ONLINE SUPPLEMENT

Obstructive Sleep Apnea Treatment and Fasting Lipids: A Comparative Effectiveness Study

Brendan T. Keenan, MS[1*], Greg Maislin, MS, MA[1,2*], Bernie Y. Sunwoo, BSc, MB, BS[1,2,3*], Erna Sif Arnardottir, PhD[4,5], Nicholas Jackson, MPH[1], Isleifur Olafsson, MD, PhD[6], Sigurdur Juliusson, MD, PhD[5], Richard Schwab, MD[1,2], Thorarinn Gislason, MD, PhD[4,5], Bryndis Benedikstdottir, MD[4,5**], Allan I. Pack, MBChB, PhD[1,2**]

[*]Co-First authors; [**]Co-senior authors; [1]Center for Sleep and Circadian Neurobiology, University of Pennsylvania, Philadelphia, PA, USA; [2]Divison of Sleep Medicine, Department of Medicine, University of Pennsylvania, Philadelphia, PA, USA; [3]Division of Pulmonary, Allergy and Critical Care, Department of Medicine, University of Pennsylvania, Philadelphia, PA, USA; [4]Department of Respiratory Medicine and Sleep, Landspitali - The National University Hospital of Iceland, Reykjavik, Iceland; [5]Faculty of Medicine, University of Iceland, Reykjavik, Iceland; [6]Department of Clinical Biochemistry, The National University Hospital of Iceland, Reykjavik, Iceland

SUPPLEMENTARY METHODS

Study Subjects

Analyses included patients diagnosed with moderate to severe (apnea hypopnea index [AHI] ≥ 15) obstructive sleep apnea (OSA) at one of the five sites performing sleep studies in Iceland and referred for positive airway pressure (PAP) treatment to the Landspitali University Hospital of Iceland in Reykjavik (the only site in Iceland providing PAP treatment) from September 2005 to December 2009 (see **Figure E1**). Over ninety percent of subjects approached agreed to participate, resulting in an initial sample of 822 patients comprising the prospective Icelandic Sleep Apnea Cohort (ISAC). For further details about diagnosis of OSA, see previous publications.[1-5] Participants completed standardized questionnaires, physical examination, a type 3 sleep study and abdominal magnetic resonance imaging (MRI) at baseline while untreated. Two years after treatment initiation, participants were invited for a follow up visit, where treatment adherence was examined and baseline assessments, apart from abdominal MRI, were repeated. Written consent was obtained from every participant, and the study protocol was approved by the National Bioethics Committee, the Data Protection Authority of Iceland and the Institutional Review Board (IRB) of the University of Pennsylvania.

Of the 822 patients initially recruited, 806 (98%) had non-missing OSA severity and lipid measures and 193 (24%) were using ATC defined lipid-lowering medications, leaving 613 (76%) patients included in our primary analyses at baseline (**Figure E1**). Comparisons between patients using and not using lipid-lowering medications are presented below (see **Table E1**). Of those 613 patients included in baseline analyses, 552 (90%) returned for a follow-up visit.

Limiting this sample to those patients that remained off lipid-lowering medications and had non-missing PAP adherence and lipid measures resulted in 491 patients. To assure that the non-users accurately represented the untreated OSA population, we removed those who were prescribed a mandibular advancement device (n=24), resulting in an analysis sample for change in lipids of 467 patients with outcome measures and adherence data: 240 adherent users, 71 partial-users, and 156 non-users. Patients determined to be "partial-users", based on the classification criteria presented below, were excluded from analyses examining the effect of PAP treatment. We then used sub-classification by propensity scores in order to select a subsample of 199 adherent and 118 non-users with minimal selection bias and balanced covariate distributions, in which we assessed the effect of PAP treatment on lipid levels.

Baseline Evaluation

Participants completed standardized questionnaires, physical examination, a type 3 sleep study, fasting morning blood samples and abdominal MRI. The standardized questionnaires were administered by trained interviewers and included: i) Demographics; ii) Medical history including a doctor diagnosis of hypertension, coronary artery disease, heart failure, stroke and diabetes mellitus; iii) Sleep history, including the Basic Nordic Sleep Questionnaire [6]; iv) Epworth Sleepiness Scale (ESS) [7]; v) Current medications coded according to the ATC drug classification system (WHO Collaborating Centre for Drug Statistics Methodology); and vi) Lifestyle habits, including smoking history, alcohol consumption and participation in exercise. Participants also underwent physical examination and standardised anthropometric measurements (height, weight, neck and waist circumferences).

OSA Severity Measures

All subjects had a type 3 sleep study prior to referral for PAP treatment. Sleep study recordings were later re-scored in a uniform manner at the Sleep Study Reading Unit of the University of Pennsylvania. The following measures of OSA severity were calculated: i) Apnea-hypopnea index (AHI), defined as the average number of apneas (defined as a ≥80% decrease in flow for ≥10 seconds) and hypopneas (defined as a ≥30% decrease in flow with a ≥4% oxygen desaturation *or* ≥50% decrease in flow for ≥10 seconds with a sudden increase in flow at the end of the event) per hour of recording; ii) Oxygen desaturation index (ODI), defined as the average number of oxygen desaturations ≥4% per hour of recording; iii) Oxygen saturation nadir (SaO_2 nadir); and iv) Percent of recording time with oxygen saturation below 90% (percent time SaO_2<90).

