Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Gottlieb DJ, Punjabi NM, Mehra R, et al. CPAP versus oxygen in obstructive sleep apnea. N Engl J Med 2014;370:2276-85. DOI: 10.1056/NEJMoa1306766

SUPPLEMENTARY APPENDIX

TABLE OF CONTENTS

Methods

- Eligibility — Page 2
- Sleep studies — Page 4
- Group assignment — Page 5
- Interventions — Page 5
- Outcomes — Page 7
- Sample size and power — Page 9
- Multiple imputation — Page 9

Outcomes

- Adverse events — Page 10

Supplementary Tables

- Table S1. Baseline biochemical measures — Page 12
- Table S2. Effect of treatment on biochemical measures — Page 13

References — Page 14

INVESTIGATORS

Daniel J. Gottlieb, Naresh M. Punjabi, Reena Mehra, Sanjay R. Patel, Stuart F. Quan, Denise C. Babineau, Russell P. Tracy, Michael Rueschman, Roger S. Blumenthal, Eldrin F. Lewis, Deepak L. Bhatt, Susan Redline

METHODS

Eligibility

Subjects were recruited from cardiology clinical practices at the four participating medical centers: Brigham and Women's Hospital, Boston, MA; Case Western Reserve University, Cleveland, OH; Johns Hopkins University, Baltimore, MD; and VA Boston Healthcare System, Boston, MA. Eligibility prior to enrollment was based on data available from a screening questionnaire and the medical record. After providing informed consent, eligible subjects underwent home sleep testing to determine eligibility for randomization into the clinical trial.

Inclusion criteria for enrollment:

- age 45-75
- Berlin Questionnaire score[1] of 2 or 3
- established coronary heart disease (prior myocardial infarction or coronary artery revascularization or angiographically documented >70% stenosis of a major coronary artery) **or** 3 or more of the following established cardiovascular risk factors:
    - hypertension (SBP >140 or DBP >90 or use of antihypertensive medication)
    - diabetes mellitus
    - obesity (BMI >30 kg/m^2)
    - dyslipidemia (total cholesterol >240 mg/dl, LDL >160 mg/dl, HDL <45 mg/dl, or taking lipid-lowering medication).

Exclusion criteria:

- diagnosed heart failure with left ventricular ejection fraction <35% or New York Heart Association Class ≥2
- poorly controlled hypertension (SBP >170 or DBP >110)

- poorly controlled diabetes (HbA1c >9.0%)
- myocardial infarction, stroke or coronary revascularization procedure within 3 months
- resting oxyhemoglobin saturation <90%
- severe chronic insomnia with reported usual sleep duration <4 hours per night
- severe sleepiness with an Epworth Sleepiness Scale[2] score ≥16 or report of falling asleep while driving within the previous 2 years
- pregnancy or a plan to become pregnant within 6 months
- smoking in the bedroom by the participant or anyone sharing a bedroom with the participant
- current use of supplemental oxygen
- current or past use of a positive airway pressure device or surgery for treatment of sleep apnea
- any uncontrolled medical problem that the investigator felt would significantly impair ability to participate in the study examinations
- inability or unwillingness to provide informed consent.

Additional criteria for randomization, based on home sleep testing:

- AHI ≥15 events per hour
- None of the following sleep study-based exclusion criteria:
    - AHI >50 events per hour
    - oxyhemoglobin saturation <85% for >10% of the recording
    - central apnea index >5 events per hour

Sleep studies

Enrolled subjects underwent home sleep testing with the Embletta Gold portable monitor (Embla Systems, Broomfield, CO) to determine the presence of obstructive sleep apnea. Randomized subjects had the study repeated after 12 weeks of intervention while using the assigned therapy. Subjects wore the device for a single night in their own homes and returned the device by express mail or courier service. The montage collected included airflow measured by both a nasal cannula-pressure transducer system and a thermal sensor, thoracic and abdominal movement by inductance plethysmography, finger pulse oximetry, body position, and a 3-lead electrocardiogram. Data were downloaded at the local sites and transferred to a central Sleep Reading Center at Case Western Reserve University for scoring by a single certified scorer. For the baseline examination, respiratory events were scored in accordance with American Academy of Sleep Medicine guidelines.[3] Apneas were defined as a >90% decrease from baseline in flow as measured by the thermal sensor, lasting at least 10 seconds, and were classified as central if there was no respiratory effort noted on plethysmography bands, and as obstructive if respiratory effort was present during the apnea. Hypopneas were defined as a >50% decrease from baseline in flow as measured by the nasal cannula-pressure transducer or thermal sensor, associated with at least a 3% fall in oxyhemoglobin saturation. The nasal cannula-pressure transducer system was used as the primary signal for identification of hypopneas, with the thermal sensor as an alternative sensor if the nasal cannula signal was missing.

