8 (13.4)	109.9 (13.4)	109.7 (13.5)
WASI Vocabulary	55.9 (9.7)	56.1 (9.7)	55.7 (9.7)
WASI Similarities	56.0 (7.9)	55.9 (7.9)	56.0 (8.0)
WASI Performance	111.5 (13.3)	112.9[2] (13.2)	109.1 (13.2)
WASI Block Design	54.0 (9.0)	55.5[2] (7.7)	52.3 (9.0)
WASI Matrix Reasoning	59.2 (8.4)	59.6[1] (8.4)	58.5 (8.3)
Primary Neurocognitive Outcomes[3]			
SWMT	0.000 (0.28)	-0.01 (0.23)	0.02 (0.28)
Pathfinder	24.43 (8.50)	24.4 (8.89)	24.5 (5.74)
BSRT	49.98 (9.14)	48.72[2] (9.12)	52.21 (8.74)
Secondary Neurocognitive Outcomes[3]			
PVT Median	251.83 (38.86)	245.54[2] (38.60)	262.92 (36.82)
PVT Mean	465.97 (329.91)	448.52[2] (36.83)	500.31 (360.22)
PASAT	154.97 (42.73)	159.32[2] (42.38)	147.34 (42.32)
Mood[3]			
HAM-D	4.43 (4.13)	3.80[2] (3.72)	5.54 (4.57)
Sleepiness[3]			
ESS	10.15 (4.35)	9.94[1] (4.23)	10.52 (4.52)
SSS	2.91 (0.98)	2.81[2] (0.96)	3.09 (1.00)
MMSE[3]	**N (%)**	**N (%)**	**N (%)**
26	14 (1.2)[5]	9 (1.2)	5 (1.1)[5]
27	79 (6.6)	53 (6.9)	26 (6.0)
28	169 (14.0)	114 (14.8)	55 (12.6)
29	401 (33.3)	263 (34.2)	138 (31.7)
30	540 (44.9)	330 (42.9)	210 (48.3)

[1]P < 0.05 Men vs. Women; [2]P < 0.01 Men vs. Women; [3]See text for abbreviations; [4]N for each variable ranges from 1143 (SWMT) to 1204 (ESS); [5]1 Subject with a MMSE score = 25 included.

severity spanning the spectrum from mild to severe would be informative in addressing some of these inconsistencies.

The Apnea Positive Pressure Long-term Efficacy Study (APPLES) is a randomized, double-blinded, sham-controlled, multi-center trial of continuous positive airway pressure (CPAP) therapy, designed to determine whether CPAP improves neurocognitive function over a 6-month test period. The present study is an analysis of the relationship between a variety of neurocognitive measures at the baseline visit (pre-CPAP) and severity of OSA in 1204 APPLES participants. We hypothesized that impairment on several neurocognitive domains would be associated with severity of OSA and that lower performance would be related to hypoxemia.

MATERIALS AND METHODS

Study Population and Protocol

A detailed description of the APPLES protocol has been previously published.[24] Briefly, at all 5 APPLES sites (Stanford University, Stanford, CA; University of Arizona, Tucson, AZ; Brigham and Women's Hospital, Boston, MA; St. Luke's Hospital, Chesterfield, MO; and St. Mary Medical Center, Walla Walla, WA), subjects were recruited into the study primarily from patients scheduled into a regular sleep clinic for evaluation of possible OSA and from local advertising. Although recruitment source was not tracked, it is estimated that initial contact with ~70% of the subjects occurred as a result of advertisement. Symptoms indicative of OSA were used as pre-screening questions. Initially, a clinical evaluation was conducted which included administering informed consent, screening questionnaires, history and physical examination and a medical assessment by a study physician. Inclusion criteria were a clinical diagnosis of OSA, as defined by American Academy of Sleep Medicine (AASM) criteria,[25] an apnea hypopnea index (AHI) ≥ 10 by polysomnography (PSG), and age ≥ 18 years. Exclusion criteria have been published previously,[24] the most important of which were non-enrollment of those with previous CPAP use, oxygen desaturation on PSG < 75% for > 10% of the recording time, history of a motor vehicle accident related to sleepiness, presence of a number of chronic medical conditions, and use of various medications with the potential to affect sleep or neurocognitive function. Furthermore, potential subjects who scored ≤ 26 on the Mini Mental State Examination (MMSE) were generally excluded, although in some cases, subjects with lower scores were allowed to participate at the discretion of the site investigator (1.2%, Table 1). Subjects subsequently returned approximately 2-4 weeks later for a 24-h sleep laboratory visit, where a diagnostic PSG was performed to confirm the presence of OSA (*vide infra*) followed by a day of neurocognitive and maintenance of wakefulness testing. Approximately 10-14 days later, the subject's AHI was received from the central PSG Scoring Center. Only those subjects with an AHI ≥ 10 events/h were considered to have OSA, and were randomized to continue participation in the APPLES study. However, for the present analysis, data from both randomized and participants not randomized on the basis of PSG exclusion criteria were included, provided that they had a baseline polysomnogram. Only data obtained during this baseline visit was used in the current analysis. Information regarding subsequent APPLES data collection is available in a previous publication.[24]

Polysomnography

The PSG montage included monitoring of the electroencephalogram (EEG, C_3-M_2, C_4-M_1, O_2-M_1 and O_1-M_2), electrooculogram (EOG, ROC-M_1, LOC-M_2), submentalis and anterior tibialis electromyograms (EMG), heart rate by 2-lead electrocardiogram, snoring intensity (anterior neck microphone), nasal pressure (nasal cannula), nasal/oral thermistor, thoracic and abdominal movement (piezo bands), and oxygen saturation (pulse oximetry). All PSGs were electronically transmitted to the Data Coordinating Center. Sleep and wakefulness were scored using Rechtschaffen and Kales criteria.[26] Apneas and hypopneas were scored using AASM (1999) criteria.[27] Briefly, an apnea was defined by a clear decrease (> 90%) from baseline in the amplitude of the nasal pressure signal lasting ≥ 10 sec. Hypopneas were identified if there was a clear decrease (> 50% but ≤ 90%)

from baseline in the amplitude of the nasal pressure signal, or if there was a clear amplitude reduction of the nasal pressure signal that did not reach the above criterion but it was associated with either an oxygen desaturation > 3% or an arousal, and the event duration was ≥ 10 seconds. Obstructive events were identified by persistence of chest or abdominal respiratory effort during flow cessation. Central events were noted if no displacement occurred on either the chest or abdominal channels. The apnea hypopnea index (AHI) was calculated as the sum of all apneic and hypopneic events divided by the hours of total sleep time (TST). The central apnea index, arousal index, periodic limb movement index, and desaturation index (based on 3% desaturation from baseline) were computed similarly.

Neurocognitive Testing

Primary outcomes

All neurocognitive testing was performed in the same predefined order for all subjects on the day after their diagnostic PSG. The testing was preceded by 2 training sessions.[24] Typically, the first occurred several days or weeks before the diagnostic PSG, and the second on the night before the PSG. The neurocognitive testing schedule is provided in Appendix 1. There were 3 prespecified primary neurocognitive outcomes, each representing 1 of 3 different neurocognitive domains: (1) attention and psychomotor function, (2) learning and memory, and (3) executive and frontal-lobe function. The primary neurocognitive outcomes were selected as a result of recommendations made by the study's neurocognitive function consulting team.[24] A brief description of each of these tests follows.

Pathfinder Number Test (Pathfinder): The Pathfinder Test assesses attention and psychomotor function, requiring the subject to scan, locate, and connect numbers in sequence. The Pathfinder Test is the computer analogue of the Trail Making Test Part A and is part of the APPLES edition of the CogScreen test battery designed to assess specific neurocognitive changes associated with OSA following CPAP therapy.[28] The participant is instructed to select consecutive numbers on a computer screen by tapping a light pen to numbers which appear in boxes in the quadrants of the screen. In addition to selecting the next number as quickly as possible, the participant is instructed to tap the center of each box to provide a measure of psychomotor coordination. The variable selected for primary analyses was the total time required to complete the test. Performance on the Trail Making Test Part A has been reported to improve with treatment of OSA.[29] The Cogscreen software is capable of one millisecond timing resolution. The ceiling for this variable is 60.00 seconds.