Blood Samples

Fasting morning blood samples were taken and serum stored at -20°C. Total and high-density lipoprotein cholesterol and triglycerides concentrations were measured using a Vitros 950 analyser (Ortho Clinical Diagnostics, Rochester, NY, USA) and the manufacturer's multilayer film dry-slide chemistry reagents and calibrators. The total analytical imprecisions, measured as coefficient of variation (CV%), for the measurements of total TC, TG and HDL-C were 2.5%, 1.6% and 3.2% respectively. LDL-C was calculated using the Friedwald equation ($LDL\text{-}C = TC - HDL\text{-}C - TG/5$). In addition to the continuous measures of lipoprotein levels, we also created

binary variables representing whether an individual fell within the "abnormal" range, based on the National Cholesterol Education Program (NCEP) Adult Treatment Panel III (ATPIII) criteria.[8] Abnormal levels were defined as greater than or equal to "borderline high" levels for TC (≥200 mg/dL), LDL-C (≥130 mg/dL) and TG (≥150 mg/dL) and less than "normal" levels for HDL-C (<40 mg/dL).[8]

Magnetic Resonance Imaging (MRI)

Abdominal MRI were performed using a 1.5T scanner with a body coil (Siemens Avanto, Germany) and images were manually analyzed in a uniform manner at the Sleep Imaging Center of the University of Pennsylvania using image analysis software (Amira 4.1.2, Mercury Computer Systems, Chelmsford, MA). Visceral (VAT) and subcutaneous (SAT) fat volumes for the abdominal compartment were quantified in cm^3 and the total abdominal fat volume calculated as the sum of VAT and SAT. Measurement reliability of volumes was assessed based on intraclass correlation coefficient (ICC) analysis. ICC values for VAT and SAT volumes were both approximately 1.0, which indicates minimal technical variability arising from differences among raters or from rater by subject interaction. MRI data was available on a subset of the baseline sample not on lipid lowering medications (n=501, see **Table E2**); reasons for failure included claustrophobia (n=66), poor picture quality (n=33), very high obesity (n=6) and nonspecific (n=7). Additional information is available in previous publications.[5]

Follow-up Evaluation and PAP Adherence

Two years after PAP treatment initiation, participants were invited for a follow-up visit where they answered the same questionnaires, underwent physical examination and anthropometric measures and had fasting morning blood samples drawn, as in baseline assessment. The mean ± SD time between baseline and follow-up visit was 774 ± 135 days. Altogether, 90.1% of the sample took part in the follow up (n=741, 596 [80%] males and 145 [20%] females).

PAP adherence at follow-up was estimated based on downloads from memory cards of mean hours and total nights of usage over the last 28 days (objective data), if available, from ResMed S8 machines (ResMed Corp. San Diego, CA, USA). Some subjects had older PAP devices which did not allow for this type of download. In these subjects, PAP adherence was determined based on subjective questionnaires. Adherent PAP users were identified as PAP use $\geq$4 hours per night and $\geq$20 out of the last 28 nights by memory card download or $\geq$60% of the night and $\geq$5 nights per week by subjective questionnaire. Partial users used PAP <4 hours per night or <20 out of the last 28 nights by memory card download or <60% of the night or <5 nights per week by subjective questionnaire. To validate the use of self-reported data where objective download data was not available, among the 355 subjects with both objective (memory cards) and self-reported data on frequency of PAP use, we compared the cut-off for objective adherent use vs. partial use to the comparable subjective cut-offs. Self-report had 98.6% sensitivity and 45.1% specificity in distinguishing adherent versus partial users. Partial users were excluded from final analyses examining the impact of adequate PAP use on changes in lipids. PAP non-users were defined as patients that had no objective use in the last 28 days, reported no current PAP usage, or had returned their device. Non-users who were prescribed a mandibular advancement device

were excluded from our analyses to assure that this group represented the untreated OSA population.

Statistical Analysis

Statistical analyses were performed using Stata Version 12, StataCorp (College Station, Texas) or SAS Software, version 9.3 (SAS Institute, Cary, NC).

Demographics

Continuous characteristics are summarized using means and standard deviations and categorical covariates using frequencies and percentages. Demographics at baseline were compared among subgroups using T-tests or analysis of variance (ANOVA), depending on the number of groups, and categorical covariates using chi-square or Fisher's exact tests. Analyses comparing demographic variables at follow-up were performed using an analysis of covariance (ANCOVA) or conditional logistic regression, adjusted for propensity score subclass and baseline covariate level.

Baseline Lipid Analyses

The primary outcomes at baseline were natural log transformed levels of lipids, including total cholesterol, LDL cholesterol, HDL cholesterol, and triglycerides. Outcomes were natural log transformed in order to achieve normality and allow for parametric analyses. Linear associations

between lipid levels and obesity or OSA severity measures were assessed using Pearson correlations and partial correlations (see **Table E3** for unadjusted results, **Table 2** for adjusted correlations). Partial correlations were adjusted for age, gender and body mass index (for OSA severity measures only).