While home sleep testing is an accepted modality for diagnosis of obstructive sleep apnea, the lack of electroencephalographic measures of sleep and arousal will result in false-negative studies in individuals whose respiratory events are accompanied by arousal but not desaturation,

or who spend a large portion of the recording awake (underestimating the frequency of events by including awake time in the denominator). Although this is of considerable clinical diagnostic importance, it does not affect the validity of the study results, since the resultant underestimation of sleep apnea severity will not lead to inclusion in the study of subjects without sleep apnea; however, some caution must be taken in extrapolating to individuals with obstructive sleep apnea whose hypopneas are not accompanied by desaturation. Moreover, the lack of sleep measures does limit our ability to evaluate potential mediation of the CPAP effect on blood pressure by changes in arousal frequency or other sleep quality measures.

Group assignment

After the baseline evaluation, subjects were randomly assigned using a stratified permuted block design to one of three arms: healthy lifestyle and sleep education alone (control), or healthy lifestyle and sleep education in addition to continuous positive airway pressure (CPAP) or nocturnal supplemental oxygen. Randomization was performed centrally using a web-based system and stratified by recruitment site and by the presence of coronary heart disease. Blinding of subjects to treatment arm was not possible. A study coordinator at each clinical site was also aware of group assignment and instructed participants in use of their assigned treatment; however, clinical site staff performing testing, and coordinating center staff editing and scoring the reactive hyperemia and blood pressure recordings, were blinded to group assignment.

Interventions

All participants received standardized healthy sleep and lifestyle education on one occasion, at the time of the baseline visit. The education was provided face-to-face by research staff, using a

slide presentation developed for this study; a copy of the slides was also provided to participants in paper or electronic format. This was supplemented by paper copies of educational materials publicly available from the American Heart Association. The education in healthy sleep habits included suggestions for maintaining a regular sleep schedule, avoiding alcohol near bedtime, and maintaining sleep duration of 7-8 hours per night. Subjects were also provided with education on a heart-healthy lifestyle, including recommendations for weight loss (for overweight and obese subjects), healthy diet, regular exercise, smoking cessation, and medication adherence.

Subjects randomized to CPAP were provided, in addition to lifestyle counseling, an Autopap REMstar device (Philips-Respironics, Inc, Murrysville, PA) with integrated humidifier. Mask fit was optimized by an experienced technician. The CPAP device was initially set at a pressure range of 4-20 cm H_2O, with auto-titration according to the device's algorithm for detecting airflow limitation and wireless transmission of usage information to the Data Coordination Center. At the end of a 7-day titration period, the Data Coordination Center identified the optimal pressure, defined as the 90^{th} percentile pressure level needed to eliminate airflow limitation events, and wirelessly reprogrammed the device to deliver this fixed pressure. Objective adherence data were obtained from the CPAP device.

Subjects randomized to supplemental oxygen received, in addition to lifestyle counseling, nightly treatment with oxygen at 2 liters/min via nasal cannula, using a stationary oxygen concentrator (EverFlo, Philips-Respironics, Murrysville, PA). This flow rate of supplemental oxygen was previously shown to reduce the number of 4% desaturations by approximately 90%.[4] Each oxygen concentrator contained a meter recording cumulative hours of use, providing an objective measure of adherence over the intervention period.

Outcomes

Participants were contacted at regular intervals to assess adverse events. At baseline and after 12 weeks, after an overnight fast of at least 12 hours, venipuncture and anthropometry were performed and endothelial function and resting blood pressure were measured. Within several days of these visits but on different days, subjects also underwent 24-blood pressure monitoring and, following the 12-week visit, home sleep testing on the assigned therapy.