Buschke Verbal Selective Reminding Test (BSRT): The BSRT is a multiple-trial, list-learning task that is a measure of verbal learning with short-term and long-term memory components.[30] For APPLES, a series of 12 unrelated words were presented over 6 selective reminding trials, or until the subject was able to recall the entire list on three consecutive trials.[31] Lastly, a delayed-recall trial was given with forewarning 30 min after the completion of the test. The BSRT has been used in OSA subjects to indicate a learning deficit in patients with moderate and severe OSA.[16] The sum recall variable, or total number of correctly recalled words over the 6 trials, was the primary variable selected for analysis. The range for this variable is a minimum score of 0 up to a maximum score of 72.

Sustained Working Memory Test (SWMT): The SWMT measures working memory which is thought to be a frontal/executive function. The test combines an N-back working memory task with simultaneous electroencephalographic (EEG) recording. The task requires the subject to sit in front of a computer monitor and maintain fixation on the middle of the computer screen. Every few seconds a simple visual stimulus appears briefly at a random screen location. Subjects are required to compare the spatial position of the initial stimulus with the position of the stimulus that occurred on a previous trial, and to press one button if the spatial position was the same as that on the previous trial, or a second button if it differed. In order to vary task difficulty, the comparison trial varies from the prior trial (1-back) to 2 trials prior (2-back). Behavioral and EEG (both background EEG and event-related potential) parameters are extracted from the resulting data. The behavioral data that are acquired provide evidence about any changes in the overt performance abilities of the subject, whereas the task-related EEG measures provide evidence about how the brain responds to changes in task difficulty. These measures are combined to yield a composite score indicating the degree of change from pretreatment baseline. The test has been shown to be sensitive to changes in alertness induced by sleep deprivation.[32] For APPLES, these data were provided as 3 sub-indices: Behavioral, Activation, and Alertness, that were combined to yield an Overall Index. The Overall Index for the mid-day administration was selected as the primary outcome variable. A positive score indicates improvement from pre-treatment baseline, while a negative score indicates worsening. Due to the fact that the SWMT is formulated as a change from baseline score, there is no baseline variable to analyze for the Overall Index. Instead, a distinct variable measuring diurnal variation between the three baseline administrations was utilized for the purpose of this manuscript.

Secondary outcomes

In addition to the primary neurocognitive outcomes, other measures included in the APPLES test battery were exploratory neurobehavioral measures and sleepiness, fatigue, mood, quality of life, and sleep and health history assessments. These are shown in Appendix 2. Secondary outcome measures included in this analysis were the Wechsler Abbreviated Scale of Intelligence (WASI), Psychomotor Vigilance Test (PVT), Paced Auditory Serial Addition Test (PASAT), and the Epworth Sleepiness Scale (ESS). In addition, the Hamilton Rating Scale for Depression (HAM-D), ESS (when not assessed as an outcome measure), and the Stanford Sleepiness Scale (SSS) were used as covariates because of their potential to affect performance on the neurocognitive outcomes. Although there were a number of secondary outcomes included in the APPLES test battery, measures were selected for this analysis based on administration in previous studies and discussions among members of the study's neurocognitive function consulting team. A brief description of these measures follows.

Wechsler Abbreviated Scale of Intelligence (WASI): The Wechsler Adult Intelligence Scale (WAIS) is a standardized and validated instrument commonly used to assess adult

intelligence. For APPLES, we administered an abbreviated version of the WAIS, the Wechsler Abbreviated Scale of Intelligence (WASI).[33] designed to provide a shorter, yet reliable intelligence measure for research settings. The WASI is standardized and linked to the WAIS-III. For this analysis, we report the Full Scale WASI IQ (WASI-Full), Verbal WASI IQ (WASI-Verbal) and Performance WASI IQ (WASI-Perf) as well as their respective components, Vocabulary WASI (WASI-Vocab), Similarities WASI (WASI-Simil), Block Design WASI (WASI-BD), and the Matrix Reasoning WASI (WASI-MATR). The WASI-Verbal is a measure of verbal comprehension. Its components are the WASI-Vocab (testing comprehension and expression of verbal vocabulary) and WASI-Simil (assessing verbal reasoning). The WASI-Perf is an assessment of perceptual organization. Its components are the WASI-BD (testing spatial perception, visual abstract processing, and problem solving) and the WASI-MATR (assessing nonverbal abstract problem solving, inductive reasoning, and spatial reasoning). The WASI-Full reflects the combined performance on both the WASI-Verbal and the WASI-Perf. The WASI uses age-based standard scores for composites, with a mean of 100 and standard deviation of 15; higher scores indicate better performance.

Psychomotor Vigilance Test (PVT): The PVT is a simple reaction time test that measures sustained attention and psychomotor function.[34] It is highly sensitive to sleepiness and yet has little practice effect.[35] The test utilized the PVT-192 device which displays a stimulus (red light) intermittently (inter-stimulus interval 2000-10000 ms). As soon as the stimulus appears, the subject is required to press a button on the device with his/her dominant hand, as quickly and as consistently as possible, to stop the reaction time counter, which counts time in milliseconds. For APPLES, a 10-min PVT was administered once in the morning and once in the afternoon. A number of outcomes can be derived from the PVT. For this analysis, the median reaction time for all correct responses (PVT Median) and the mean of the slowest 10% of reaction times (PVT Mean) for the first PVT administration are used.

Paced Auditory Serial Addition Test (PASAT): The PASAT assesses auditory information processing speed and flexibility, as well as calculation ability.[36] Subjects listen to a recording of spoken single numerical digits spaced every 2.4 sec (trial 1), every 2.0 sec (trial 2), every 1.6 sec (trial 3), or every 1.2 sec (trial 4). Their task is to add each new digit to the one immediately prior to it, and verbally respond with the sum. For APPLES, the 60-item version was utilized (60 possible sums per trial). The variable selected for analysis is the total number of correct sums for all trials. The minimum score is 0 and the maximum score is 240.

Hamilton Rating Scale for Depression (HAM-D): The HAM-D is a validated 21-item clinician-administered assessment of the severity of depression.[37] APPLES used a modified version of this test, the GRID Hamilton Rating Scale for Depression which was developed through a broad-based international consensus process to both simplify and standardize administration and scoring in clinical practice and research. In this scale, 17 items (e.g., depressed mood, suicide, work and anhedonia, retardation, agitation, gastrointestinal or general somatic symptoms, hypochondriasis, loss of insight or weight) are scored using either a 3- or 5-point scale based on intensity and frequency, and are summed to provide a single score.

Epworth Sleepiness Scale (ESS): The ESS is a validated self-administered questionnaire that asks an individual to rate his or her probability of falling asleep on a scale of increasing probability from 0 to 3 in 8 different situations.[38] The scores for the 8 questions are summed to obtain a single score from 0 to 24 that is indicative of self-reported sleep propensity.

Stanford Sleepiness Scale (SSS): The SSS asks a person to rate current moment sleepiness on a scale of one to seven.[39] Each numerical rating has an associated descriptor, for example a rating of 1 is described as "feeling active, vital, alert, or wide awake," while a rating of 7 is described as "no longer fighting sleep, sleep onset soon; having dream-like thoughts." For APPLES the SSS was administered at 10:00, 12:00, 14:00, and 16:00; the variable analyzed was the mean score from these 4 trials.

Data Analysis

Data from continuous and interval variables were reviewed to determine whether there were any that had extremely skewed distributions. As a consequence of this inspection, log transformation was performed to normalize the distribution of the following variables: PVT Median, PVT Mean, sleep latency, and % total sleep time with desaturation less than 85% (O_2 Saturation < 85%). Subsequently, 60% of the cohort was randomly selected for further exploration of the data. Bivariate associations between neurocognitive measures, and demographic, sleep and respiratory variables were calculated using analysis of variance, Student's unpaired *t*-test, Mann-Whitney U Test, and Pearson correlation coefficients as appropriate. Except for calculation of the mean data, no further analyses were performed for the SWMT because this measure was developed only for use as an indicator of change from baseline. For each of the other neurocognitive measures, multivariate models were constructed using analysis of covariance incorporating as independent variables only those that had significant bivariate associations with the dependent variable. Some of the sleep and respiratory variables were highly inter-correlated. In these cases, the single variable which had either the best bivariate association or resulted in the best adjusted r^2 was used. After multivariate models were constructed using the 60% exploratory sample, the models then were validated using the remaining 40% of the sample. Final models displayed in the tables reflect results derived only from the 40% validation cohort. Data for continuous and interval variables are expressed as mean ± SD and for categorical variables as a percentage. $P \leq 0.05$ was considered statistically significant. Analyses were performed using SPSS 17.0 (SPSS, Inc., Chicago, IL).