Propensity Score Designed Observational Cohort

In order to assess the impact of PAP treatment on two-year changes in lipid levels, we created a purposefully designed observational study based on the propensity score (PS) distributions in positive airway pressure adherent patients and non-users, using an established sequential heuristic, presented and discussed in detail in previous literature [9-19], particularly Maislin & Rubin [11]. The resulting sample contains participants divided into 'propensity score quintiles', within each of which there is a relative balance in the observed covariates. This balance with respect to measured covariates is similar to what would be expected had patients been randomized, allowing for causal inferences to be made from the non-randomized group comparisons. As discussed in the manuscript, one important limitation of the propensity score methodology is the inability to address unmeasured confounders, which should be balanced in a true randomized controlled trial. To mitigate the impact of unmeasured confounders, we included as many relevant measured variables as possible, as unmeasured variables are controlled for to the extent that they are correlated with included variables.

As discussed above, our original observational cohort consisted of 240 adherent PAP users and 156 non-users that were not taking lipid-lowering medications. Propensity scores for our

purposefully designed observational study were based on a number of important covariates at baseline, including: age, gender, BMI, current smoking status, hypertension, cardiovascular disease, diabetes, participation in exercise, excessive alcohol use, Epworth Sleepiness Scale (ESS), sleep medication use, and OSA severity (apnea-hypopnea index, oxygen desaturation index, SaO_2 nadir, and percentage of sleep time with $SaO_2<90$). We also controlled for baseline levels of our four lipid measures (total cholesterol, LDL cholesterol, HDL cholesterol and triglycerides), as these were strongly correlated with two-year changes. We used single imputation to fill in few missing values at baseline for ESS, smoking status, diabetes, exercise participation and excessive alcohol use; less than 2.3% of the sample was missing values for any one variable. After implementing the heuristic, we identified a subsample of 199 (83%) adherent and 118 (76%) non-users with sufficient propensity score overlap that met the model assumptions assuring covariate balance and allowing for causal inference. A comparison of patients included and excluded from the propensity score designed observational cohort is presented in **Table E4.**

In our purposefully designed observational study, there was a clear reduction in the bias due to covariate imbalance after adjusting for propensity score subclass (mean bias = -0.04, p=0.873) when compared to the unadjusted bias (mean bias = 0.63, p<0.0001); there was no significant group by subclass interaction in bias estimates (p=0.938). In addition, our resulting sample met the 'three basic distributional conditions' described by Rubin.[13] The fact that these conditions are met allow us to conclude that the measured covariate distributions are the same for the adherent and non-users, and allow inference to proceed as if patients had been randomized to

treatment. The resulting covariate balance is illustrated using a "Love Plot", as described by Ahmed and colleagues.[9]

Differences between changes in lipid levels in PAP adherent and non-users

We assessed differences in two-year lipid changes between PAP adherent patients and non-users included in the propensity score subclass matched sample described above. Subject specific two-year changes in lipid levels were calculated as within subject follow-up levels minus baseline levels. To examine whether there was an association between change in lipid levels and PAP adherence, we used an analysis of covariance (ANCOVA), comparing the predicted least square mean changes between adherent and non-users. Results are presented as predicted least square mean change ± standard error (SE). Models were adjusted for propensity score quintile subclass, which was included in the model as a 4 degree of freedom variable. Given the strong correlations observed between baseline lipid levels and the magnitude of two-year change, we additionally included baseline lipid levels in these models in order to control for residual differences within propensity score subclass. Additionally, we assessed whether there was a difference in the change in proportion of patients with "abnormal" fasting lipid levels between PAP adherent patients and non-users using a generalized estimating equations (GEE) model, adjusted for propensity score subclass. To determine whether any observed associations between PAP use and lipid level changes was moderated by obesity, we examined whether there was evidence for a significant interaction between PAP and BMI groups by including main effect (PAP usage and BMI group) and product ([PAP usage] x [BMI group]) terms in our regression models. Secondarily, given the limited power of interaction analyses, we performed analyses

identical to those described above in the overall sample within each BMI group, regardless of whether the interaction term was significant.

Subset Analyses

In addition to the primary analyses described above, we performed two additional exploratory analyses examining the effect of PAP usage on changes in lipids within relevant subsets of our propensity score subclass matched sample: 1) among patients with abnormal lipid levels at baseline and 2) limited to patients with the most severe hypoxia at baseline (defined as $\geq 75^{th}$ percentile of percent time $SaO_2<90$). Analyses were performed using similar methods to those described for the primary analyses.

Significance Level

Analyses for the associations between baseline lipoprotein levels and obesity or OSA severity measures were evaluated with a type I error rate of $\alpha=0.05$. Similarly, for analyses assessing the relationship between lipid change and PAP usage, a $p<0.05$ was considered evidence of a significant association.

Power Calculations

Our sample size for analysis provided adequate power to detect statistical significance ($p<0.05$) in our primary analyses. For analyses examining the correlations between obesity, OSA severity,

and lipid levels at baseline, we had >90% power to detect even a small effect size (Pearson ρ^2 = 0.02). Similarly, for our analyses examining the relationship between changes in lipid levels and PAP usage, we had >90% power to find an association if PAP adherence explained as small as 2.5% of the variance in lipid change, after accounting for the impact of propensity score subclass and baseline lipid level. Within our BMI defined strata, where the sample is reduced to approximately 100 participants, the power to find a small effect (R^2=0.025) was reduced to 46%, but we had 90% power to find a small to moderate effect size of 7% variance explained.

SUPPLEMENTAL RESULTS

Comparison of lipid-lowering medication users and non-users. A comparison of lipid lowering medication users and non-users is presented in **Table E1**. Primary analyses in the manuscript were restricted to the 613 non-users. Patients that reported lipid-lowering medication use were older, had more severe OSA (higher AHI and higher percentage time SaO_2<90%), and greater prevalence of hypertension, diabetes and cardiovascular disease compared non-users. Lipid-lowering medication users had similar BMI, HDL-C and TG levels as non-users, but lower TC and LDL-C.

