CAUSAL EFFECT OF SLEEP DISTURBANCE ON COGNITIVE DECLINE IN OLDER ADULTS

by

Yuanyuan Jiao

BS, Shandong Normal University, China, 2001

PhD, University of Wyoming, 2012

Submitted to the Graduate Faculty of

Department of Biostatistics

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Yuanyuan Jiao

It was defended on

August 8th, 2018

and approved by

Ying Ding, PhD
Assistant Professor, Department of Biostatistics
Graduate School of Public Health
University of Pittsburgh

Yu Cheng, PhD
Associate Professor, Department of Statistics
Associate Professor, Department of Biostatistics
University of Pittsburgh

Thesis Director: Chung-Chou H. Chang, PhD
Professor, Department of Biostatistics
Professor, Clinical and Translational Science Institute
Graduate School of Public Health and School of Medicine
University of Pittsburgh

Chung-Chou H. Chang, PhD

CAUSAL EFFECT OF SLEEP DISTURBANCE ON COGNITIVE DECLINE IN OLDER ADULTS

Yuanyuan Jiao, MS

University of Pittsburgh, 2018

ABSTRACT

Prior studies have shown that sleep disturbance is closely associated with cognitive decline in older adults. However, one cannot use standard regression models to verify the causal relationship between sleep disorder and cognitive dysfunction. In this study, by combining propensity score weighting and honest causal tree technique, we balanced baseline characteristics between individuals with and without a certain type of sleep disorder, effectively partitioned older adults into groups based on the baseline conditions, and estimated heterogeneity in sleep disturbance impacts on cognitive function. We analyzed the data collected from the first nine waves of an ongoing community-based cohort study and the propensity score weighting causal tree model showed the causal effect of sleep disturbance on cognitive decline in various types of sleep disorder and cognitive domains. Sleep disorders caused faster decline in the memory and visuospatial domains. In addition, these causal relationship showed different effects among people with different sociodemigraphic or baseline health conditions, including age, gender, self-reported general health, systolic blood pressure (BP), diastolic BP, exercise, subjective memory complaint, and other baseline cognitive domain scores. Our findings advance the knowledge in cognitive dysfunction among the elderly and allow us to validate sleep disturbance as a therapeutic target for treating cognitive decline in older adults.

PUBLIC HEALTH SIGNIFICANCE: Sleep deprivation and cognitive impairment are common among older adults yet the causal relationship between sleep disturbance and cognitive decline remains controversial. Causal tree method employed in this study directly clarified the causal effect of sleep deprivation on cognitive degeneration, thus improves our understanding of the underlying mechanisms for cognitive impairment among the elderly also helps clinicians with diagnosis and prognosis. In addition, the modifiable moderators examined in this study can help clinicians and public health practitioners find appropriate prevention and treatments for sleep disturbances.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

First of all, I want to extend my sincerest thanks to my thesis supervisor, chair of my thesis committee, Dr. Chung-Chou H. Chang, for giving me the opportunity to undertake this research study, and for providing me her heartful support, extraordinary scientific guidance, inspiration, and suggestions during the whole course of this thesis.

I would like to express my gratitude to Victor Talisa, for providing indispensable advice in handling data and constructing models.

I would also like to express my gratitude to my academic advisor, Dr. Gary M. Marsh, for providing support and guidance through my academic study.

I have great pleasure in acknowledging my gratitude to Dr. Ying Ding, for sparing her valuable time whenever I approached her and showing me the way ahead.

I would like to thank Dr. Yu Cheng, for her valuable comments and advice on my thesis.

I would like to thank Dr. Mary Ganguli, for providing data source for my thesis.

I warmly thank all the faculty members, staff and friends in Biostatistics Master Program at University of Pittsburgh who have been so helpful and cooperative in giving their support at all times to help me achieve my goal.

My acknowledgement would be incomplete without thanking the biggest source of my strength, my family. The love and support of my husband **Qi Guan**, and my daughters **Fiona Guan & Celina Guan**, and the love and care of my parents **Zhenwu Jiao & Yuhua Fang**, my

sister **Li Jiao**, and her husband **Karl Gai**, and my nephew **Steven Gai**, have all made a tremendous contribution in helping me follow my dream of getting this degree.

1.0 INTRODUCTION

Sleep deprivation is a well-recognized health issue among the elderly. Approximately half of the older adults suffer from a sleep disorder with estimated 25% show symptoms of insomnia [1]. Insomnia is characterized by reduced sleep time, shorter sleep duration, and worsened sleep quality. Sleep disturbances commonly observed in older people include difficulty falling asleep, difficulty staying asleep, early morning awakening, excessive daytime sleepiness, excessive daytime napping, and sleep apnea. Life habits (including exercise, use of tobacco, and alcohol), sleep habits, genetic, physiological, and psychological factors may cause insomnia in the elderly.

A large body of literature have shown that sleep insufficiency and disturbance are closely associated with cognitive impairment or dementia [2]. Evidence from cross-sectional studies and prospective studies in older adults supports the association of poor sleep quality with worse cognitive performance in attention function, visuospatial skills, and memory [3-10]. Several prospective studies have revealed that people having excessive daytime sleepiness are at an increased risk of global cognitive decline [11, 12] and dementia [13]. Moreover, retrospective studies, and prospective studies have also found the association of insomnia [14, 15] and sleep-disordered breathing [16-21], with reduced cognitive function. However, some cross-sectional studies and prospective studies have not confirmed this association [5, 16, 22-25].

These inconsistent findings might be due partly to the variability in measurement of sleep quantity and quality, and the measurement of cognitive function, partly to heterogeneity in age grouping, duration of follow-up, and methods of analysis.

Regression models are widely used to study the linkage between sleep disturbance and cognition decline, however, such application of the models is questionable and the relationship it was intended to establish is weak, because the regression models target on the average effect of exposure (e.g. sleep disturbance) across the whole sample subjects and ignore heterogeneity in sample characteristic. With the development in machine learning methods, causal tree model [26] was proposed to use observational data in estimating the effect of an exposure in each of the subgroups with similar effects. Causal tree method could effectively partition subjects into subgroups with similar exposure effect based on the given baseline covariates. Therefore, this method can be used to analyze heterogeneous exposure effects based on a set of baseline subject characteristics.

In this study, we assessed whether sleep disturbance in older adults increases the risk of cognitive decline during a 9-year follow-up period. Causal tree method was used to effectively partition older adults into groups and estimate average sleep-disturbance effects for different subgroups, i.e., older adults in the same subgroup having similar sleep-disturbance effects. In addition, propensity score weighting derived from generalized boosted models (GBMs) was incorporated into the causal tree model to reduce the differences of baseline variables between groups. In this study, we used the propensity score weighting causal tree model to examine the causal relationship between sleep disturbance and cognitive decline in older adults and to identify subgroups with similar causal effects based on the baseline conditions. We also investigate whether any of the causal relationship was moderated by the known risk factors of

cognitive decline including age, sex, general health condition, blood pressures, exercise, subjective memory complaint, and baseline cognitive domain scores.