RESULTS

Table 2 shows the demographic data for all participants who had a baseline visit and those who were subsequently randomized stratified by gender. There were 1204 participants who had a baseline polysomnogram; 1098 were ultimately randomized. As expected, both cohorts were obese, and there was a greater proportion of men. In addition, approximately 75% of the participants were Caucasian and the majority were married. Importantly, most participants were relatively well educated, with

Table 2—Baseline demographics for all participants and subsequently randomized participants stratified by gender

	Total Mean (SD) or Count (%)	Men Mean (SD) or Count (%)	Women Mean (SD) or Count (%)	Total Randomized Mean (SD) or Count (%)	Men Mean (SD) or Count (%)	Women Mean (SD) or Count (%)
Participant Enrollment						
Total Number of Participants	1204	769 (63.9)	435 (36.1)	1098	719 (65.5)	379 (34.5)
Site Enrollment						
Stanford University (%)	303 (25.2)	188 (24.4)	115 (26.4)	280 (25.5)	179 (24.9)	101 (26.6)
University of Arizona (%)	346 (28.7)	232 (30.2)	114 (26.2)	314 (28.6)	216 (30.0)	98 (25.9)
St. Mary Medical Center (%)	117 (9.7)	74 (9.6)	43 (9.9)	106 (9.7)	68 (9.5)	38 (10.0)
St. Luke's Hospital (%)	193 (16.0)	132 (17.2)	61 (14.0)	177 (16.1)	121 (16.8)	56 (14.8)
Brigham and Women's Hospital (%)	245 (20.3)	143 (18.6)	102 (23.4)	221 (20.1)	135 (18.8)	86 (22.7)
Demographic Data						
Age (y)	50.7 (12.6)	50.3 (12.7)	51.5 (12.2)	51.5 (12.2)	51.0 (12.4)	52.4 (11.7)
BMI (kg/m^2)	32.0 (7.3)	31.3 (6.4)[2]	33.3 (8.6)	32.2 (7.1)	31.3 (6.1)[2]	34.0 (8.5)
Highest Grade Level (y)	15.5 (2.6)	15.6 (2.6)	15.4 (2.6)	15.5 (2.6)	15.6 (2.6)[1]	15.3 (2.5)
Ethnicity						
White (%)	903 (75.0)	590 (65.3)[4]	313 (34.7)	835 (76.0)	553 (76.9)[4]	282 (74.4)
Native American (%)	20 (1.7)	11 (1.4)	9 (2.1)	18 (1.6)	10 (1.4)	8 (2.1)
Asian Pacific Islander (%)	72 (6.0)	52 (6.8)	20 (4.6)	61 (5.6)	47 (6.5)	14 (3.7)
Black (%)	119 (9.9)	59 (7.7)	60 (13.8)	103 (9.4)	54 (7.5)	49 (12.9)
Hispanic (%)	83 (6.9)	54 (7.0)	29 (6.7)	76 (6.9)	52 (7.2)	24 (6.3)
Other (%)	7 (0.6)	3 (0.4)	4 (0.9)	5 (0.5)	3 (0.4)	2 (0.5)
Marital Status						
Married (%)	695 (57.7)	505 (65.7)[3]	190 (43.7)	634 (57.7)	475 (66.1)[3]	159 (42.0)
Single (%)	297 (24.7)	174 (22.6)	123 (28.3)	262 (23.9)	158 (22.0)	104 (27.4)
Divorced (%)	181 (15.0)	87 (11.3)	94 (21.6)	171 (15.6)	83 (11.5)	88 (23.2)
Widowed (%)	31 (2.6)	3 (0.4)	28 (6.4)	31 (2.8)	3 (0.4)	28 (7.4)
Smoking						
Current Smokers (%)	152 (12.7)	107 (13.9)	45 (10.3)	136 (12.5)	98 (13.7)	38 (10.1)
Alcohol Use						
Consumption > 7 Servings/Week (%)	161 (13.4)	134 (17.7)[2]	27 (6.2)	154 (14.2)	130 (18.3)[2]	24 (6.4)
Servings/Week	3.0 (4.6)	3.7 (5.0)	1.8 (3.3)	3.1 (4.7)	3.7 (5.1)[2]	1.8 (3.4)
Caffeine Use						
Consumption > 14 Servings/Week (%)	555 (46.1)	375 (48.8)[2]	180 (41.4)	510 (47.3)	351 (49.9)[1]	159 (42.3)
Servings/Week	17.8 (15.7)	18.8 (16.2)	15.9 (14.7)	17.8 (15.5)	18.9 (16.2)[2]	15.8 (14.0)
Self Report Sleep Time						
Hours per Weeknight	6.88 (1.25)	6.84 (1.19)	6.95 (1.33)	6.87 (1.21)	6.85 (1.17)	6.90 (1.30)
Hours per Weekend Night	7.57 (1.44)	7.56 (1.37)	7.61 (1.57)	7.56 (1.42)	7.55 (1.34)	7.59 (1.55)

[1]P < 0.05 Men vs. Women; [2]P < 0.01 Men vs. Women; [3]P < 0.01 Difference in proportions in Marital Status between genders; [4]P < 0.01 Difference in proportions in Ethnicity between genders.

an average of approximately 15.5 years of education. There were few gender differences, although women on average had a higher body mass index (BMI), and in comparison to men were less likely to be married, or to consume alcohol or caffeinated beverages. In addition, Black women were more likely to be participants than Black men. Addition of the 206 participants who were not randomized did not change the overall demographics of the baseline cohort.

Baseline neurocognitive measures stratified by gender are shown in Table 1. There were no differences between the cohort that included all 1204 participants who had a baseline polysomnogram and the cohort limited only to the 1098 who were randomized. Thus, data from only the former are shown. It is notable that overall, the participants had above average intelligence as reflected in their WASI scores, and that men scored significantly higher on the WASI-Full and WASI-Perf as well as both components of the WASI-Perf. Men also performed better on both PVT variables and the PASAT. In contrast, women scored better on the BSRT. Women also were slightly sleepier and scored higher on the HAM-D.

Displayed in Table 3 are the baseline polysomnographic data stratified by gender. Except for more extreme values for the respiratory variables, data from the cohort with 1204 participants was not different than the one that was limited only to the

Table 3—Baseline polysomnographic data for all participants[3] stratified by gender

	Total Mean (SD)	Men Mean (SD)	Women Mean (SD)
Sleep Architecture			
Total Recording Time (min)	500.5 (40.4)	499.8 (38.0)	501.6 (44.4)
Total Sleep Time (TST, min)	377.8 (67.6)	376.8 (67.7)	379.6 (67.4)
Sleep Efficiency (%)	78.4 (13.0)	78.3 (13.2)	78.7 (12.3)
Sleep Latency (min)	19.2 (22.8)	17.3 (20.8)[2]	22.7 (25.6)
REM Latency (min)	137.8 (82.2)	134.1 (80.5)[1]	144.3 (84.7)
Wake After Sleep Onset (min)	85.0 (52.9)	87.5 (54.7)[1]	80.5 (49.2)
Stage 1 (%TST)	18.5 (14.5)	20.0 (15.6)[2]	15.9 (12.0)
Stage 2 (% TST)	60.7 (13.6)	59.7 (14.7)[2]	62.3 (11.1)
Stage 3 (% TST)	2.6 (4.8)	2.1 (4.5)[2]	3.3 (5.4)
Stage 4 (% TST)	0.6 (2.2)	0.4 (1.8)[2]	0,9 (2.7)
Stage REM (% TST)	17.6 (7.1)	17.6 (7.1)	17.5 (7.1)
Respiratory Indices			
Apnea Hypopnea Index (AHI, events/h TST)	38.2 (26.7)	40.8 (26.7)	33.5 (26.1)[2]
AHI < 10[4]	92 (7.6)	38 (4.9)	54 (12.4)
10 < AHI ≤ 15[4]	144 (12.0)	81 (10.5)	63 (14.5)
15 < AHI ≤ 30[4]	345 (28.7)	224 (29.1)	121 (27.8)
30 < AHI ≤ 50[4]	285 (23.7)	180 (23.4)	105 (24.1)
50 < AHI ≤ 75[4]	204 (16.9)	145 (18.9)	59 (13.6)
AHI > 75[4]	134 (11.1)	101 (13.1)	33 (7.6)
Central Apnea Index (events/h TST)	1.1 (4.3)	1.4 (4.0)	0.6 (4.9)[2]
Minimum Oxygen Saturation (%)	81.2 (8.8)	81.0 (8.6)	81.6 (9.0)
O_2 Saturation < 85% (% TST)	2.6 (8.1)	3.2 (9.2)	1.6 (5.5)[2]
Oxygen Desaturation Index (events/h TST)	25.5 (25.1)	27.3 (26.0)	22.3 (23.3)[2]
Sleep Fragmentation			
Arousal Index (arousals/h TST)	28.9 (21.0)	30.9 (22.1)[2]	25.4 (18.2)
Periodic Leg Movement Index (events/h TST)	6.7 (15.6)	7.4 (16.0)[1]	5.5 (14.8)

[1]P < 0.05 Men vs. Women; [2]P < 0.01 Men vs. Women; [3]N = 1204 except for Stage REM (N = 1194, 10 participants without REM sleep) and Minimum Oxygen Saturation (N = 1202, 2 participants with missing data); [4] Data represent N (% of N).