Comparison of patients with and without abdominal MRI. Abdominal MRIs were available in 501 (82%) participants: 66 (59%) failed due to claustrophobia, 33 (29%) due to MRI picture quality, 6 (5%) were too obese, and 7 (6%) due to unspecified reasons. Subjects with no MRI data were slightly younger, more obese with greater OSA severity and had higher frequencies of hypertension, diabetes mellitus, and smoking (p<0.05) (**Table E2**). Significant differences in

lipid measures were also observed, as they had lower total, LDL, and HDL cholesterol, but higher triglyceride levels ($p<0.05$). Due to the more limited sample with available abdominal MRI measurements and that MRI was not repeated at follow-up, BMI was used as our primary measure of severity of obesity.

Unadjusted Correlations between lipid levels, obesity and OSA. Unadjusted associations with baseline lipid levels for clinical and MRI obesity measures and OSA severity are shown in **Table E3**. Correlations between baseline lipid levels and obesity measures adjusted for age and gender are shown in **Table 2** in the manuscript. We observed significant negative correlations with TC, LDL-C, and HDL-C and positive correlations with TG for all clinical obesity measures. All MRI measures were positively correlated with TG. Both total and visceral abdominal fat were negatively correlated with HDL-C. Prior to adjustment for age, gender and obesity, positive correlations were seen between OSA severity measures and TG. ODI also correlated with TC and LDL-C, while the percent time $SaO_2<90$ correlated with LDL-C. We note that the unadjusted correlations between OSA severity and lipid levels are relatively weak compared to those with obesity measures, and are no longer significant after adjustment for age, gender and BMI (see **Table 2** in manuscript).

Characteristics of participants in propensity score designed observational cohort. A comparison of patients included and excluded from the propensity score designed observational study is presented in **Table E4**. We see that, on average, excluded full users were older ($p<0.001$), heavier ($p=0.035$), had more severe OSA (all $p<0.0001$), were more likely to have hypertension ($p<0.0001$), cardiovascular disease ($p=0.049$), diabetes ($p=0.004$) and excessive

alcohol use (p=0.015), and were less likely to be smokers (p=0.045) when compared to full users that were included in the propensity score matched sample. On the other hand, compared to those included, excluded non-users had less subjective sleepiness (p=0.017), less severe OSA (all p<0.012), and were more likely to be female (p=0.001) and smokers (p=0.008). While there were no differences in lipid levels between excluded and included full-users, non-users that were excluded from the propensity score matched sample had higher baseline levels of total (p=0.015) and LDL (p=0.014) cholesterol at baseline.

Relationship between PAP usage and proportion with abnormal lipids. We examined the effect of PAP treatment on changes in the proportion of patients with abnormal lipid levels (**Table E5**). No significant differences were observed between PAP adherent and non-users in the overall sample. There was suggestive evidence of PAP by BMI group interaction for abnormal HDL-C (p=0.096), and within the BMI<30 strata, a greater decrease in the proportion of adherent patients with abnormal HDL-C compared to non-users (p=0.024). Given no difference in two-year HDL-C mean changes in this subgroup (**Table 4** in manuscript, both groups had an estimated 7 unit increase), we examined whether there were a higher proportion of adherent patients with baseline HDL of 33-40 mg/dL. While more adherent patients fell in this range, the difference was non-significant (p=0.588). We found no differences in the change in proportion abnormal between adherent and non-users within BMI strata for other lipid measures.

Differences in lipid changes between PAP adherent and non-users in patients abnormal at baseline. The results from our exploratory analysis on the effect of PAP adherence on lipid change within the subset of patients with abnormal lipid levels at baseline are presented in **Table**

E6. Within our propensity score designed observational study, we found no significant differences in two-year lipid level change between PAP adherent patients and non-users. Interestingly, there was a significant PAP by BMI interaction for LDL-C change (p=0.041). This interaction is likely driven by directional differences in the observed (non-significant) effect in the BMI 30-35 strata (non-users have larger decreases) compared to other BMI strata, where adherent patients have larger decreases.

Differences in lipid changes between PAP adherent and non-users in most hypoxic subset. We restricted our analysis sample to only those with the most severe hypoxia, defined as being in the top quartile of percent time SaO_2<90. Within this sample, no significant differences between PAP adherent patients and non-users in changes in lipid levels were observed (**Table E7**).