2.0 METHODS

2.1 STUDY POPULATION

We used data collected from the Monongahela-Youghiogheny Healthy Aging Team (MYHAT) study which was a community-based cohort study investigating risk and protective factors on cognitive function among the elderly. MYHAT participants were aged 65 or older registered voters randomly selected and recruited from the Monongahela-Youghiogheny area of the Allegheny County of Southwestern Pennsylvania [27]. The inclusion criteria are older individuals who were not residing in a long-term care institution at the date of recruitment and the exclusion criteria include: previously severe impairment in vision and/or hearing, decisional incapacity, and severe health condition. Participants are followed annually to monitor health condition and behavior change and to take a set of neuropsychological cognitive examinations. Details of the MYHAT study have been previously described and published [27]. The MYHAT study was approved by the University of Pittsburgh Institutional Review.

The MYHAT study enrolled 1,982 individuals, of whom we excluded 375 individuals who do not have 2 or more cognitive domain test scores during the 9-year follow up; 675 individuals do not have data on disruptive sleep questions; 7 individuals do not have data on the rest 6 types of sleep quality questions; and 121 individuals do not have complete information on the demographic or baseline health-related covariates. Our study dataset does not include

disruptive sleep variable that has too many missing data, and our final analytic dataset includes 804 participants (Figure 1).

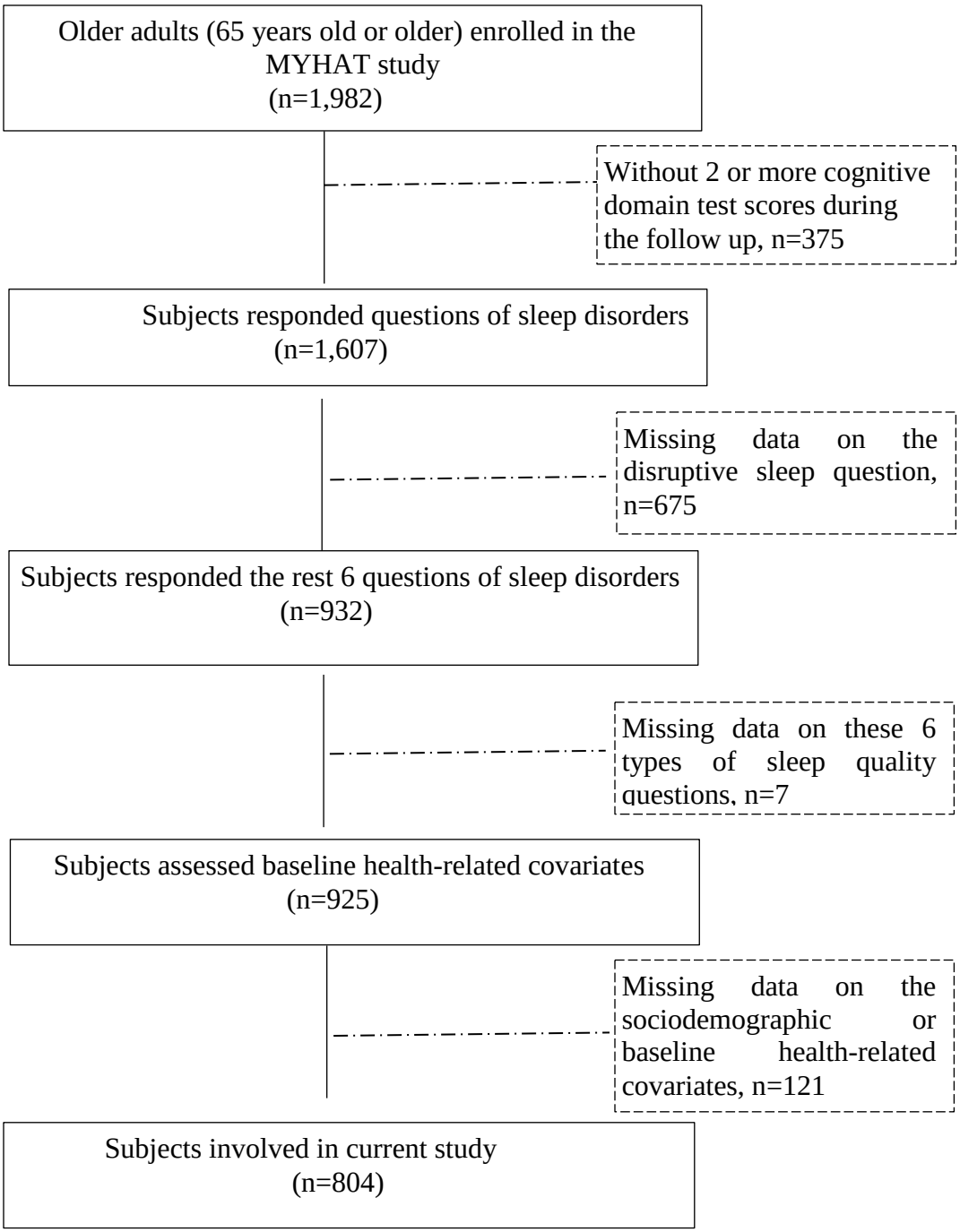


Figure 1 Flowchart of subjects included in the current study

2.2 PRIMARY EXPOSURE VARIABLE

MYHAT Participants were administered six questions related to the quality of sleep: (1) Do you take longer than half an hour to fall sleep (labeled as *Difficulty falling asleep* subsequently)?, (2) Do you wake up during the night(including to go to the bathroom) and find that it takes you more than half an hour to go back to sleep (labeled as *Difficulty staying asleep* subsequently)?, (3) Do you wake up earlier than you want to and find that you can't go back to sleep (labeled as *Waking up earlier than desired* subsequently)?, (4) Do you ever fall asleep while actively doing something during the day (labeled as *Excessive daytime sleepiness* subsequently)?, (5) Do you doze off in most of the days during a week (labeled as *Excessive daytime napping* subsequently)?, and (6) Since our last visit, have you been told that you have sleep apnea (labeled as *Sleep apnea* subsequently)? The response categories for Questions 1-4 listed above included 0=never/rarely, 1=sometimes, 2=usually, or 3=never sleep/awake all night. The response categories for Questions 5-6 included 0=no, or 1=yes. Although all questions were administered at the baseline and annual follow-up, our primary exposure variable only used baseline sleeping patterns and converted each question into a binary response: 0=no/never or 1=other. For participants who did not answer a particular question, the variable was coded as missing.

2.3 PRIMARY OUTCOME VARIABLES

Cognitive functions are measured annually via five cognitive domains z-scores which include attention/processing speed, executive functions, memory, language, and visuospatial skills.

Detailed descriptions of the tests related to these 5 domains can be found in previous study [27]. The primary outcome variable is the longitudinal cognitive change, which was computed by the estimated slope of cognitive domain scores over the 9-year study period from a linear mixed effects model with random intercept and random slope of time.