1098 who were randomized. Thus, the table shows only summary data from former group. In comparison to recently published normative data from a general population, sleep quality was suboptimal with moderately reduced sleep efficiency, increased percent stages 1 and 2 sleep, increased arousal index, and correspondingly decreased stages 3, 4, and REM sleep.[40] Sleep and REM latencies were longer in women. Women also had less percent stage 1 and greater percent stages 3 and 4 sleep. With respect to respiratory variables, the mean AHI in all participants of 38.2/h was indicative of relatively severe OSA with greater severity in men than women. Additionally, 28% of all participants had an AHI > 50/h. The O_2 Saturation < 85% and the desaturation index, but not the minimum O_2 saturation, were more severe in men as well. The central apnea index was quite low, suggesting that almost all apneic and hypopneic events were obstructive. There were no significant correlations between age and AHI ($r = 0.039$, $P = 0.171$) or oxygen desaturation index ($r = -0.008$, $P = 0.789$). However, weak negative correlations were noted between age and O_2 Saturation < 85% ($r = -0.081$, $P = 0.005$) and Minimum O_2 Saturation ($r = -0.071$, $P = 0.014$).

Bivariate relationships between neurocognitive measures and demographic factors (data not shown) generally found strong and consistent associations with years of education; greater amounts of education corresponding with better performance. Less consistent, but in many cases highly significant were associations with race (Non-Hispanic Whites with better performance than Other Ethnicity), gender (men with better performance), and age (better or worse performance depending on test instrument). Mood as reflected by the HAM-D score also impacted performance on other neurocognitive measures with higher scores associated with worse performance (data not shown). In contrast, the effect of sleepiness and alterations in sleep architecture were inconsistent with the strongest bivariate associations found with total sleep time (lower amounts with worse performance), sleep latency (longer latency with worse performance), and percent Stage 1 sleep (higher percent with worse performance).

In Table 4 are displayed the correlation coefficients between respiratory variables and neurocognitive measures from the 60% random sample used as the exploratory dataset. As can be observed, there are consistent but weak associations between the AHI and desaturation indices, and the WASI-Full, WASI-Verbal, and PASAT. There also were some inconsistent and weak correlations between the AHI and desaturation indices, and a few of the other neurocognitive measures, most notably the 2 PVT indicators and the ESS. The highest correlation coefficient noted was 0.266 between the Log PVT Mean and O_2 Saturation < 85%. Of particular note is the absence of associations between any of the respiratory variables and the WASI-Perf and the Pathfinder.

Table 5 shows the final multivariate models for the neurocognitive variables computed from the 40% of the cohort reserved for the validation dataset. For every measure except for the BSRT, ESS, and the 2 reaction time (PVT) indicators, years of education was the strongest predictor of performance. For the WASI measures, race other than Non-Hispanic White was a consistent negative impact factor on performance as well. However, the amount of variance explained by these models was modest with r^2 ranging from 0.039 to 0.296. Furthermore, the impact of respiratory and sleep variables was quite limited. For the 2 primary outcome variables analyzed, % Stage 1 Sleep negatively impacted performance on the Pathfinder Test, but there was no effect from any other respiratory or sleep variable. For secondary outcomes, statistically significant contributions by oxygen desaturation indices to poorer performance were found on the WASI-Block Design, ESS, PVT Mean, and the PASAT. Borderline statistical significance was noted for WASI-Vocab and WASI-Perf. Overall AHI was not related to neurocognition on any outcomes (For outcomes

without significant associations, data shown are with desaturation indices). However, the contribution of either AHI or oxygen desaturation to the overall variance explained in the models was < 2% in all cases. Of additional interest is that sleepiness assessed by either the ESS or SSS only affected performance on the PVT Median.

To determine whether there were associations between high values of AHI and neurocognitive function, comparisons of neurocognitive function were made using only participants with an AHI < 10 events/h or AHI ≥ 30 events/h. No significant associations were observed (data not shown). This analysis was repeated using only participants with an AHI < 10 events/h or AHI ≥ 50 events/h. Performance on the Pathfinder Test and the BSRT was worse for those with an AHI ≥ 50 events/h on bivariate analyses (Pathfinder: 22.8 ± 0.9 vs. 25.0 ± 0.5 sec, P = 0.03; BSRT: 52.1 ± 1.5 vs. 48.6 ± 0.6 words correct, P = 0.03). However, these findings were no longer significant after multivariate analyses. A similar analysis was performed to determine if there was an impact of severe oxygen desaturation. For this analysis, only participants who met one of the 2 following criteria were used. Those who had 2% or more of their PSG total sleep time less than 85% were classified as severe oxygen desaturation (Severe Desaturation) and those who never had any PSG saturation less than 90% were considered as having minimal or no desaturation (No Desaturation). As can be shown in Table 6, performance on the WASI-Full, WASI-Perf, WASI-BD, WASI-MATR, and PVT Mean Reaction Time was negatively impacted by severe oxygen desaturation. Slight trends for worse performance also were noted for WASI-Verbal, WASI-Simil, PVT Median Reaction Time, and PASAT (data not shown, all 0.10 < P < 0.20). To investigate whether those in the severe oxygen desaturation group had chronic lung disease, the prevalence of asthma and chronic obstructive lung disease was compared between the Severe Desaturation and No Desaturation groups. No differences in prevalence were observed although the number with these conditions was small. However, waking oxygen saturations were lower in the Severe Desaturation group (96.2% ± 1.1% vs. 92.1% ± 2.4%, P < 0.001), and BMI was higher (38.8 ± 8.2 vs. 27.4 ± 4.3 kg/m^2, P < 0.001). Furthermore, 10 of those in the Severe Desaturation group had waking oxygen saturations below 90%.

DISCUSSION

In this study using a large cohort of participants entering a randomized study of the impact of long-term CPAP on neurocognitive functioning, cross-sectional analysis of the associations between indices of sleep quality and OSA, and various measures of neurocognitive performance showed only weak and inconsistent relationships at baseline. Sleep quality as assessed by laboratory polysomnography had relatively little relationship with neurocognitive performance. Although respiratory indices of OSA severity appeared to be more associated with worse neurocognitive function than sleep quality indicators, their effect still was quite modest overall. Nevertheless, there was some evidence that oxygen desaturation negatively correlated with performance to a slight degree on a broad range of neurocognitive measures. This was particularly evident for those participants with severe oxygen desaturation (i.e., desaturation < 85%).

Table 4—Correlation between neurocognitive and sleep disordered breathing measures (N = 722)[1]

Neurocognitive Measures[2]	Sleep Disordered Breathing Measures			
	AHI	Minimum SpO_2	% TST < 85% Saturation	Desaturation Index
WASI Full				
r	-0.100	0.123	-0.062	-0.130
p	0.008	0.001	0.099	< 0.001
n	711	710	711	711
WASI Verbal				
r	-0.122	-0.128	-0.075	-0.150
p	0.001	0.001	0.046	< 0.001
n	712	711	712	712
WASI Vocabulary				
r	-0.108	0.126	-0.068	-0.137
p	0.004	0.001	0.069	< 0.001
n	712	711	712	712
WASI Similarities				
r	-0.113	0.102	-0.063	-0.134
p	0.003	0.007	0.095	< 0.001
n	712	711	712	712
WASI Performance				
r	-0.047	0.08	-0.033	-0.078
p	0.210	0.033	0.385	0.037
n	711	710	711	711
WASI Block Design				
r	-0.028	0.085	-0.004	-0.043
p	0.462	0.024	0.916	0.255
n	712	711	712	712
WASI Matrix Reasoning				
r	-0.056	0.040	-0.039	-0.092
p	0.132	0.290	0.304	0.014
n	711	710	711	711
Pathfinder				
r	0.043	-0.049	-0.031	0.021
p	0.252	0.188	0.404	0.575
n	713	712	713	713
BSRT				
r	-0.104	0.094	-0.024	-0.079
p	0.005	0.012	0.524	0.035
n	713	712	713	713
LOG PVT Median				
r	0.028	-0.036	0.112	0.069
p	0.464	0.335	0.003	0.065
n	706	705	706	706
LOG PVT Mean				
r	0.116	-0.126	0.266	0.163
p	0.002	0.001	< 0.001	< 0.001
n	706	705	706	706
PASAT				
r	-0.100	0.141	-0.084	-0.110
p	0.008	< 0.001	0.027	0.004
n	703	702	703	702
ESS				
r	0.128	-0.086	0.187	0.118
p	0.001	0.022	< 0.001	0.001
n	722	721	722	722

[1]Analyses performed on the 60% exploratory dataset; [2]Abbreviations of neurocognitive measures defined in the text.