SUPPLEMENTAL REFERENCES

1. Arnardottir ES, Janson C, Bjornsdottir E, Benediktsdottir B, Juliusson S, Kuna ST, Pack AI, Gislason T. Nocturnal Sweating - a Common Symptom of Obstructive Sleep Apnea: The Icelandic Sleep Apnea Cohort. *BMJ Open* 2013: 3(5).
2. Arnardottir ES, Maislin G, Jackson N, Schwab RJ, Benediktsdottir B, Teff K, Juliusson S, Pack AI, Gislason T. The role of obesity, different fat compartments and sleep apnea severity in circulating leptin levels: the Icelandic Sleep Apnea Cohort study. *Int J Obes (London)* 2013: 37(6): 835-842.
3. Arnardottir ES, Maislin G, Schwab RJ, Staley B, Benediktsdottir B, Olafsson I, Juliusson S, Romer M, Gislason T, Pack AI. The interaction of obstructive sleep apnea and obesity on the inflammatory markers C-reactive protein and interleukin-6: the icelandic sleep apnea cohort. *Sleep* 2012: 35(7): 921-932.
4. Bjornsdottir E, Janson C, Gislason T, Sigurdsson JF, Pack AI, Gehrman P, Benediktsdottir B. Insomnia in untreated sleep apnea patients compared to controls. *J Sleep Res* 2012: 21(2): 131-138.
5. Maislin G, Ahmed MM, Gooneratne N, Thorne-Fitzgerald M, Kim C, Teff K, Arnardottir ES, Benediktsdottir B, Einarsdottir H, Juliusson S, Pack AI, Gislason T, Schwab RJ. Single slice vs. volumetric MR assessment of visceral adipose tissue: reliability and validity among the overweight and obese. *Obesity (Silver Spring)* 2012: 20(10): 2124-2132.
6. Partinen M, Gislason T. Basic Nordic Sleep Questionnaire (BNSQ): a quantitated measure of subjective sleep complaints. *J Sleep Res* 1995: 4(S1): 150-155.
7. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* 1991: 14(6): 540-545.
8. National Cholesterol Education Program Expert Panel on Detection E, Treatment of High Blood Cholesterol in A. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. *Circulation* 2002: 106(25): 3143-3421.
9. Ahmed A, Husain A, Love TE, Gambassi G, Dell'Italia LJ, Francis GS, Gheorghiade M, Allman RM, Meleth S, Bourge RC. Heart failure, chronic diuretic use, and increase in mortality and hospitalization: an observational study using propensity score methods. *Eur Heart J* 2006: 27(12): 1431-1439.
10. D'Agostino RB, Jr. Propensity scores in cardiovascular research. *Circulation* 2007: 115(17): 2340-2343.
11. Maislin G, Rubin DB. Design of Non-Randomized Medical Device Trials Based on Sub-Classification Using Propensity Score Quintiles, Topic Contributed Session on Medical Devices. *Proceedings of the Joint Statistical Meetings* 2010: 2182-2196.
12. Rosenbaum P, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983: 70: 41-44.
13. Rubin DB. Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation. *Health Services & Outcomes Research Methodology* 2001: 2: 169-188.
14. Rubin DB. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Stat Med* 2007: 26(1): 20-36.
15. Rubin DB. For Objective Causal Inference, Design Trumps Analysis. *The Annals of Applied Statistics* 2008: 2(3): 808-840.

16. D'Agostino RB, Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med* 1998: 17(19): 2265-2281.
17. Lieu TA, Au D, Krishnan JA, Moss M, Selker H, Harabin A, Taggart V, Connors A, Lung CER. Comparative Effectiveness Research in Lung Diseases and Sleep Disorders Recommendations from the National Heart, Lung, and Blood Institute Workshop. *Am J Resp Crit Care Med* 2011: 184(7): 848-856.
18. Yue LQ. Statistical and regulatory issues with the application of propensity score analysis to nonrandomized medical device clinical studies. *J Biopharm Stat* 2007: 17(1): 1-13; discussion 15-17, 19-21, 23-17 passim.
19. Zhou Z, Lam P. Discussion of: Statistical and regulatory issues with the application of propensity score analysis to nonrandomized medical device clinical studies. *J Biopharm Stat* 2007: 17(1): 25-27.

Table E1: Characteristics of lipid-lowering medication users and non-users at baseline

Characteristic	Lipid-Lowering Medication Use		
	No (N=613)	**Yes (N=193)**	**p**[†]
Age (years)	**52.8±10.6**	**59.5±8.9**	**<0.0001**
Male	80.4%	83.9%	0.275
BMI (kg/m^2)	33.5±5.8	33.7±5.4	0.726
Current Smoker	22.4%	16.7%	0.088
Excessive Alcohol	3.6%	3.1%	>0.999
Hypertension	**37.4%**	**74.6%**	**<0.0001**
Cardiovascular Disease	**4.3%**	**46.1%**	**<0.0001**
Diabetes Mellitus	**3.8%**	**24.9%**	**<0.0001**
Participate in Exercise	**60.3%**	**72.1%**	**0.003**
Epworth Sleepiness Scale	11.8±5.1	11.2±5.0	0.147
AHI (events/hour)	**44.0±20.6**	**47.5±20.7**	**0.039**
ODI (events/hour)	34.9±20.3	37.5±19.9	0.115
SaO_2 Nadir	76.2±8.1	76.1±7.7	0.843
Percent Time SaO_2<90	**13.1±19.9**	**16.7±19.3**	**0.023**
Total Cholesterol (mg/dL)	**206.6±40.8**	**159.2±35.4**	**<0.0001**
Total Cholesterol ≥200[‡]	**58.2%**	**13.5%**	**<0.0001**
LDL Cholesterol (mg/dL)	**150.9±37.0**	**103.6±31.4**	**<0.0001**
LDL Cholesterol ≥130[‡]	**70.8%**	**19.2%**	**<0.0001**
HDL Cholesterol (mg/dL)	40.3±11.3	40.2±10.7	0.878
HDL Cholesterol <40[‡]	58.1%	57.5%	0.890
Triglycerides (mg/dL)	176.0±87.6	177.4±72.5	0.826
Triglycerides ≥150[‡]	57.7%	58.0%	0.945