2.4 COVARIATES

Covariates of interest in this study included variables measured at the baseline, which included age (categorized as 65-74, 75-84, and 85+ years old), sex (male or female), education (categorized as less than, equal to, and greater than high school education), self-reported general health (categorized as poor/fair, good, and very good), baseline cognitive domain z-scores (included attention, executive, language, memory, and visuospatial domains), systolic blood pressure (BP), diastolic BP, smoking status (categorized as never smoker or ever smoker), alcohol consumption (categorized as never drinker or ever drinker), having history of cerebrovascular diseases (no/yes), having history of cardiovascular diseases (no/yes), exercise (categorized as never or ever), total number of prescription medication use (categorized as <3 or ≥3), Instrumental Activities of Daily Living (IADL) score (categorized as 0 or >0), subjective memory complaints (categorized as 0 or >0), modified CES-D depression score (categorized as <3 or ≥3), and history of hypertension (categorized as never or ever).

Cerebrovascular diseases included stroke and TIA (mini-stroke). Both stroke and TIA were categorized as never or ever. Cerebrovascular disease was categorized as never (if never had a stroke or a TIA) and ever (if ever had a stroke or a TIA). Cardiovascular diseases included

heart attack, congestive heart failure, irregular heartbeat, and cardiac arrest. If ever had any of these conditions, the variable was coded as 1, otherwise was coded as 0.

2.5 STATISTICAL ANALYSES

2.5.1 Descriptive statistics

We performed descriptive statistical analysis using Stata 14 (StataCorp, College Station, TX) to (1) analyze prevalence of the six sleep disturbance patterns (Table 1); (2) analyze the relationship between sleep disturbance and other baseline covariates (Table 2); and (3) examine the reliability of randomly splitting participants into the training and estimation sets in the honest causal tree analysis that will be described in Section 3.3.1 (Table 4).

Continuous variables and categorical variables were summarized as mean±standard deviation (SD) and number of participants with percentage, respectively. We used chi-square (for categorical variables), t-test, or Kruskal-Wallis test (for continuous variables) to assess the difference between groups. The significance level for all analyses was set at $P < 0.05$.

2.5.2 Propensity score method

To remove selection bias associated with nonrandomized or observational studies, we estimated the causal effect of sleep disturbance on cognitive decline by incorporating propensity score weights in the main analysis.

Propensity score for an individual i, denoted as $e(X_i)$ or $P(T_i|X_i)$, is a conditional probability of being exposed given his/her baseline characteristics, where T_i is the exposure status ($T = 1$ for exposed, $T = 0$ for nonexpsoed) and X_i is a vector of observed baseline covariates. In our study, the exposure group included individuals having a sleep problem and the nonexposure group included individuals without the sleep problem. Applying propensity score weighting in the main analysis model will result in a more unbiased estimate of the exposure effect because the pre-exposure variables were much similar (or balanced) between the exposure and the nonexposure groups [28]. Note that using propensity score weighting method to estimate the causal exposure effect replies on three key assumptions: consistency (i.e., an individual's potential outcomes under his/her observed exposure history is precisely her observed outcomes), exchangeability (i.e., no unmeasured confounders), and positivity (i.e., no individuals of one pattern of pre-exposure covariates are all exposed or all nonexposed). Although none of these assumptions can be verified using the observed data, we checked whether individuals in any set of baseline covariate values all had sleep disturbance or none had sleep disturbance (i.e., check for positivity).

Traditional approach of propensity score estimation is to fit a logistic regression model of exposure status on the observed baseline covariates. Recently researchers proposed the use of machine learning methods and proved that these methods reduce bias and the mean-squared error more than the traditional logistic regression method does. Among the proposed machine learning methods, generalized boosted model (GBM) is the most frequently used in estimating the propensity scores[29].

GBM is a nonparametric, automated, data-adaptive algorithm combining classification and regression tree (CART) with boosting, which can be used to estimate nonlinear and

interactive effects between the exposure status and the observed pre-exposure baseline covariates [30] using a smoother fit. The final model consists of several regression trees by starting with a simple regression tree and iteratively adding another tree. At each iteration the new tree is chosen to be added if it provides the best fit to the residuals of the model resulted from the previous iteration. At the last iteration, all included trees are used to obtain an overall piecewise constant function. [30]

By applying the GBM-based propensity score weighting method, one can reduce the differences in the baseline covariates between the exposed and the nonexposed groups. We used absolute standardized mean difference (ASMD) to measure the degree of balancing between the two groups after weighting. ASMD or effect sizes *d* is defined as the absolute weighted group mean difference divided by the unweighted standard deviation in the exposed group. We ran the GBM algorithm using the R package `twang` with the maximum iteration number (`n.tree`) set to 10,000 in order to minimize the ASMD. Because our goal is to identify heterogeneity of the effects of sleep disturbance, that is, we are interested in estimating the effect of sleep disturbance in different subgroups, especially for those who are more likely to have sleep problems, we employed *the average treatment effect across the treated groups* (ATT) instead of *average treatment effect* (ATE) in the propensity score analysis.

We used the R package `twang` to implement the propensity score estimation via GBM. To optimize the likelihood function at each iteration in GBM, we allowed four-way interactions (`interaction.depth` = 4 in `twang`) between all covariates. We performed the metric ASMD (`es.mean` in `twang`) to assess the balance and summarize across covariates. Shrinkage was set to be 0.01 to yield smooth fits. To evaluate the success of propensity score weighting, we

examined whether the exposed and the nonexposed groups have similar distributions for all covariates.

2.5.3 Causal tree and honest causal tree

The causal tree method [26] was derived from the traditional CART approach for investigating heterogeneity in exposure effects, which recursively partition subjects into subgroups according to the observed characteristics. A causal tree model focuses on estimating conditional average exposure effects instead of the prediction outcomes. To avoid overfitting, the *honest* causal tree method was proposed. The honest causal tree is a two-step procedure that randomly divides subjects into the training subsample and the estimation subsample, where the training subsample is used to split the tree and the estimation subsample is used to estimate the exposure effects.

We applied the honest causal tree method with GBM-based propensity score weighting to estimate the causal effect of each type of sleep disturbance on cognitive changes measured by the slopes of change for the 5 cognitive domain scores.

Let N_{1j} and N_{0j} be the numbers of individuals in the exposure group ($T = 1$) and nonexposure group ($T = 0$) for the jth leaf of the tree [1]. Note: $N_{1j} + N_{0j} = N_j$, sample size for the jth leaf. Using the propensity score weighting method, we estimated the outcome difference between the exposure and the nonexposure groups by:

$$\hat{O}(X_j) = \bar{Y}_{1j} - \bar{Y}_{0j} = \left(\sum_{\mathrm{T}=1} \frac{Y_{1j}}{e(X_j)}\right) - \left(\sum_{\mathrm{T}=0} \frac{e(X_j)}{1-e(X_j)}\right)^{-1} \left(\sum_{\mathrm{T}=0} \frac{Y_0}{1-e(X_j)}\right),$$

where Y_{1j} and Y_{0j} indicate outcomes for subjects who exposed and who did not expose, respectively; and $e(X_j)$ denotes the propensity of being exposed given the observed characteristics X_j,for leaf j.