We observed that increased % Stage 1 sleep was weakly associated with poorer performance on the Pathfinder test. The Pathfinder assesses attention and psychomotor function. Although the magnitude of the association was modest, this finding suggests that diminished sleep quality in those with OSA contributes to the difficulty in maintaining attention and concentration often observed in these patients. To our knowledge, there have been no previous reports of objectively measured sleep quality negatively impacting performance on either of the Trail Making tasks. However, some, but not all previous studies have found a relationship between sleep restriction or deprivation and poorer performance on Trail Making.[41] Our results would be consistent with those studies finding such a negative relationship.

Notwithstanding the stage 1 sleep Pathfinder correlation, overall sleep quality and sleep duration had little relationship with neurocognitive function. Although sleep deprivation negatively affects neurocognitive performance on a variety of domains including learning, executive function and vigilance,[42] the disturbances in sleep architecture and duration observed in the cohort as a whole were relatively mild, and were likely insufficient to result in a demonstrable effect on the measures used in the study.

In contrast to the relative lack of association between sleep quality and sleep duration, and neurocognitive performance, we found consistent, albeit weak associations between several indices of oxygen desaturation and results from WASI-Block Design, PVT Mean, the ESS and the PASAT. As reviewed by Beebe and colleagues,[17] most studies have not found that OSA affects overall intelligence and verbal abilities. While our failure to demonstrate any associations with OSA severity on the WASI-Full, WASI-Verbal, and WASI-Perf are consistent with these previous studies, the trend towards poorer performance on the WASI-Vocab does suggest that OSA can negatively affect performance on these domains to a small degree. The preponderance of previous studies, however, indicate that OSA impairs performance on visual ability, executive function, motor speed, and vigilance. Our finding that performance on the PVT and PASAT was diminished is consistent with these previous observations although their overall contribution in relation to other factors such as previous education and ethnicity was small. Most previous studies of the impact of OSA on neurocognitive function have utilized participants recruited from clinical settings, and thus may have selected those who had clinically evident neurocognitive impairment. In addition, recent studies have indicated that the brain attempts to compensate to avoid decrements in performance in the setting of untreated OSA.[43,44] It is possible that in a population of previously untreated individuals, compensatory mechanisms partially masked neurocognitive impairment. Nevertheless, the small effect observed is consistent with results from another large general population cohort.[20]

Another perspective on the strength of the associations found in this study can be found by examining the effect sizes derived from the multivariate analyses. As shown in Table 5, the

Table 5—Regression models for factors associated with neurocognition (N = 482[9])

Factor	B[6]	SE[7]	95% CI		Partial η Squared[8]	P
			Lower	Upper		
WASI Full						
Years Education	2.203	0.220	1.861	2.725	0.190	< 0.01
Race[1]	-8.393	1.340	-11.026	-5.759	0.078	< 0.01
AHI	-0.034	0.220	-0.077	0.010	0.005	0.13
			Adjusted r^2 = 0.279			
WASI Verbal						
Years Education	2.422	0.218	1.994	2.850	0.211	< 0.01
Race[1]	-7.904	1.350	-10.560	-5.250	0.069	< 0.01
AHI	-0.024	0.022	-0.067	0.019	0.003	0.27
			Adjusted r^2 = 0.296			
WASI Vocabulary						
Years Education	1.640	0.160	1.326	1.953	0.186	< 0.00
Race[1]	-6.117	0.980	-8.040	-4.190	0.078	< 0.00
Desaturation Index	-0.033	0.017	-0.066	0.001	0.008	0.06
			Adjusted r^2 = 0.280			
WASI Similarities						
Years Education	1.278	0.129	1.023	1.532	0.174	< 0.00
Race[1]	-4.132	0.804	-5.710	-2.550	0.054	< 0.00
Desaturation Index	-0.012	0.014	-0.039	0.015	0.002	0.38
			Adjusted r^2 = 0.248			
WASI Performance						
Years Education	1.600	0.238	1.132	2.068	0.089	< 0.01
Race[1]	-6.910	1.454	-9.769	-4.054	0.046	< 0.01
Desaturation Index	-0.043	0.026	-0.094	0.007	0.006	0.09
			Adjusted r^2 = 0.154			
WASI Block Design						
Years Education	0.852	0.156	0.545	1.158	0.056	< 0.01
Race[1]	-7.370	1.932	-11.170	-3.570	0.033	< 0.01
Gender[2]	5.961	2.369	1.423	10.499	0.020	0.01
Marital Status[5]	5.159	2.521	0.201	10.113	0.013	0.04
Minimum SpO_2	0.141	0.051	0.041	0.240	0.017	0.01
			Adjusted r^2 = 0.163			
WASI Matrix Reasoning						
Years Education	0.959	0.151	0.662	1.255	0.080	< 0.01
Race[1]	-4.430	0.921	-6.240	-2.620	0.047	< 0.01
AHI	-0.200	0.015	-0.050	0.010	0.004	0.19
			Adjusted r^2 = 0.143			

[1]Comparison is Non-Hispanic White (Reference Group) vs. Other Ethnicity/Race; [2]Reference group is Women; [3] PVT Median Reaction Time was log transformed, and coefficients are expressed as change in log seconds per unit change in the independent variable; [4]PVT Mean Reaction Time was log transformed, and coefficients are expressed as change in log seconds per unit change in the independent variable; [5]Reference group is Not Married; [6]Estimated regression coefficient; [7]Estimated standard error; [8]Partial eta squared is an estimate of effect size; [9]Result computed from 40% validation sample.

Table 5 continues on the following page

Table 5 *(continued)*—Regression models for factors associated with neurocognition (N = 482[9])

Factor	B[6]	SE[7]	95% CI		Partial η Squared[8]	P
			Lower	Upper		
Pathfinder						
Years Education	-0.307	0.102	-0.508	-0.106	0.019	< 0.01
% Stage 1 Sleep	0.063	0.024	0.017	0.109	0.015	0.02
Age	0.234	0.021	0.192	0.276	0.208	< 0.01
AHI	0.005	0.011	-0.017	0.028	0.000	0.63
			Adjusted r^2 = 0.245			
BSRT						
Age	-0.235	0.033	-0.299	0.170	0.099	< 0.01
Gender[2]	-3.227	1.639	-6.447	-0.007	0.008	< 0.05
Years Education	0.779	0.153	0.478	1.080	0.053	< 0.01
Race[1]	4.949	1.438	2.123	7.776	0.025	< 0.01
Minimum SpO_2	0.019	0.052	-0.083	0.121	0.000	0.72
			Adjusted r^2 = 0.180			
PVT Median Reaction Time[3]						
Age	< 0.000	< 0.000	< 0.000	0.001	0.008	0.05
Stanford Sleepiness Scale	0.007	0.003	0.002	0.012	0.014	0.01
Gender[2]	-0.022	0.005	-0.032	-0.012	0.039	< 0.01
Desaturation Index	< 0.000	< 0.000	< 0.000	< 0.000	0.000	0.75
			Adjusted r^2 = 0.065			
PVT Mean Reaction Time[4]						
%TST < 85% Saturation	0.004	0.001	0.002	0.007	—	< 0.00
			Adjusted r^2 = 0.039			
PASAT						
Years Education	3.606	0.718	2.196	5.016	0.058	< 0.01
Age	-1.215	0.156	-1.522	-0.909	0.121	< 0.01
Minimum SpO_2	0.536	0.243	0.058	1.014	0.009	0.03
Race[1]	-14.730	6.709	-27.916	-1.545	0.015	0.03
			Adjusted r^2 = 0.185			
Epworth Sleepiness Scale						
HamD Total Score	0.130	0.046	0.040	0.220	0.017	< 0.01
Minimum SpO_2	-0.061	0.023	-0.107	-0.016	0.015	< 0.01
Total Sleep Time	0.007	0.003	0.001	0.012	0.012	0.02
			Adjusted r^2 = 0.048			