Significant differences shown in **bold**. [†]p-values from t-test and chi-square or Fisher's exact test; [‡]Abnormal cutoffs based on the NCEP ATPIII published criteria[8]; BMI: body mass index; AHI: apnea-hypopnea index; ODI: oxygen-desaturation index; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Table E2. Descriptive characteristics of patients with and without MRI data

Characteristic*	Have MRI (N=501)	Missing MRI (N=112)	p†
Age (years)	**53.3 ± 10.3**	**50.6 ± 11.5**	**0.023**
Male (%)	80.0%	82.1%	0.612
BMI (kg/m^2)	**32.5 ± 5.0**	**38.1 ± 7.1**	**<0.0001**
Current Smoker (%)	**20.8%**	**29.5%**	**0.048**
Excessive Alcohol (%)	3.0%	6.3%	0.098
Hypertension (%)	**35.3%**	**46.4%**	**0.028**
Cardiovascular Disease (%)	3.8%	6.3%	0.297
Diabetes Mellitus (%)	**2.6%**	**8.9%**	**0.004**
Participate in Exercise (%)	**62.6%**	**50.0%**	**0.014**
Epworth Sleepiness Scale	12.0 ± 5.1	11.1 ± 5.2	0.108
AHI (events/hour)	**42.6 ± 19.2**	**50.3 ± 25.0**	**0.003**
ODI (events/hour)	**33.1 ± 18.9**	**42.8 ± 24.0**	**<0.001**
SaO_2 Nadir	**76.8 ± 7.8**	**73.5 ± 9.2**	**0.001**
Percent Time SaO_2<90	**11.6 ± 16.2**	**19.7 ± 22.9**	**0.001**
Total Cholesterol (mg/dL)	**208.8 ± 38.8**	**197.0 ± 47.7**	**0.016**
Total Cholesterol ≥ 200 (%)‡	**60.3%**	**49.1%**	**0.030**
LDL Cholesterol (mg/dL)	**153.0 ± 35.2**	**141.6 ± 43.2**	**0.010**
LDL Cholesterol ≥ 130 (%)‡	**74.3%**	**55.4%**	**<0.001**
HDL Cholesterol (mg/dL)	**40.9 ± 11.6**	**37.7 ± 9.5**	**0.003**
HDL Cholesterol <40 (%)‡	**55.9%**	**67.9%**	**0.020**
Triglycerides (mg/dL)	**170.1 ± 82.0**	**202.1 ± 106.0**	**0.003**
Triglycerides ≥ 150 (%)‡	**55.3%**	**68.8%**	**0.009**

Significant differences shown in **bold**. *Results presented as mean ± standard deviation or N (percentage); †p-value from t-test (for continuous variables) and chi-square or Fisher's exact test (for categorical variables); ‡Abnormal cutoffs based on the NCEP ATPIII published criteria.[8] *Abbreviations:* MRI: Magnetic Resonance Imagine; BMI: body mass index; AHI: apnea-hypopnea index; ODI: oxygen desaturation index; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Table E3: Unadjusted Pearson correlations between obesity and OSA severity measures and natural log transformed lipid measures

Measure	Total Cholesterol		LDL Cholesterol		HDL Cholesterol		Triglycerides	
	rho	p	*rho*	p	*rho*	p	*rho*	p
BMI (kg/m^2)	**-0.14**	**0.001**	**-0.14**	**0.001**	**-0.18**	**<0.0001**	**0.29**	**<0.0001**
Weight (kg)	**-0.22**	**<0.0001**	**-0.21**	**<0.0001**	**-0.28**	**<0.0001**	**0.29**	**<0.0001**
Neck Cir. (cm)	**-0.21**	**<0.0001**	**-0.18**	**<0.0001**	**-0.33**	**<0.0001**	**0.35**	**<0.0001**
Waist Cir. (cm)	**-0.15**	**<0.001**	**-0.16**	**<0.001**	**-0.19**	**<0.0001**	**0.34**	**<0.0001**
Waist-to-hip Ratio	**-0.12**	**0.002**	**-0.11**	**0.006**	**-0.21**	**<0.0001**	**0.30**	**<0.0001**
Total Abdominal fat (cm^3)	-0.03	0.527	-0.06	0.209	**-0.09**	**0.049**	**0.30**	**<0.0001**
SAT (cm^3)	-0.01	0.873	-0.04	0.320	-0.02	0.648	**0.21**	**<0.0001**
VAT (cm^3)	-0.05	0.249	-0.05	0.243	**-0.16**	**<0.001**	**0.31**	**<0.0001**
AHI (events/hour)	-0.01	0.895	-0.03	0.484	0.02	0.671	**0.11**	**0.007**
ODI (events/hour)	0.07	0.100	**-0.08**	**0.046**	-0.07	0.091	**0.16**	**<0.001**
SaO_2 Nadir	0.05	0.184	0.06	0.151	0.06	0.146	**-0.09**	**0.022**
Percent Time SaO_2<90‡	-0.06	0.116	**-0.08**	**0.036**	-0.03	0.523	**0.11**	**0.005**

Significant correlations are shown in **bold**; ‡percent time SaO_2<90 natural log transformed for normality; *Abbreviations:* OSA: obstructive sleep apnea; BMI: body mass index; Cir.: circumference; SAT: subcutaneous abdominal fat; VAT: visceral abdominal fat; AHI: apnea-hypopnea index; ODI: oxygen desaturation index; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Table E4: Covariate comparisons for patients included in and excluded from final propensity score designed cohort