Let S_{1j}^2 and S_{0j}^2 denote the within-leaf variance of outcome Y_j among exposed and the nonexposed individuals, respectively. Therefore, variance of $\hat{O}(X_j)$ can be estimated by

$$Var\{\hat{O}(X_j)\} = \frac{S_{1j}^2}{N_{1j}} + \frac{S_{0j}^2}{N_{0j}}.$$

Honest causal tree, which is different from the CART, splits at a node with minimum expected mean squared error EMSE(S^{tr}, S^{est}) over the training and the estimation sets. That is, for a node, we minimize

$$\text{EMSE}(S^{tr}, S^{est}) = -\left(\frac{1}{N_{tr}}\right)\sum \hat{O}(X)^2 + \left(\frac{1}{N_{tr}} + \frac{1}{N_{est}}\right)\sum \text{Var}[\hat{O}(X)],$$

where S^{tr} and S^{est} denote the training samples and the estimation samples, respectively; and N^{tr} and N^{est} denote the sample size of the training and the estimation samples, respectively.

We utilized the R package `CausalTree` to analyze heterogeneity effects of sleep disturbance on the decline of cognitive function. We randomly split the observations in half to form a training set and an estimation data set. The minimum leaf size was set to having at least thirty exposure and thirty nonexposure subjects. When pruning the tree, we picked the optimum complexity parameter as the minimum cross-validation (CV), `error(minerr)` and set CV error ≤ miner +1 SD.

3.0 RESULTS

3.1 DESCRIPTIVE

Nearly one third of the older adults reported to have sleep disturbance at baseline (Table 1). For each type of the sleep disturbance, 40.30% had difficulty falling asleep, 42.41% had difficulty staying asleep, 27.49% had the problem of waking up earlier than desired, 30.35% had the problem of excessive or uncontrolled daytime sleepiness, 38.81% had the issue of excessive daytime napping, and 9.08% had self-reported sleep apnea.

Table 1. Descriptive statistics for sleeping deprivation at baseline

Baseline sleeping pattern	**Present**	
Difficult in falling asleep,n(%)	324	(40.30)
Difficult in staying asleep,n(%)	341	(42.41)
Earlier waking up than desired,n(%)	221	(27.49)
Excessive daytime sleepiness,n(%)	244	(30.35)
Excessive daytime napping,n(%)	312	(38.81)
Sleep apnea,n(%)	73	(9.08)

N=804 total participants

Many of the baseline covariates are associated with sleep quality (Table 2). Older adults aged 75-84 were more likely to experience early awakenings than desired ($P<0.05$). Difficulty falling asleep ($P<0.001$) and excessive daytime napping ($P=0.01$) were more common in older

women than in older men. Individuals with high school degree or being nonsmokers were more likely to suffer from having difficulty falling asleep ($P<0.05$). Self-reported good health was related to the presence of sleep disturbance, including difficulty falling asleep, difficulty staying asleep, waking up earlier than desired, and sleep apnea. Older adults with sleep disturbance, except sleep apnea, tended to have modified CES-D depression score <3 (less depressed) or have subjective memory complaints. Subjects without daily activity dependence (IADL score=0) suffered from excessive daytime napping and sleep apnea. The elderly who took more prescription medication (≥ 3) were more likely to have difficulty falling asleep and report sleep apnea.

Health conditions could disrupt sleep in older people. The subjects having systolic BP tended to have difficulty falling asleep ($P<0.05$) and excessive daytime sleepiness ($P<0.01$). Diabetes were related to difficulty falling asleep ($P<0.05$), difficulty staying asleep ($P<0.05$), excessive daytime napping ($P<0.01$), and sleep apnea ($P<0.001$). Those who having hypertension tended to suffer from difficulty falling asleep ($P<0.001$), excessive daytime sleepiness ($P=0.05$), and sleep apnea ($P<0.05$). Excessive daytime napping was common in older adults having history of heart attack or congestive heart failure ($P<0.05$). Irregular heartbeat was related to difficulty falling asleep ($P<0.05$), difficulty staying asleep ($P<0.05$), and waking up earlier than desired ($P<0.01$).

Sleep deprivation were also related to cognitive impairment in older adults. Older adults with excessive daytime sleepiness tended to have lower baseline cognitive scores in executive function ($P<0.05$), language ($P<0.001$), and memory domains ($P<0.001$). In addition, lower executive domain score was associated with difficulty falling asleep ($P<0.05$).

Table 2. Descriptive of sleeping pattern and other covariates at baseline

Baseline variable	Difficult in falling asleep (n=804)			Difficult in staying asleep (n=804)			Earlier waking up than desired (n=804)			Excessive daytime sleepiness (n=804)			Excessive daytime napping (n=804)			Sleep apnea (n=804)		
	No	Yes	P value	No	Yes	P value	No	Yes	P value	No	Yes	P value	No	Yes	P value	No	Yes	P value
Age in years,n(%)			0.239*			0.737*			0.025*			0.359*			0.158*			0.052*
65-74	192 (40.00)	119 (36.73)		184 (39.74)	127 (37.24)		209 (35.85)	102 (46.15)		225 (40.18)	86 (35.25)		202 (41.06)	109 (34.94)		273 (37.35)	38 (52.05)	
75-84	238 (49.58)	159 (49.07)		226 (48.81)	171 (50.15)		303 (51.97)	94 (42.53)		272 (48.57)	125 (51.23)		230 (46.75)	167 (53.53)		368 (50.34)	29 (39.73)	
85+	50 (10.42)	46 (14.20)		53 (11.45)	43 (12.61)		71 (12.18)	25 (11.31)		63 (11.25)	33 (13.52)		60 (12.20)	36 (11.54)		90 (12.31)	6 (8.22)	
Female gender, n(%)	292 (60.83)	241 (74.38)	<0.001*	296 (63.93)	237 (69.50)	0.099*	376 (64.49)	157 (71.04)	0.080*	381 (68.04)	152 (62.30)	0.113*	343 (69.72)	190 (60.90)	0.010*	491 (67.17)	42 (57.53)	0.097*
Education, n(%)			0.023*			0.116*			0.439*			0.366*			0.429*			0.070*
<HS	65 (13.54)	40 (12.35)		62 (13.39)	43 (12.61)		73 (12.52)	32 (14.48)		67 (11.96)	38 (15.57)		62 (12.60)	43 (13.78)		91 (12.45)	14 (19.18)	
HS	195 (40.63)	163 (50.31)		192 (41.47)	166 (48.68)		255 (43.74)	103 (46.61)		254 (45.36)	104 (42.62)		228 (46.34)	130 (41.67)		334 (45.69)	24 (32.88)	
>HS	220 (45.83)	121 (37.35)		209 (45.14)	132 (38.71)		255 (43.74)	86 (38.91)		239 (42.68)	102 (41.80)		202 (41.06)	139 (44.55)		306 (41.86)	35 (47.95)	
Self-reported geneal health,n(%)			0.014*			<0.001*			0.001*			0.067*			0.143*			0.033*
Poor/Fair	61 (12.71)	54 (16.67)		54 (11.66)	61 (17.89)		68 (11.66)	47 (21.27)		70 (12.50)	45 (18.44)		61 (12.40)	54 (17.31)		98 (13.41)	17 (23.29)	
Good	211 (43.96)	162 (50.00)		194 (41.90)	179 (52.49)		272 (46.66)	101 (45.70)		261 (46.61)	112 (45.90)		231 (46.95)	142 (45.51)		338 (46.24)	35 (47.95)	
Very good	208 (43.33)	108 (33.33)		215 (46.44)	101 (29.62)		243 (41.68)	73 (33.03)		229 (40.89)	87 (35.66)		200 (40.65)	116 (37.18)		295 (40.36)	21 (28.77)	
Ever smoking, n(%)	255 (53.13)	147 (45.37)	0.031*	244 (52.70)	158 (46.33)	0.074*	285 (48.89)	117 (52.94)	0.304*	288 (51.43)	114 (46.72)	0.220*	233 (47.36)	169 (54.17)	0.060*	364 (49.79)	38 (52.05)	0.713*
Ever drinking, n(%)	416 (86.67)	284 (87.65)	0.682*	400 (86.39)	300 (87.98)	0.508*	506 (86.79)	194 (87.78)	0.709*	493 (88.04)	207 (84.84)	0.214*	432 (87.80)	268 (85.90)	0.432*	636 (87.00)	64 (87.67)	0.871*
Exercise, n(%)	302 (62.92)	209 (64.51)	0.646*	296 (63.93)	215 (63.05)	0.798*	367 (62.95)	144 (65.16)	0.561*	367 (65.54)	144 (59.02)	0.077*	305 (61.99)	206 (66.03)	0.247*	459 (62.79)	52 (71.23)	0.153*
mCESD score >=3, n(%)	23 (4.79)	59 (18.21)	<0.001*	31 (6.70)	51 (14.96)	<0.001*	48 (8.23)	34 (15.38)	0.003*	49 (8.75)	33 (13.52)	0.040*	41 (8.33)	41 (13.14)	0.028*	71 (9.71)	11 (15.07)	0.149*
IADL score >0, n(%)	39 (8.13)	30 (9.26)	0.573*	41 (8.86)	28 (8.21)	0.747*	52 (8.92)	17 (7.69)	0.579*	42 (7.50)	27 (11.07)	0.097*	29 (5.89)	40 (12.82)	0.001*	58 (7.93)	11 (15.07)	0.038*
No of Rx medications	337 (70.21)	249 (76.85)	0.038*	326 (70.41)	260 (76.25)	0.066*	416 (71.36)	170 (76.92)	0.113*	406 (72.50)	180 (73.77)	0.709*	348 (70.73)	238 (76.28)	0.084*	522 (71.41)	64 (87.67)	0.002*