[1]Comparison is Non-Hispanic White (Reference Group) vs. Other Ethnicity/Race; [2]Reference group is Women; [3] PVT Median Reaction Time was log transformed, and coefficients are expressed as change in log seconds per unit change in the independent variable; [4]PVT Mean Reaction Time was log transformed, and coefficients are expressed as change in log seconds per unit change in the independent variable; [5]Reference group is Not Married; [6]Estimated regression coefficient; [7]Estimated standard error; [8]Partial eta squared is an estimate of effect size; [9]Result computed from 40% validation sample.

partial η squared values for % Stage 1 on the Pathfinder Test, Minimum SpO_2 on the PASAT and ESS were 0.015, 0.009, and 0.012, respectively. Partial η squared is a marker of effect size, and values of approximately 0.01 are considered small.[45] In contrast, in one recent study, the impact of obesity on risk of hypertension was found to have a medium effect size.[46]

Mechanisms most commonly invoked to explain the association between OSA and neurocognition are the negative effects of sleep disordered breathing events on sleep continuity and/or the impact of repetitive hypoxemic events. In this study, oxygen desaturation indices explained virtually all of the observed associations between OSA severity and neurocognitive performance. Furthermore, when we restricted analysis only to those with severe OSA (AHI ≥ 30 or ≥ 50), we found no or few associations on bivariate analyses, but did observe that relatively severe oxygen desaturation was predictive of worse performance on several of our measures. In contrast, there were no significant associations with AHI in our adjusted models. Similar findings have been noted by others in large population cohorts.[22,23] Not surprisingly, as we observed, relatively severe oxygen desaturation during a PSG is associated with lower waking oxygen saturations. In addition, these subjects had a higher BMI. Therefore, it is possible that obesity hypoventilation syndrome was present as well as obstructive sleep apnea, and increased the risk of worse neurocognitive performance. Because arterial blood gases were not measured as part of this study, we could not determine whether any of our subjects were hypoventilating. It is possible that the strength of the associations we found are underestimated because those with very severe oxygen desaturation were excluded. Nevertheless, taken as a whole, our data indicate that decrements in neurocognitive performance related to OSA are predominantly a result of hypoxemia, rather than the frequency of sleep disordered breathing events or reductions in sleep quality. Our results are consistent with those reported from the Sleep Heart Health Study,[20] as well as intermittent hypoxia models of sleep disordered breathing in experimental animals.[47] Alternatively, it has been suggested by others that sleepiness, most likely caused by sleep fragmentation, is the mechanism underlying neurocognitive dysfunction in OSA.[48] This was not supported by our results.

The most powerful factor explaining better neurocognitive performance in our study was educational level. Non-Hispanic White ethnicity was an important positive predictor on many measures as well. These findings are not surprising given the large amount of data demonstrating a correlation between intelligence testing and educational attainment as well as poorer performance by many minority groups on standardized intelligence tests. Although there were some gender differences in neurocognitive performance when the data were unadjusted, multivariate analyses found that gender was not an important factor on most neurocognitive measures, and have not been found to be important in previous studies in large populations.

There are several limitations and caveats to our results. First, although we studied a large cohort of participants, the analysis is cross-sectional. Thus, causation cannot be unequivocally demonstrated. Second, inasmuch as the data utilized were collected as part of the baseline visit for a randomized controlled trial of therapeutic vs. sub-therapeutic CPAP in OSA individuals, virtually all of the participants in the cohort had some degree of

Table 6—Regression models for factors associated with severe oxygen desaturation[9]

Factor	B[6]	SE[7]	95% CI		Partial η Squared[8]	P
			Lower	Upper		
WASI Full						
Race[1]	-7.350	2.374	-12.047	-2.653	0.070	0.002
Years Education	2.164	0.430	1.314	3.014	0.166	< 0.001
Desaturation[3]	6.461	2.164	2.180	10.743	0.066	0.003
			Adjusted r^2 = 0.287			
WASI Performance						
Race[1]	-7.557	2.547	-12.598	-2.516	0.065	0.004
Years Education	1.535	0.461	0.623	2.448	0.080	0.001
Desaturation[3]	7.828	2.322	3.233	12.423	0.082	0.001
			Adjusted r^2 = 0.220			
WASI Block Design						
Race[1]	-5.071	1.678	-8.390	-1.751	0.068	0.003
Years Education	0.902	0.302	0.305	1.499	0.066	0.003
Desaturation[3]	4.340	1.530	1.310	7.366	0.060	0.005
Gender	4.527	1.472	1.613	7.440	0.070	0.003
Marital Status[4]	3.176	1.455	0.296	6.055	0.036	0.031
			Adjusted r^2 = 0.260			
WASI Matrix Reasoning						
Race[1]	-3.473	1.683	-6.803	-0.143	0.041	0.032
Years Education	0.993	0.305	0.390	1.595	0.077	0.001
Desaturation[3]	4.558	1.534	1.523	7.594	0.065	0.004
			Adjusted r^2 = 0.173			
PVT Mean Reaction Time[5]						
Desaturation[3]	0.073	0.030	0.013	0.132	0.000	0.017
			Adjusted r^2 = 0.043			

[1]Comparison is Non-Hispanic White (Reference Group) vs. Other Ethnicity/Race; [2]Reference group is Women; [3]Reference group is ≥ 2% of total sleep time with oxygen saturation < 85%; [4]Reference group is Not Married; [5]PVT Mean Reaction Time was log transformed, and coefficients are expressed as change in log seconds per unit change in the independent variable; [6]Estimated regression coefficient; [7]Estimated standard error; [8]Partial eta squared is an estimate of effect size; [9]N = 47 (minimal or no desaturation); N = 83 (severe desaturation); See text for definitions.

OSA. Therefore, there was a paucity of participants who did not have any OSA. If neurocognitive impairment in OSA is more of a threshold rather than a "dose-response" type of effect, it is possible that our inability to demonstrate strong associations between OSA severity and neurocognition may be explained by the lack of an adequate comparison group. Third, our study population was relatively well educated and had above average intelligence. There is some suggestion that persons with higher levels of intelligence may be more resistant to any adverse impact of OSA on neurocognition.[49] Fourth, reviews of the relationship between OSA and neurocognitive function indicate that executive function is one domain that is frequently impaired in those with OSA.[17,50] Although one of our primary outcome measures, the SWMT is sensitive to changes in executive function, it was not designed to be used in cross-sectional analysis, and thus not included in this analysis. Furthermore, our neurocognitive test battery, while comprehensive, did not utilize every measure previously shown to be impacted by OSA. Moreover, due to limitations in the time allowed for testing it was not possible to conduct a long duration vigilance task (e.g., a 14-20 minute continuous performance test). Thus, it is possible that the associations between OSA severity and neurocognitive function may be stronger than we have observed. Fifth, participants in our study were recruited both from clinical venues as well as from the general population through advertisement. Individuals in the latter category may have had fewer symptoms, and thus have had less neurocognitive impact from their OSA. Unfortunately, recruitment source was not collected as part of the study although it is estimated that 70% were recruited from the general population. Sixth, although the mean AHI in this study was 38.2/hour, hypopneas were identified if there was a decrement in nasal pressure, and an associated > 3% oxygen desaturation or an arousal. In contrast, the Sleep Heart Health Study used a thermistor and allowed only associated 4% oxygen desaturations to identify hypopneas.[4,5] Thus, the severity of OSA in our cohort may have been less than in other studies. Last, the variance explained by all of the variables included our models was relatively small. This suggests a number of other factors not included in our study are important in explaining neurocognitive performance in the setting of OSA.

In summary, in a large cohort of individuals with OSA recruited for a large randomized controlled intervention trial of CPAP, associations between severity of OSA as well as objectively determined sleep quality and duration, and neurocognitive function were weak and inconsistent. When impairment was observed, oxygen desaturation rather than the number of sleep-disordered breathing events affected neurocognitive performance. These findings indicate that the relative impact of OSA on neurocognition may be minor in most individuals with this condition.

ACKNOWLEDGMENTS

APPLES was funded by HL068060 from the National Heart, Lung and Blood Institute**.** The APPLES pilot studies were supported by grants from the American Academy of Sleep Medicine and the Sleep Medicine Education and Research Foundation to Stanford University and by the National Institute of Neurological Disorders and Stroke (N44NS002394) to SAM Technology.