Baseline Characteristic	Adherent Users			Non-Users		
	Included (N=199)	Excluded (N=41)	p*	Included (N=118)	Excluded (N=38)	p*
Age (years)	**51.8 ± 10.4**	**58.4 ± 10.7**	**0.0003**	52.8 ± 10.0	53.3 ± 10.5	0.7774
Male (%)	163 (81.9%)	36 (87.8%)	0.3611	**95 (80.5%)**	**20 (52.6%)**	**0.0007**
BMI (kg/m^2)	**33.9 ± 5.9**	**36.1 ± 5.2**	**0.0349**	33.1 ± 5.9	32.8 ± 6.3	0.8010
Current Smoker (%)	**41 (20.6%)**	**3 (7.3%)**	**0.0453**	**24 (20.3%)**	**16 (42.1%)**	**0.0075**
Excessive Alcohol (%)	**5 (2.5%)**	**5 (12.2%)**	**0.0150**	3 (2.5%)	3 (7.9%)	0.1555
Hypertension (%)	**66 (33.2%)**	**32 (78.1%)**	**<0.0001**	34 (28.8%)	6 (15.8%)	0.1098
Cardiovascular Disease (%)	**5 (2.5%)**	**4 (9.8%)**	**0.0486**	2 (1.7%)	3 (7.9%)	0.0935
Diabetes Mellitus (%)	**3 (1.5%)**	**5 (12.2%)**	**0.0044**	1 (0.9%)	2 (5.3%)	0.1473
Participate in Exercise (%)	117 (58.8%)	26 (63.4%)	0.5830	69 (58.4%)	26 (68.4%)	0.2745
Epworth Sleepiness Scale	12.0 ± 5.0	13.8 ± 6.0	0.0533	**11.5 ± 4.8**	**9.4 ± 4.6**	**0.0174**
AHI (events/hour)	**44.8 ± 19.6**	**65.1 ± 19.9**	**<0.0001**	**40.8 ± 19.8**	**33.2 ± 13.4**	**0.0088**
ODI (events/hour)	**36.1 ± 19.3**	**60.0 ± 19.2**	**<0.0001**	**32.1 ± 19.0**	**21.9 ± 10.7**	**<0.0001**
SaO_2 Nadir	**76.2 ± 7.5**	**68.9 ± 8.8**	**<0.0001**	77.5 ± 7.3	79.0 ± 9.4	0.3734
Percent Time $SaO_2<90^{\dagger}$	**2.1 ± 1.0**	**3.5 ± 0.8**	**<0.0001**	**1.8 ± 1.0**	**1.3 ± 1.0**	**0.0114**
Total Cholesterol (mg/dL)	204.7 ± 41.6	202.6 ± 40.9	0.7697	**204.9 ± 35.4**	**224.6 ± 48.0**	**0.0153**
LDL Cholesterol (mg/dL)	149.4 ± 37.3	145.6 ± 38.8	0.5618	**149.7 ± 33.8**	**166.4 ± 42.0**	**0.0141**
HDL Cholesterol (mg/dL)†	40.0 ± 12.4	40.5 ± 9.9	0.8411	39.8 ± 10.3	41.5 ± 12.6	0.3959
Triglycerides (mg/dL)†	174.8 ± 86.1	189 ± 108.7	0.4348	176.0 ± 102.9	191.2 ± 102.0	0.4261

Significant differences shown in **bold**; *****p-value from t-test for continuous covariates or chi-square or Fisher's exact test for categorical covariates, where appropriate, comparing included and excluded participants; †variable natural log transformed for normality in matching heuristic. *Abbreviations:* BMI: body mass index; AHI: apnea-hypopnea index; ODI: oxygen desaturation index; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Table E5: Differences in two-year lipid changes between PAP Adherent and non-users within the PS designed cohort

Lipid	BMI Group*,†	*N (%) Abnormal*				*p*‡
		Adherent		***Non-Users***		
		Baseline	*Follow-Up*	*Baseline*	*Follow-Up*	
Total Cholesterol	*Overall*	*112 (56.3%)*	*128 (64.3%)*	*65 (55.1%)*	*75 (63.6%)*	*0.9842*
	<30	36 (64.3%)	40 (71.4%)	32 (68.1%)	34 (72.3%)	0.5419
	30-35	40 (61.5%)	47 (72.3%)	12 (36.4%)	19 (57.6%)	0.5299
	≥35	36 (46.2%)	41 (52.6%)	21 (55.3%)	22 (57.9%)	0.8624
LDL Cholesterol	*Overall*	*141 (70.9%)*	*151 (75.9%)*	*86 (72.9%)*	*86 (72.9%)*	*0.5471*
	<30	43 (76.8%)	46 (82.1%)	38 (80.9%)	38 (80.9%)	0.8111
	30-35	51 (78.4%)	55 (84.6%)	19 (57.6%)	22 (66.7%)	0.9861
	≥35	47 (60.3%)	50 (64.1%)	29 (76.3%)	26 (68.4%)	0.2801
HDL Cholesterol	*Overall*	*121 (60.8%)*	*61 (30.7%)*	*68 (57.6%)*	*41 (34.8%)*	*0.2163*
	<30	**31 (55.4%)**	**12 (21.4%)**	**23 (48.9%)**	**18 (38.3%)**	**0.0241**
	30-35	36 (55.4%)	17 (26.2%)	19 (57.6%)	9 (27.3%)	0.9584
	≥35	54 (69.2%)	32 (41.0%)	26 (68.4%)	14 (36.8%)	0.7792
Triglycerides	*Overall*	*119 (59.8%)*	*109 (54.8%)*	*65 (55.1%)*	*58 (49.2%)*	*0.7349*
	<30	26 (46.4%)	28 (50.0%)	17 (36.2%)	18 (38.3%)	0.6251
	30-35	39 (60.0%)	37 (56.9%)	20 (60.6%)	16 (48.5%)	0.5642
	≥35	54 (69.2%)	44 (56.4%)	28 (73.7%)	24 (63.2%)	0.7338