Table 2 Continued

>=3,n(%)																		
Subjective memory complaint, n(%)	295 (61.46)	222 (68.52)	0.040*	284 (61.34)	233 (68.33)	0.041*	358 (61.41)	159 (71.95)	0.005 *	342 (61.07)	175 (71.72)	0.004*	296 (60.16)	221 (70.83)	0.002*	464 (63.47)	53 (72.60)	0.121*
Health outcome history																		
Systolic blood pressure, mean(SD)	130.54 (15.05)	133.27 (16.08)	0.032#	131.33 (14.99)	132.05 (16.23)	0.514#	131.59 (15.58)	131.78 (15.41)	0.844 #	130.58 (15.20)	134.08 (16.00)	0.003#	131.45 (14.99)	131.94 (16.35)	0.800#	131.60 (15.54)	132.01(15.44)	0.889#
Diastolic blood pressure, mean(SD)	72.37 (8.51)	73.43 (8.16)	0.088#	72.62 (8.47)	73.04 (8.27)	0.371#	72.53 (8.64)	73.50 (7.64)	0.247 #	72.61 (8.31)	73.21 (8.56)	0.179#	72.63 (8.52)	73.05 (8.17)	0.247#	72.71 (8.35)	73.63 (8.76)	0.266#
Stroke, n(%)	15 (3.13)	11 (3.40)	0.832#	16 (3.46)	10 (2.93)	0.679#	16 (2.74)	10 (4.52)	0.203 #	19 (3.39)	7 (2.87)	0.699#	16 (3.25)	10 (3.21)	0.971#	21 (2.87)	5 (6.85)	0.078#
TIA,n(%)	38 (7.92)	39 (12.04)	0.051*	42 (9.07)	35 (10.26)	0.570*	58 (9.95)	19 (8.60)	0.561 *	51 (9.11)	26 (10.66)	0.493*	46 (9.35)	31 (9.94)	0.783*	69 (9.44)	8 (10.96)	0.676*
Diabetes, n(%)	123 (25.62)	61 (18.83)	0.024*	118 (25.49)	66 (19.35)	0.041*	141 (24.19)	43 (19.46)	0.154 *	127 (22.68)	57 (23.36)	0.832*	95 (19.31)	89 (28.53)	0.002*	155 (21.20)	29 (39.73)	<0.001*
Hypertension ,n(%)	314 (65.42)	254 (78.40)	<0.001 *	315 (68.03)	253 (74.19)	0.058*	403 (69.13)	165 (74.66)	0.124 *	384 (68.57)	184 (75.41)	0.050*	340 (69.11)	228 (73.08)	0.228*	508 (69.49)	60 (82.19)	0.023*
Heartattack,n (%)	58 (12.08)	51 (15.74)	0.137*	56 (12.10)	53 (15.54)	0.158*	77 (13.21)	32 (14.48)	0.638 *	76 (13.57)	33 (13.52)	0.986*	56 (11.38)	53 (16.99)	0.024*	97 (13.27)	12 (16.44)	0.451*
Congestive heart failure, n(%)	28 (5.83)	25 (7.72)	0.270*	26 (5.62)	27 (7.92)	0.215*	34 (5.83)	19 (8.60)	0.097 *	33 (5.89)	20 (8.20)	0.150*	25 (5.08)	28 (8.97)	0.042*	44 (6.02)	9 (12.33)	0.134*
Irregular heartbeat, n(%)	125 (26.04)	108 (33.33)	0.037*	119 (25.70)	114 (33.43)	0.027*	153 (26.24)	80 (36.20)	0.005 *	154 (27.50)	79 (32.38)	0.113*	130 (26.42)	103 (33.01)	0.057*	204 (27.91)	29 (39.73)	0.101*
Cardiac arrest,n(%)	8 (1.67)	6 (1.85)	0.565*	6 (1.30)	8 (2.35)	0.219*	10 (1.72)	4 (1.81)	0.368 *	8 (1.43)	6 (2.46)	0.168*	7 (1.42)	7 (2.24)	0.268*	11 (1.50)	3 (4.11)	0.206*
Cognitive domain score																		
Attention, mean(SD)	0.14 (0.77)	0.15 (0.72)	0.464#	0.12 (0.76)	0.17 (0.73)	0.231#	0.13 (0.74)	0.19 (0.76)	0.269 #	0.17 (0.75)	0.09 (0.75)	0.123#	0.18 (0.74)	0.10 (0.75)	0.202#	0.14 (0.74)	0.21 (0.79)	0.539#
Executive ,mean(SD)	0.17 (0.72)	0.19 (0.66)	0.975#	0.15 (0.74)	0.22 (0.64)	0.229#	0.17 (0.70)	0.22 (0.67)	0.429 #	0.21 (0.71)	0.11 (0.66)	0.023#	0.21 (0.68)	0.14 (0.71)	0.164#	0.17 (0.69)	0.32 (0.76)	0.129#
Language, mean(SD)	0.18 (0.71)	0.16 (0.69)	0.467#	0.16 (0.74)	0.19 (0.64)	0.874#	0.16 (0.72)	0.22 (0.66)	0.443 #	0.23 (0.68)	0.04 (0.73)	<0.001 #	0.17 (0.71)	0.18 (0.69)	0.895#	0.17 (0.70)	0.25 (0.73)	0.299#
Memory, mean(SD)	0.25 (0.67)	0.15 (0.68)	0.047#	0.23 (0.69)	0.18 (0.66)	0.159#	0.24 (0.70)	0.15 (0.62)	0.093 #	0.27 (0.68)	0.08 (0.66)	<0.001 #	0.24 (0.65)	0.17 (0.71)	0.139#	0.21 (0.68)	0.22 (0.65)	0.984#
Visuospatial, mean(SD)	0.13 (1.00)	0.13 (1.02)	0.905#	0.13 (1.04)	0.13 (0.95)	0.866#	0.13 (1.01)	0.12 (1.00)	0.912 #	0.18 (1.02)	0.03 (0.97)	0.055#	0.14 (1.00)	0.12 (1.02)	0.558#	0.12 (0.99)	0.27 (1.15)	0.263#