In addition, APPLES investigators gratefully recognize the vital input and support of Dr. Sylvan Green who died before the results of this trial were analyzed, but was instrumental in its design and conduct.

Administrative Core

Clete A. Kushida, MD, PhD; Deborah A. Nichols, MS; Eileen B. Leary, BA, RPSGT; Pamela R. Hyde, MA; Tyson H. Holmes, PhD; Daniel A. Bloch, PhD; William C. Dement, MD, PhD

Data Coordinating Center

Daniel A. Bloch, PhD; Tyson H. Holmes, PhD; Deborah A. Nichols, MS; Rik Jadrnicek, Microflow; Ric Miller,

Microflow; Usman Aijaz, MS; Aamir Farooq, PhD; Darryl Thomander, PhD; Chia-Yu Cardell, RPSGT; Emily Kees; Michael E. Sorel, MPH; Oscar Carrillo RPSGT; Tami Crabtree, MS; Ray Balise, PhD; Tracy Kuo, PhD

Clinical Coordinating Center

Clete A. Kushida, MD, PhD; William C. Dement, MD, PhD; Pamela R. Hyde, MA; Rhonda M. Wong, BA; Pete Silva; Max Hirshkowitz, PhD; Alan Gevins, DSc; Gary Kay, PhD; Linda K. McEvoy, PhD; Cynthia S. Chan, BS; Sylvan Green, MD

Clinical Centers

Stanford University

Christian Guilleminault, MD; Eileen B. Leary, BA, RPSGT; David Claman, MD; Stephen Brooks, MD; Julianne Blythe, PA-C, RPSGT; Jennifer Blair, BA; Pam Simi; Ronelle Broussard, BA; Emily Greenberg, MPH; Bethany Franklin, MS; Amirah Khouzam, MA; Sanjana Behari Black, BS, RPSGT; Viola Arias, RPSGT; Romelyn Delos Santos, BS; Tara Tanaka, PhD

University of Arizona

Stuart F. Quan, MD; James L. Goodwin, PhD; Wei Shen, MD; Phillip Eichling, MD; Rohit Budhiraja, MD; Charles Wynstra, MBA; Cathy Ward; Colleen Dunn, BS; Terry Smith, BS; Dane Holderman; Michael Robinson, BS; Osmara Molina, BS; Aaron Ostrovsky; Jesus Wences; Sean Priefert; Julia Rogers, BS; Megan Ruiter, BS; Leslie Crosby, BS, RN

St. Mary Medical Center

Richard D. Simon Jr, MD; Kevin Hurlburt, RPSGT; Michael Bernstein, MD; Timothy Davidson, MD; Jeannine Orock-Takele, RPSGT; Shelly Rubin, MA; Phillip Smith, RPSGT; Erica Roth, RPSGT; Julie Flaa, RPSGT; Jennifer Blair, BA; Jennifer Schwartz, BA; Anna Simon, BA; Amber Randall, BA

St. Luke's Hospital

James K. Walsh, PhD; Paula K. Schweitzer, PhD; Anup Katyal, MD; Rhody Eisenstein, MD; Stephen Feren, MD; Nancy Cline; Dena Robertson, RN; Sheri Compton, RN; Susan Greene; Kara Griffin, MS; Janine Hall, PhD

Brigham and Women's Hospital

Daniel J. Gottlieb, MD, MPH; David P. White, MD; Denise Clarke, BSc, RPSGT; Kevin Moore, BA; Grace Brown, BA; Paige Hardy, MS; Kerry Eudy, PhD; Lawrence Epstein, MD; Sanjay Patel, MD

Sleep HealthCenters

For the use of their clinical facilities to conduct this research.

DISCLOSURE STATMENT

Respironics, Inc. supplied the CPAP and sham CPAP devices for this study. There was no other industry support. Dr. Walsh has received research support through grants to his institution from Pfizer, Merck, Somaxon, Evotec, Actelion, Vanda, Neurogen, Sanofi-Aventis, Ventus Medical, Respironics, and Jazz. He has consulted for Pfizer, Sanofi-Aventis, Cephalon, Schering-Plough/Organon, Neurocrine, Takeda, Actelion, Sepracor, Jazz, Respironics, Transcept, Neurogen, GlaxoSmithKline, Somaxon, Eli Lilly, Evotec, Merck, Kingsdown, Vanda, Ventus Medical, and Somnus. Dr. Schweitzer has been the principal investigator by studies sponsored through grants to his institution from Neurogen, Pfizer, Merck, Somnus, and Ventus Medical. Dr. Simon has received research support from Ventus Medical and has participated in speaking engagements for Asante Communications. Dr. Kay has received research support from Cephalon and Merck; he is a 50% owner of CogScreen, which provided licenses and tests to the APPLES study at cost. Dr. Dement has participated in paid speaking engagements. Dr. Kushida Has received research support through grants to Stanford University from Ventus Medical, Merck, Respironics, Jazz, Cephalon, GlaxoSmithKline, Pacific Medico, and Xenoport. Dr. Hirshkowitz has participated in speaking engagements for Cephalon and Takeda and has consulted for Cephalon and Sanofi-Aventis.

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, Palta M, Dempsey J, Peppard PE, Nieto FJ, Hla KM. Burden of sleep apnea: rationale, design, and major findings of the Wisconsin Sleep Cohort study. WMJ 2009;108:246-9.
3. Peker Y, Hedner J, Norum J, Kraiczi H, Carlson J. Increased incidence of cardiovascular disease in middle-aged men with obstructive sleep apnea: a 7-year follow-up. Am J Respir Crit Care Med 2002;166:159-65.
4. Punjabi NM, Caffo BS, Goodwin JL, et al. Sleep-disordered breathing and mortality: a prospective cohort study. PLoS Med 2009;6:e1000132.
5. Redline S, Yenokyan G, Gottlieb DJ, et al. Obstructive sleep apnea hypopnea and incident stroke: The Sleep Heart Health Study. Am J Respir Crit Care Med 2010;182:269-77.
6. Yaggi HK, Concato J, Kernan WN, Lichtman JH, Brass LM, Mohsenin V. Obstructive sleep apnea as a risk factor for stroke and death. N Engl J Med 2005;353:2034-41.
7. Marin JM, Carrizo SJ, Vicente E, Agusti AG. Long-term cardiovascular outcomes in men with obstructive sleep apnoea-hypopnoea with or without treatment with continuous positive airway pressure: an observational study. Lancet 2005;365:1046-53.
8. Young T, Peppard P, Palta M, et al. Population-based study of sleep-disordered breathing as a risk factor for hypertension. Arch Intern Med 1997;157:1746-52.
9. Decary A, Rouleau I, Montplaisir J. Cognitive deficits associated with sleep apnea syndrome: a proposed neuropsychological test battery. Sleep 2000;23:369-81.
10. Engleman HM, Kingshott RN, Martin SE, Douglas NJ. Cognitive function in the sleep apnea/hypopnea syndrome (SAHS). Sleep 2000;Sleep, 23(Suppl 4):S102-8.
11. Dealberto MJ, Pajot N, Courbon D, Alperovitch A. Breathing disorders during sleep and cognitive performance in an older community sample: the EVA Study. J Am Geriatr Soc 1996;44:1287-94.
12. Greenberg GD, Watson RK, Deptula D. Neuropsychological dysfunction in sleep apnea. Sleep 1987;10:254-62.
13. Kotterba S, Rasche K, Widdig W, et al. Neuropsychological investigations and event-related potentials in obstructive sleep apnea syndrome before and during CPAP-therapy. J Neurol Sci 1998;159:45-50.
14. Montplaisir J, Bedard MA, Richer F, Rouleau I. Neurobehavioral manifestations in obstructive sleep apnea syndrome before and after treatment with continuous positive airway pressure. Sleep 1992;15:S17-9.
15. Naegele B, Pepin JL, Levy P, Bonnet C, Pellat J, Feuerstein C. Cognitive executive dysfunction in patients with obstructive sleep apnea syndrome (OSAS) after CPAP treatment. Sleep 1998;21:392-7.