Significant associations between PAP usage and change shown in **bold**; *The propensity score matched sample included 199 adherent (56 with BMI<30; 65 with BMI 30-35; 78 with BMI≥35) and 118 non-users (47 with BMI<30; 33 with BMI 30-35; 38 with BMI≥35); †p-values for an interaction between PAP and BMI group: p=0.724 for total cholesterol, p=0.844 for LDL, p=0.096 for HDL, p=0.849 for triglycerides; ‡p-value from generalized estimating equation (GEE) model comparing adherent and non-users within purposefully designed observational study, adjusted for propensity score subclass; *Abbreviations*: BMI: body mass index; LDL: low-density lipoprotein; HDL: high-density lipoprotein; LS: least squares; SE: standard error

Table E6: Differences in two-year lipid changes in those with abnormal lipid levels at baseline between PAP adherent and non-users within the propensity score designed observational study

Lipid	BMI Group*,†	*LS Mean ± SE Change*		*p*‡
		Adherent	*Non-Users*	
Total Cholesterol	*Overall*	*-9.1 ± 2.7*§	*-1.9 ± 3.6*	*0.126*
	<30	-10.1 ± 5.5	1.4 ± 5.9	0.183
	30-35	1.8 ± 4.2	-3.7 ± 7.8	0.548
	≥35	-19.9 ± 4.1§	-6.5 ± 5.5	0.065
LDL Cholesterol	*Overall*	*-9.4 ± 2.2*§	*-6.3 ± 2.9*§	*0.403*
	<30	-13.7 ± 4.3§	-4.9 ± 4.6	0.180
	30-35	0.6 ± 3.4	-11.3 ± 5.7§	0.080
	≥35	-15.5 ± 3.7§	-6.2 ± 4.9	0.150
HDL Cholesterol	*Overall*	*8.9 ± 0.8*§	*8.7 ± 1.1*§	*0.911*
	<30	9.5 ± 1.2§	6.4 ± 1.4§	0.119
	30-35	10.9 ± 1.4§	9.8 ± 1.9§	0.651
	≥35	6.9 ± 1.4§	10.8 ± 2.2§	0.157
Triglycerides	*Overall*	*-15.4 ± 8.7*	*-17.9 ± 11.9*	*0.866*
	<30	-2.2 ± 24.6	12.5 ± 31.3	0.731
	30-35	7.2 ± 17.2	-3.9 ± 24.5	0.717
	≥35	-42.2 ± 9.3§	-38.2 ± 13.2§	0.810

Significant differences shown in **bold**; *The propensity score matched sample included 199 adherent (56 with BMI<30; 65 with BMI 30-35; 78 with BMI≥35) and 118 non-users (47 with BMI<30; 33 with BMI 30-35; 38 with BMI≥35); †p-values for an interaction between PAP and BMI group: p=0.139 for total cholesterol, **p=0.041** for LDL, p=0.123 for HDL, p=0.508 for triglycerides; ‡p-value from ANCOVA comparing adherent and non-users within purposefully designed observational study, adjusted for propensity score subclass and baseline lipid level; §Within PAP group estimate of lipid change significantly different from 0 (p<0.05). *Abbreviations*: BMI: body mass index; LDL: low-density lipoprotein; HDL: high-density lipoprotein; LS: least squares; SE: standard error

Table E7. Differences in two-year lipid changes between PAP Adherent and Non-users with percent time SaO_2<90 greater than the 75th percentile

Lipid Change	*LS Mean ± SE Change*		*p**
	Adherent (n=46)	*Non-Users (n=17)*	
Total Cholesterol	0.3 ± 4.7	4.0 ± 7.9	0.689
LDL Cholesterol	-3.8 ± 4.0	2.9 ± 6.8	0.410
HDL Cholesterol	4.6 ± 1.2	4.6 ± 2.0	0.998
Triglycerides	-5.9 ± 11.1	-40.9 ± 18.7	0.119

[*]p-value from ANCOVA comparing adherent and non-users within purposefully designed observational study, adjusted for propensity score subclass and baseline lipid level; [‡]within PAP group estimate of lipid change significantly different from 0 ($p<0.05$). *Abbreviations*: BMI: body mass index; LDL: low-density lipoprotein; HDL: high-density lipoprotein; LS: least squares; SE: standard error.

Figure E1: Diagram illustrating the analysis population. The chart presented below illustrates the flow of patients from initial recruitment to inclusion in the baseline analysis sample (n=613) and the propensity score designed observational cohort (n=317; 199 PAP adherent and 118 non-users). *Abbreviations*: OSA: obstructive sleep apnea; PAP: positive airway pressure; MAD: mandibular advancement device.