Age- Age at baseline; Self-reported general health-Subjective health assessment at baseline; mCESD score-Modified CES-D depression score; IADL score-Score on the IADL questions at baseline; No of Rx medications-Total number of prescription medications at baseline; Attention-Standardized attention domain score at baseline; Executive-Standardized executive function domain score at baseline; Language-Standardized language domain score at baseline; Memory-Standardized memory domain score at baseline; Visuospatial-Standardized visuospatial domain score at baseline.

*-For categorical variables: p-value was calculated using Chi-square test (Fisher exact test if sample size is too small); # - For continuous variables: p-value was calculated using Kruskal_Wallis test; bold P represents statistically significant difference between groups.

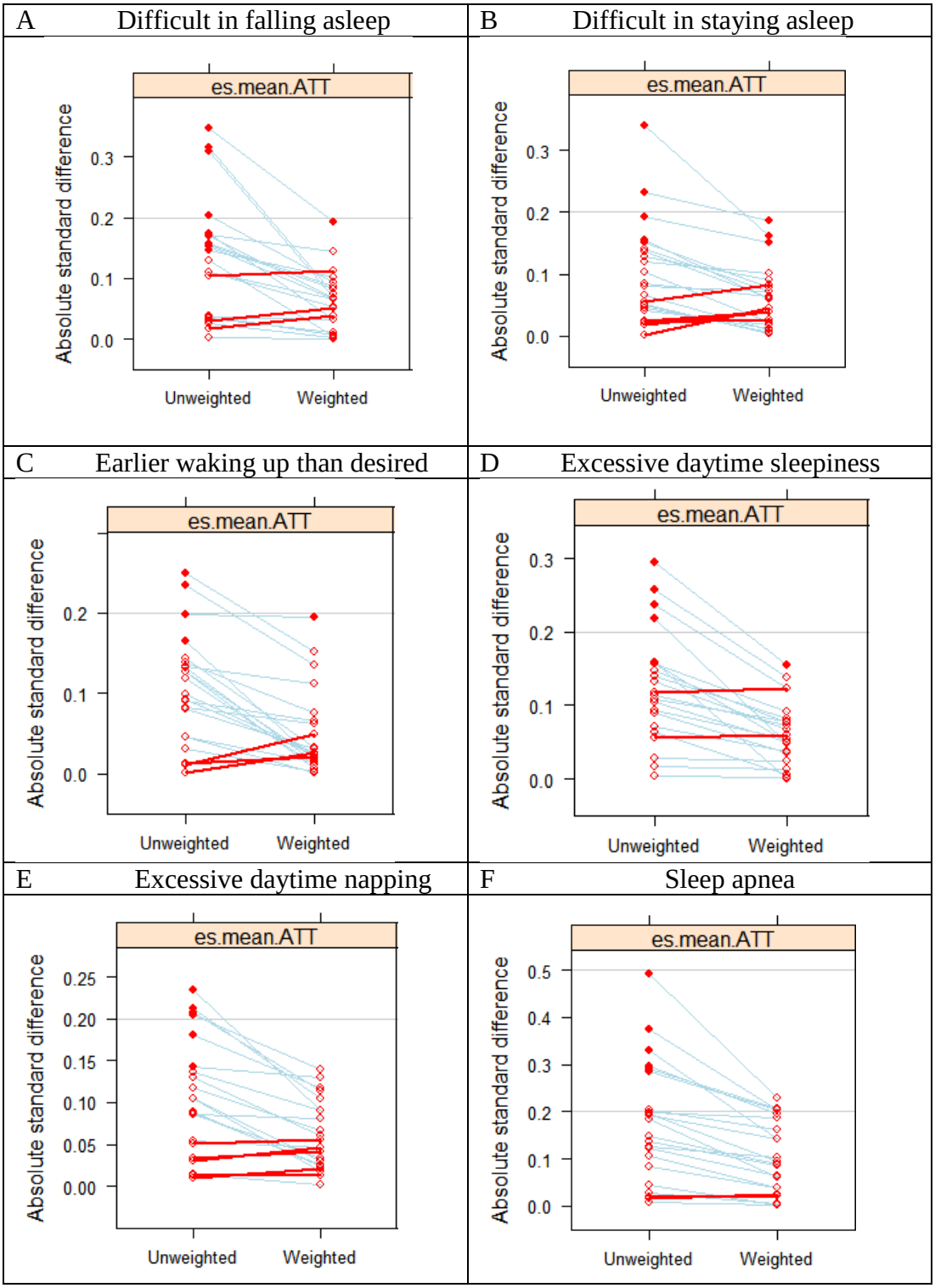
3.2 GENERALIZED BOOSTED MODEL(GBM) PROPENSITY SCORE WEIGHTS

3.2.1 GBM propensity score weights resulting in balanced distributions of covariates

In this study, ASMD values greater than 0.2 were considered moderate imbalance.

Unweighted analysis in Figure 2 and Table 3 showed substantial mean differences between the exposure (with sleep disturbance) and nonexposure (without sleep disturbance) groups, expressed by the imbalanced variables (imbalanced variables for *difficulty falling asleep*: gender, general health, mCES-D score, and hypertension; imbalanced variables for *difficulty staying asleep*: general health, and mCES-D score; imbalanced variables for *waking up earlier than desired*: general health and subjective memory complaints; imbalanced variables for *excessive daytime sleepiness*: subjective memory complaints, systolic BP, baseline language domain score, and baseline memory domain score; imbalanced variables for *excessive daytime napping*: IADL score, subjective memory complaints, diabetes, and heart failure; imbalanced variables for *sleep apnea*: age, general health, number of prescription medications, subjective memory complaints, diabetes, hypertension, and heart failure) with ASMD larger than 0.2.