Blood pressure was measured during a single 24-hour period using the Spacelabs 90207 Ambulatory Blood Pressure monitor (Spacelabs Medical Inc., Deerfield, WI). The monitor was programmed to measure blood pressure every 20 minutes between 6 AM and 10 PM, and every 30 minutes between 10 PM and 6 AM. A diary was used to record bed and rise times on the day of the recording. All evaluable blood pressures recorded during the reported awake (or asleep) period were averaged to obtain mean daytime (or nighttime) values. At least 10 valid daytime and 4 valid nighttime readings were required. Mean pressure was calculated at each reading as 1/3 systolic plus 2/3 diastolic pressure, and the 24-hour mean pressure calculated as a weighted average of the daytime and nighttime mean values, with weights determined by the percentage of time reportedly spent in each state as determined by a sleep diary. Nocturnal non-dipping blood pressure was defined as a mean nocturnal blood pressure >90% of the daytime value.

Following venipuncture, blood was processed locally and stored at -80˚C before shipment on dry ice to the Biochemistry Core at the Laboratory of Clinical Biochemistry Research (LCBR), University of Vermont, where the biochemical assays were performed. Serum glucose and lipid panel were measured using a colorimetric reflectance spectrophotometric method on the Vitros 950IRC instrument (Johnson & Johnson Clinical Diagnostics, Rochester, NY), at the Clinical Chemistry Laboratory at Fletcher Allen Health Care. Serum insulin and amino-terminal pro-B-

type natriuretic peptide (BNP) were measured using a using the Roche Elecsys 2010 analyzer (Roche Diagnostics, Indianopolis, IN) which utilizes an electrochemiluminescence immunoassay method. Serum high-sensitivity C-reactive protein (CRP) was measured using the BNII nephelometer (Siemens, Inc., Deerfield, IL) utilizing a particle enhanced immunonepholometric assay. For subjects in whom these measurements were below the lower limit of detection, the minimum value observed in all subjects was used in its place.

While still fasting and after 5 minutes of rest in a quiet environment, endothelial function was determined as the reactive hyperemia index measured using the Endo-PAT device (Itamar Medical, Caesarea, Israel), as previously described[5,6]. Briefly, disposable tonometry probes are placed on the index finger of each hand and a blood pressure cuff is placed on the non-dominant upper arm. Following a 5-minute baseline recording, the cuff is inflated to 60 mmHg above systolic blood pressure, with a minimum pressure of 200 mmHg and a maximum of 300 mmHg. After 5 minutes of occlusion, the cuff pressure is rapidly released and arteriolar pulse volume recorded from both index fingers. Reactive hyperemia index is calculated as the average increase from baseline in pulse volume measured in the previously occluded hand between 90 and 120 seconds following cuff deflation, adjusted for changes from baseline noted in the non-occluded hand. Additional measures included height, measured with the subject in stocking feet using a wall-mounted stadiometer, and weight, measured with a calibrated digital scale.

A Data Coordination and Sleep Reading Center at Case Western Reserve University and Brigham and Women's Hospital regularly monitored data and provided central scoring and processing of data blinded to intervention assignment. Data files from the Endo-PAT, ambulatory blood pressure, portable sleep monitoring, and CPAP devices were centrally edited, scored, and monitored for quality at the Coordinating Center.

Sample size and power

A planned sample size of 118 subjects per group was justified by conservatively assuming that 12-week follow-up data would only be available for 85% of subjects and that the coefficient of variation for each endpoint at 12 weeks would be either 0.1 or 0.5. Under these assumptions, a sample size of 100 subjects per arm was needed to detect a 3.9% to 17.1% decrease in each endpoint respectively with 80% power, based on a family-wise error rate of 0.15, using a two-sample t-test and a two-sided Bonferroni-corrected significance level of 0.05 for each pairwise comparison. This significance level was chosen *a priori* to permit the detection of reasonable effect sizes with a sample that could be feasibly recruited within the time frame of the 2-year grant that funded this study, and was approved by the external Data and Safety Monitoring Board prior to the start of recruitment. Additionally, an interim analysis of all blood pressure endpoints was performed in December 2010 when 121 subjects had completed the 12-week visit. To maintain a Type I error rate of 0.05 for each pairwise comparison of all blood pressure endpoints, the significance level for the interim and final analyses was set to 0.0008 and 0.0497 respectively by applying the Lan-DeMets method for group sequential trials using an O'Brien-Fleming spending function.[7]