16. Naegele B, Thouvard V, Pepin JL, et al. Deficits of cognitive executive functions in patients with sleep apnea syndrome. Sleep 1995;18:43-52.
17. Beebe DW, Groesz L, Wells C, Nichols A, McGee K. The neuropsychological effects of obstructive sleep apnea: a meta-analysis of norm-referenced and case-controlled data. Sleep 2003;26:298-307.
18. Redline S, Strauss ME, Adams N, et al. Neuropsychological function in mild sleep-disordered breathing. Sleep 1997;20:160-7.
19. Kim HC, Young T, Matthews CG, Weber SM, Woodward AR, Palta M. Sleep-disordered breathing and neuropsychological deficits. A population-based study. Am J Respir Crit Care Med 1997;156:1813-9.
20. Quan SF, Wright R, Baldwin CM, et al. Obstructive sleep apnea-hypopnea and neurocognitive functioning in the Sleep Heart Health Study. Sleep Med 2006;7:498-507.
21. Weaver TE. Outcome measurement in sleep medicine practice and research. Part 2: assessment of neurobehavioral performance and mood. Sleep Med Rev 2001;5:223-36.
22. Boland LL, Shahar E, Iber C, et al. Measures of cognitive function in persons with varying degrees of sleep-disordered breathing: the Sleep Heart Health Study. J Sleep Res 2002;11:265-72.
23. Telakivi T, Kajaste S, Partinen M, Koskenvuo M, Salmi T, Kaprio J. Cognitive function in middle-aged snorers and controls: role of excessive daytime somnolence and sleep-related hypoxic events. Sleep 1988;11:454-62.
24. Kushida CA, Nichols DA, Quan SF, et al. The Apnea Positive Pressure Long-term Efficacy Study (APPLES): rationale, design, methods, and procedures. J Clin Sleep Med 2006;2:288-300.
25. American Academy of Sleep Medicine. International classification of sleep disorders, 2nd ed.: diagnostic and coding manual. Westchester, IL: American Academy of Sleep Medicine, 2005.
26. Rechtschaffen A, Kales A. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Washington, DC: US Government Printing Office, National Institute of Health Publication, 1968.
27. American Academy of Sleep Medicine Taskforce. Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. Sleep 1999;22:667-89.
28. Kay GG. Cogscreen. http://www.cogscreen.com/overview.shtml; Last accessed May 7, 2010.
29. Jokic R, Klimaszewski A, Crossley M, Sridhar G, Fitzpatrick MF. Positional treatment vs continuous positive airway pressure in patients with positional obstructive sleep apnea syndrome. Chest 1999;115:771-81.
30. Buschke H. Selective reminding for analysis of memory and learning. J Verbal Learn Verbal Behav 1973;12:534-50.
31. Hannay HJ, Levin HS. Selective reminding test: an examination of the equivalence of four forms. J Clin Exp Neuropsychol 1985;7:251-63.
32. Smith ME, McEvoy LK, Gevins A. The impact of moderate sleep loss on neurophysiologic signals during working-memory task performance. Sleep 2002;25:784-94.
33. Wechsler D. Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: Psychological Corp., 1999.
34. Dinges DF, Powell JW. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. Behav Res Methods Instrum Comput 2007;17:652-5.
35. 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.
36. Gronwall DMA. Paced auditory serial-addition task: a measure of recovery from concussion. Percept Mot Skills 1977;44:367-73.
37. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56-62.
38. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 1991;14:540-5.
39. Hoddes E, Dement W, Zarcone V. The development and use of the Stanford Sleepiness Scale (SSS). Psychophysiology 1972;9:150.
40. Walsleben JA, Kapur VK, Newman AB, et al. Sleep and reported daytime sleepiness in normal subjects: the Sleep Heart Health Study. Sleep 2004;27:293-8.
41. Samkoff JS, Jacques CH. A review of studies concerning effects of sleep deprivation and fatigue on residents' performance. Acad Med 1991;66:687-93.
42. Durmer JS, Dinges DF. Neurocognitive consequences of sleep deprivation. Semin Neurol 2005;25:117-29.
43. Castronovo V, Canessa N, Strambi LF, et al. Brain activation changes before and after PAP treatment in obstructive sleep apnea. Sleep 2009;32:1161-72.
44. Aloia MS, Sweet LH, Jerskey BA, Zimmerman M, Arnedt JT, Millman RP. Treatment effects on brain activity during a working memory task in obstructive sleep apnea. J Sleep Res 2009;18:404-10.
45. Eta-squared, http://en.wikiversity.org/wiki/Eta-squared
46. John U, Meyer C, Hanke M, Volzke H, Schumann A. Smoking status, obesity, and hypertension in a general population sample: a cross-sectional study. Q J Med 2006;99:407-15.
47. Gozal D, Daniel JM, Dohanich GP. Behavioral and anatomical correlates of chronic episodic hypoxia during sleep in the rat. J Neurosci 2001;21:2442-50.
48. Verstraeten E. Neurocognitive effects of obstructive sleep apnea syndrome. Curr Neurol Neurosci Rep 2007;7:161-6.
49. Alchanatis M, Zias N, Deligiorgis N, Amfilochiou A, Dionellis G, Orphanidou D. Sleep apnea-related cognitive deficits and intelligence: an implication of cognitive reserve theory. J Sleep Res 2005;14:69-75.
50. Saunamaki T, Jehkonen M. A review of executive functions in obstructive sleep apnea syndrome. Acta Neurol Scand 2007;115:1-11.

Appendix 1—Neurocognitive testing schedule

TIME	Other	MWT	Neurocognitive Test Battery							
			SWMT	BSRT	SDC/DR	PF N/C	VSC	SAT	PVT	PASAT
NNC Session 1										
9:30–10:00 am	Qs									
10:00–10:25 am		X								
10:25–10:50 am			X							
10:50–11:05 am				X						
11:05–11:35 am	Psych 1									
11:35–11:40 am				recall						
Break (20- in lab)										
SNC Session 2										
12:00–12:25 pm		X								
12:25–12:50 pm			X							
Lunch Break (25)										
1:15–1:50 pm					X	X	X	X		
Break (10- in lab)										
SNC Session 3										
2:00–2:25 pm		X								
2:25–2:50 pm			X							
2:50–3:10 pm	SAQLI									
3:10–3:25 pm									X	
3:25–3:45 pm										X
Break (15- in lab)										
SNC Session 4										
4:00–4:25 pm		X								
4:25–4:30 pm	SIs									
4:30–4:45 pm	Psych 2									

Other: Qs = various self-administered questionnaires (BDI, POMS, QWB-SA. ESS); Psych 1 = administration of Hamilton Rating Scale for Depression and Mini International Neuropsychiatric Interview modules; SAQLI = Calgary Sleep Apnea Quality of Life Index; Sis = Administration of standardized instructions to subject; Psych 2 = finish administration of psychological interviews if not completed during Psych 1; MWT = maintenance of wakefulness test; SWMT = Sustained Working Memory Test; BSRT = Buschke Verbal Selective Reminding Test (plus 30-min recall); SDC/DR = Symbol Digit Coding (plus 20-30-min recall [SDCDR]); PF N/C = Pathfinder Number/ Pathfinder Combined; VSC = Visual Sequence Comparison; SAT = Shifting Attention Test; PVT = Psychomotor Vigilance Task; PASAT = Paced Auditory Serial Addition Test. All CogScreen tests are administered between 1:15–1:50 pm as one computerized module (SDC, SDCDR, PF N/C, VSC, and SAT).

Each subject has 70 minutes of "break time" incorporated into his or her day of neurocognitive testing.

Appendix 2—Additional assessments and measures included in APPLES

Tests of Attention and Psychomotor Function
- Symbol Digit Coding (CogScreen computer analogue of the Digit Symbol Substitution Test)
- Visual Sequence Comparison (Component of APPLES CogScreen package)
- Shifting Attention Test Instruction Condition (Component of APPLES CogScreen package)

Tests of Learning and Memory
- Symbol Digit Coding Delayed Recall Task

Tests of Executive and Frontal-Lobe Function
- Pathfinder Combined (CogScreen computer analogue of Trails Making B)
- Shifting Attention Test Discovery Condition (Component of APPLES CogScreen package)

Psychological Mood and Quality of Life Assessment
- Profile of Mood States
- Beck Depression Inventory
- Mini International Neuropsychiatric Interview
- Hamilton Rating Scale for Depression
- Calgary Sleep Apnea Quality of Life Index
- Quality of Well-Being Scale, Self-Administered

Sleep Assessments
- Fatigue Scale: Levels of Alertness
- Morning Questionnaire
- Bedtime Questionnaire
- Morningness-Eveningness Questionnaire (Derived from Horne and Ostberg)
- Sleep Habits Questionnaire (Modified from the Sleep Heart Health Study)
- Sleep Log

Other Measures (On Initial Screening Only)
- Clinical Screening Questionnaire
- Mini Mental State Examination
- Judgement of Line Orientation