After applying propensity score weighting, as shown in Figure 2 and Table 3, differences in baseline covariates between the exposure and nonexposure groups diminished substantially. No variable had an effect size (measured by ADMD) over 0.2, indicating balance in covariates after applying GBM-based propensity score weights to the data.



For panels A-F, each dot depicts unweighted (left side, before applying propensity score weights) or weighted (right side, after applying propensity score weights) absolute standard mean difference (ASMD) of a particular covariate at baseline between the exposure (with sleep disturbance) and nonexposure (without sleep disturbance) groups. Open circle and closed circles indicate statistically nonsignificant ($P \geq 0.05$) and significant ($P < 0.05$) mean difference of a

baseline covariate between the exposure group and nonexposure groups, respectively. Corresponding variates were connected with lines. Absolute standard mean value for each variable is equal to absolute weighted group mean difference divided by the unweighted standard deviation for the exposure group and was summarized in Table 3.

Figure 2. Reduced absolute standard mean difference in baseline covariates between older adults with and without sleep disturbance after GBM propensity score weighting

Table 3A. Baseline covariates characteristics and group difference before and after GBM propensity score weighting

	Difficult in falling asleep								Difficult in staying asleep							
	Unweighted					Propensity Score Weighted			Unweighted					Propensity Score Weighted		
	Nonexposure		Exposure			Nonexposure			Nonexposure		Exposure			Nonexposure		
Covariate	**Mean**	**SD**	**Mean**	**SD**	**Effect size**	**Mean**	**SD**	**Effect size**	**Mean**	**SD**	**Mean**	**SD**	**Effect size**	**Mean**	**SD**	**Effect size**
Demographics																
Age(yrs)	**1.704**	**0.646**	**1.775**	**0.678**	**0.104**	**1.699**	**0.628**	**0.112**	**1.717**	**0.658**	**1.754**	**0.663**	**0.055**	**1.699**	**0.647**	**0.083**
Female	**0.608**	**0.489**	**0.744**	**0.437**	**0.310**	**0.714**	**0.453**	**0.069**	**0.639**	**0.481**	**0.695**	**0.461**	**0.121**	**0.649**	**0.478**	**0.101**
Education	**2.323**	**0.700**	**2.250**	**0.660**	**-1.500**	**2.285**	**0.679**	**-0.054**	**2.317**	**0.697**	**2.261**	**0.668**	**-0.085**	**2.303**	**0.665**	**-0.063**
Self-reported geneal health	**2.306**	**0.684**	**2.167**	**0.688**	**-0.203**	**2.231**	**0.683**	**-0.093**	**2.348**	**0.679**	**2.117**	**0.68**	**-0.339**	**2.226**	**0.666**	**-0.160**
Ever smoking	**0.531**	**0.500**	**0.454**	**0.499**	**-0.156**	**0.495**	**0.500**	**-0.083**	**0.527**	**0.500**	**0.463**	**0.499**	**-0.127**	**0.508**	**0.500**	**-0.090**
Ever drinking	**0.867**	**0.340**	**0.877**	**0.329**	**0.030**	**0.860**	**0.347**	**0.050**	**0.864**	**0.343**	**0.880**	**0.326**	**0.049**	**0.878**	**0.328**	**0.006**
Exercise	**0.629**	**0.484**	**0.645**	**0.479**	**0.033**	**0.640**	**0.481**	**0.011**	**0.639**	**0.481**	**0.630**	**0.483**	**-0.018**	**0.649**	**0.478**	**-0.039**
mCESD score >=3	**0.048**	**0.214**	**0.182**	**0.387**	**0.347**	**0.108**	**0.310**	**0.192**	**0.067**	**0.250**	**0.150**	**0.357**	**0.231**	**0.083**	**0.276**	**0.186**
IADL score	**0.081**	**0.274**	**0.093**	**0.290**	**0.039**	**0.094**	**0.293**	**-0.006**	**0.089**	**0.284**	**0.082**	**0.275**	**-0.023**	**0.093**	**0.291**	**-0.039**
No of Rx medications	**0.702**	**0.458**	**0.769**	**0.422**	**0.157**	**0.732**	**0.443**	**0.087**	**0.704**	**0.457**	**0.762**	**0.426**	**0.137**	**0.736**	**0.441**	**0.062**
Subjective memory complaint	**0.615**	**0.487**	**0.685**	**0.465**	**0.152**	**0.655**	**0.476**	**0.066**	**0.613**	**0.487**	**0.683**	**0.466**	**0.15**	**0.647**	**0.478**	**0.078**
Health outcome history																
Cerebro	**0.106**	**0.308**	**0.145**	**0.353**	**0.110**	**0.122**	**0.327**	**0.066**	**0.119**	**0.324**	**0.126**	**0.332**	**0.022**	**0.117**	**0.322**	**0.027**
Diabetes	**0.256**	**0.437**	**0.188**	**0.392**	**-0.174**	**0.201**	**0.401**	**-0.033**	**0.255**	**0.436**	**0.194**	**0.396**	**-0.155**	**0.217**	**0.413**	**-0.060**
Hypertension	**0.654**	**0.476**	**0.784**	**0.412**	**0.315**	**0.753**	**0.432**	**0.075**	**0.680**	**0.467**	**0.742**	**0.438**	**0.141**	**0.710**	**0.454**	**0.072**
Cardiac	**0.356**	**0.479**	**0.441**	**0.497**	**0.171**	**0.370**	**0.483**	**0.143**	**0.350**	**0.477**	**0.446**	**0.498**	**0.193**	**0.371**	**0.484**	**0.150**
Systolic blood pressure	**130.540**	**15.050**	**133.269**	**16.080**	**0.170**	**132.158**	**14.180**	**0.069**	**131.335**	**14.986**	**132.053**	**16.233**	**0.044**	**131.867**	**14.899**	**0.011**
Diastolic blood pressure	**72.369**	**8.512**	**73.426**	**8.158**	**0.130**	**73.380**	**7.963**	**0.006**	**72.616**	**8.471**	**73.038**	**8.267**	**0.051**	**73.003**	**8.358**	**0.004**

Table 3A Continued

Cognitive domain score																
Attention	**0.140**	**0.769**	**0.152**	**0.716**	**0.017**	**0.179**	**0.715**	**-0.038**	**0.124**	**0.762**	**0.172**	**0.729**	**0.065**	**0.179**	**0.712**	**-0.009**
Executive	**0.175**	**0.721**	**0.192**	**0.657**	**0.026**	**0.193**	**0.640**	**-0.002**	**0.154**	**0.737**	**0.219**	**0.635**	**0.103**	**0.205**	**0.639**	**0.022**
Language	**0.183**	**0.714**	**0.159**	**0.687**	**-0.036**	**0.180**	**0.647**	**-0.032**	**0.162**	**0.744**	**0.188**	**0.645**	**0.041**	**0.200**	**0.642**	**-0.018**
Memory	**0.253**	**0.674**	**0.153**	**0.678**	**-0.147**	**0.222**	**0.652**	**-0.102**	**0.235**	**0.687**	**0.182**	**0.662**	**-0.080**	**0.230**	**0.651**	**-0.073**
Visuospatial	**0.131**	**0.996**	**0.133**	**1.020**	**0.002**	**0.134**	**0.971**	**-0.001**	**0.131**	**1.043**	**0.133**	**0.985**	**0.002**	**0.176**	**0.973**	**-0.046**