Multiple imputation

In order to estimate the impact of missing data on the primary outcome measure, 24-hour mean arterial pressure, the analysis was repeated using multiple imputation of missing values. Multiple imputation was performed under a fully conditional specification using a Markov Chain Monte Carlo method implemented in the R library, mice.[8] Imputations for continuous data were drawn from a normal linear model using the standard non-informative prior while imputations

for binary data were drawn from a logistic regression model. Variables used for imputation included those in both the primary and secondary analyses, given by 24-hour MAP at baseline and at 12 weeks, treatment arm, site, CAD, age, race, sex, BMI, AHI, and use of ACE inhibitors, beta blockers, alpha blockers, diuretics, and calcium blockers and the time between baseline and 12-week follow-up. A total of 15 multiple imputations was selected because a general rule of thumb is to select the number of imputations using the average missing data rate; in the primary and secondary analyses for 24-hour MAP at 12 weeks, 11.6% and 12.9% of subjects were removed from the primary and secondary analyses, respectively, due to missing data.

OUTCOMES

Adverse events

The 90-day incidence rates of non-serious and serious adverse events following randomization in each treatment arm were calculated by dividing the total number of events in each treatment arm by the total number of days at risk among all subjects randomized to that treatment arm and then multiplying this quantity by 90. An exact 95% CI based on a Poisson distribution was also calculated for each rate. The 90-day incidence rate of non-serious adverse events was 0.31 (95% CI, 0.20 to 0.36) in the control group, 0.23 (95% CI, 0.14 to 0.35) in the CPAP group, and 0.29 (95% CI, 0.19 to 0.43) in the supplemental oxygen group. The respective 90-day incidence rates of serious adverse events were 0.11 (95% CI, 0.049 to 0.20), 0.046 (95% CI, 0.012 to 0.12), and 0.084 (95% CI, 0.034 to 0.17). There were no deaths. There was one episode of unstable angina, one myocardial infarction, one percutaneous coronary intervention for worsening angina, and one stroke, all in the control group. There were three episodes of atrial fibrillation, two in the supplemental oxygen group and one in the CPAP group, and a single episode of an

unspecified tachyarrhythmia requiring hospitalization, occurring in the control group. There was a single motor vehicle accident, occurring in the supplemental oxygen group.

Table S1. Baseline biochemical characteristics of the study sample[1]			
Characteristic	**HLSE**	**CPAP**	**NSO**
Glucose[2], mg/dL	100.1 (13.3) N= 59	101.4 (17.2) N= 62	102.5 (29.9) N= 61
Insulin[2,3], µU/ml	17.3 (12.3) N= 59	19.6 (15.2) N= 62	19.3 (12.2) N= 61
HOMA-IR[2,4]	4.4 (3.8) N= 59	5.2 (4.9) N= 62	5.1 (4.6) N= 61
Total cholesterol, mg/dl	163.3 (37.2) N= 101	166.2 (31.9) N= 98	168.0 (38.3) N= 99
LDL cholesterol, mg/dl	90.1 (25.5) N= 95	92.6 (25.3) N= 95	91.6 (31.8) N= 95
HDL cholesterol, mg/dl	43.6 (12.8) N= 101	44.1 (13.1) N= 98	44.5 (11.4) N= 99
Triglycerides, mg/dl	146.0 (90.6) N= 101	151.1 (78.3) N= 98	159.3 (98.9) N= 99
BNP[5], pg/ml	209.2 (337.6) N= 101	202.2 (315.6) N= 98	203.9 (385.8) N= 99
C-reactive protein, µg/ml	4.0 (5.1) N= 100	3.6 (5.8) N= 98	4.0 (5.3) N= 99

[1] Values are mean (SD).

[2] Subjects that took diabetic medications at baseline are excluded.

[3] Two subjects with non-detectable insulin at baseline were assigned values of 0.2 µU/ml.

[4] HOMA-IR calculated using the following formula: glucose*insulin/405.

[5] Four subjects with non-detectable NT-pro-BNP at baseline were assigned values of 5 pg/ml.

Abbreviations: HLSE = healthy lifestyle and sleep education; CPAP = continuous positive airway pressure; NSO = nocturnal supplemental oxygen; HOMA-IR = homeostatic model assessment-insulin resistance; LDL = low-density lipoprotein; HDL = high-density lipoprotein; BNP = amino-terminal pro-B-type natriuretic peptide; hs-CRP = high-sensitivity C-reactive protein

Table S2. Effect of treatment on biochemical measures[1,2]

Response	CPAP vs. HLSE	NSO vs. HLSE	CPAP vs. NSO
Glucose	0.99 (0.95, 1.03)	1.00 (0.97, 1.04)	0.99 (0.95, 1.03)
Insulin	0.91 (0.79, 1.06)	0.93 (0.80, 1.08)	0.98 (0.85, 1.13)
HOMA-IR[3]	0.90 (0.76, 1.07)	0.93 (0.79, 1.10)	0.97 (0.82, 1.14)
Total cholesterol	1.02 (0.98, 1.05)	1.00 (0.96, 1.03)	1.02 (0.98, 1.05)
LDL cholesterol	1.00 (0.95, 1.05)	0.98 (0.93, 1.03)	1.02 (0.97, 1.07)
HDL cholesterol	1.03 (0.99, 1.07)	1.01 (0.98, 1.05)	1.01 (0.98, 1.05)
Triglycerides	1.06 (0.97, 1.16)	1.03 (0.94, 1.13)	1.03 (0.94, 1.13)
BNP	0.93 (0.78, 1.12)	1.01 (0.84, 1.20)	0.93 (0.78, 1.10)
C-reactive protein	0.80 (0.65, 0.97)[4]	0.91 (0.74, 1.11)	0.88 (0.72, 1.07)

[1] Geometric mean ratio of measures at 12 weeks, for all evaluable subjects, adjusted for study site, presence of coronary artery disease, and baseline biochemical measure. Values are given as adjusted geometric mean ratio between arms (95% CI).

[2] Sample size for glucose, insulin and HOMA-IR: HLSE 59, CPAP 62, NSO 61, as those on glucose lowering medications were excluded; for total and HDL cholesterol, triglycerides and BNP: HLSE 101, CPAP 98, NSO 99; for LDL cholesterol: HLSE 95, CPAP 95, NSO 95; for C-reactive protein: HLSE 100, CPAP 98, NSO 99.

[3] HOMA-IR calculated using the following formula: glucose*insulin/405.

[4]For comparison of CPAP vs. HLSE, p=0.026; for all other comparisons p >0.05.

Abbreviations: see Table S1.

REFERENCES

1. Netzer NC, Stoohs RA, Netzer CM, Clark K, Strohl KP. Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. Ann Intern Med 1999;131:485-91.

2. Johns MW. A new method for measuring daytime sleepiness: The Epworth Sleepiness Scale. Sleep 1991;14:540-5.

3. Iber C, Ancoli-Israel S, Chesson A, Quan SF. American Academy of Sleep Medicine. The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications, 2nd ed. Westchester, IL: American Academy of Sleep Medicine; 2007.

4. Teramoto S, Kume H, Yamamoto H, et al. Effects of oxygen administration on the circulating vascular endothelial growth factor (VEGF) levels in patients with obstructive sleep apnea syndrome. Intern Med 2003;42:681-5.

5. Hamburg NM, Keyes MJ, Larson MG, et al. Cross-sectional relations of digital vascular function to cardiovascular risk factors in the Framingham Heart Study. Circulation 2008;117:2467-74.

6. Seif F, Patel SR, Walia H, et al. Association between obstructive sleep apnoea severity and endothelial dysfunction in an increased background of cardiovascular burden. J Sleep Res 2013;doi:10.1111/jsr.12026.

7. DeMets DL, Lan KK. Interim analysis: The alpha spending function approach. Stat Med 1994;13:1341-52.

van Buuren S, Groothuis-Oudshoorn K. Mice: Multivariate Imputation by Chained Equations in R. J Stat Softw 2011;45:1-67.