Table 4B. Baseline covariates characteristics and group difference before and after GBM propensity score weighting

	Earlier waking than desired								**Excessive daytime sleepiness**							
	Unweighted					**Propensity Score Weighted**			**Unweighted**					**Propensity Score Weighted**		
	Nonexposure		**Exposure**			**Nonexposure**			**Nonexposure**		**Exposure**			**Nonexposure**		
Covariate	**Mean**	**SD**	**Mean**	**SD**	**Effect size**	**Mean**	**SD**	**Effect size**	**Mean**	**SD**	**Mean**	**SD**	**Effect size**	**Mean**	**SD**	**Effect size**
Demographics																
Age(yrs)	**1.763**	**0.652**	**1.652**	**0.675**	**-0.166**	**1.663**	**0.616**	**-0.017**	**1.711**	**0.657**	**1.783**	**0.665**	**0.108**	**1.731**	**0.640**	**0.078**
Female	**0.645**	**0.479**	**0.710**	**0.455**	**0.144**	**0.695**	**0.461**	**0.033**	**0.680**	**0.467**	**0.623**	**0.486**	**-0.118**	**0.682**	**0.466**	**-0.122**
Education	**2.312**	**0.683**	**2.244**	**0.690**	**-0.098**	**2.262**	**0.660**	**-0.026**	**2.307**	**0.673**	**2.262**	**0.712**	**-0.063**	**2.266**	**0.684**	**-0.005**
Self-reported geneal health	**2.300**	**0.666**	**2.118**	**0.729**	**-0.250**	**2.229**	**0.671**	**-0.152**	**2.284**	**0.674**	**2.172**	**0.717**	**-0.156**	**2.208**	**0.687**	**-0.050**
Ever smoking	**0.489**	**0.500**	**0.529**	**0.500**	**0.081**	**0.514**	**0.500**	**0.031**	**0.514**	**0.500**	**0.467**	**0.500**	**-0.094**	**0.494**	**0.500**	**-0.053**
Ever drinking	**0.868**	**0.339**	**0.878**	**0.328**	**0.030**	**0.879**	**0.327**	**-0.003**	**0.880**	**0.325**	**0.848**	**0.359**	**-0.089**	**0.861**	**0.346**	**-0.036**
Exercise	**0.630**	**0.483**	**0.652**	**0.478**	**0.046**	**0.652**	**0.477**	**-0.001**	**0.655**	**0.476**	**0.590**	**0.493**	**-0.132**	**0.624**	**0.485**	**-0.068**
mCESD score >=3	**0.082**	**0.275**	**0.154**	**0.362**	**0.198**	**0.083**	**0.276**	**0.195**	**0.088**	**0.283**	**0.135**	**0.343**	**0.139**	**0.107**	**0.309**	**0.083**
IADL score	**0.089**	**0.285**	**0.077**	**0.267**	**-0.046**	**0.081**	**0.273**	**-0.015**	**0.075**	**0.264**	**0.111**	**0.314**	**0.113**	**0.088**	**0.283**	**0.073**

Table 3B Continued

No of Rx medications	**0.714**	**0.452**	**0.769**	**0.422**	**0.132**	**0.763**	**0.426**	**0.015**	**0.725**	**0.447**	**0.738**	**0.441**	**0.029**	**0.748**	**0.434**	**-0.024**
Subjective memory complaint	**0.614**	**0.487**	**0.719**	**0.450**	**0.234**	**0.658**	**0.475**	**0.136**	**0.611**	**0.488**	**0.717**	**0.451**	**0.236**	**0.662**	**0.474**	**0.123**
Health outcome history																
Cerebro	**0.122**	**0.327**	**0.122**	**0.328**	**0.001**	**0.114**	**0.318**	**0.025**	**0.121**	**0.327**	**0.123**	**0.329**	**0.005**	**0.122**	**0.328**	**0.002**
Diabetes	**0.242**	**0.429**	**0.195**	**0.397**	**-0.119**	**0.199**	**0.399**	**-0.011**	**0.227**	**0.419**	**0.234**	**0.424**	**0.016**	**0.228**	**0.420**	**0.012**
Hypertension	**0.691**	**0.462**	**0.747**	**0.436**	**0.127**	**0.744**	**0.437**	**0.006**	**0.686**	**0.465**	**0.754**	**0.432**	**0.158**	**0.754**	**0.431**	**-0.001**
Cardiac	**0.372**	**0.484**	**0.439**	**0.497**	**0.134**	**0.383**	**0.487**	**0.112**	**0.382**	**0.486**	**0.410**	**0.493**	**0.056**	**0.380**	**0.486**	**0.060**
Systolic blood pressure	**131.587**	**15.578**	**131.778**	**15.407**	**0.012**	**132.104**	**13.713**	**-0.021**	**130.577**	**15.200**	**134.078**	**16.001**	**0.219**	**133.304**	**14.388**	**0.048**
Diastolic blood pressure	**72.527**	**8.638**	**73.502**	**7.639**	**0.128**	**73.645**	**7.727**	**-0.019**	**72.612**	**8.306**	**73.213**	**8.557**	**0.070**	**73.525**	**7.977**	**-0.036**
Cognitive domain score																
Attention	**0.126**	**0.742**	**0.194**	**0.763**	**0.090**	**0.144**	**0.712**	**0.066**	**0.168**	**0.747**	**0.090**	**0.747**	**-0.105**	**0.128**	**0.670**	**-0.051**
Executive	**0.166**	**0.704**	**0.222**	**0.675**	**0.082**	**0.180**	**0.621**	**0.062**	**0.213**	**0.710**	**0.110**	**0.657**	**-0.157**	**0.169**	**0.621**	**-0.091**
Language	**0.157**	**0.721**	**0.217**	**0.655**	**0.093**	**0.202**	**0.610**	**0.024**	**0.230**	**0.684**	**0.042**	**0.730**	**-0.258**	**0.142**	**0.620**	**-0.138**
Memory	**0.236**	**0.697**	**0.150**	**0.616**	**-0.139**	**0.197**	**0.603**	**-0.075**	**0.272**	**0.675**	**0.077**	**0.661**	**-0.294**	**0.179**	**0.598**	**-0.154**
Visuospatial	**0.135**	**1.009**	**0.125**	**0.997**	**-0.010**	**0.076**	**0.932**	**0.049**	**0.176**	**1.016**	**0.032**	**0.974**	**-0.148**	**0.107**	**0.909**	**-0.077**

Table 5C. Baseline covariates characteristics and group difference before and after GBM propensity score weighting

	Excessive daytime napping								**Sleep apnea**							
	Unweighted					**Propensity Score Weighted**			**Unweighted**					**Propensity Score Weighted**		
	Nonexposure		**Exposure**			**Nonexposure**			**Nonexposure**		**Exposure**			**Nonexposure**		
Covariate	**Mean**	**SD**	**Mean**	**SD**	**Effect size**	**Mean**	**SD**	**Effect size**	**Mean**	**SD**	**Mean**	**SD**	**Effect size**	**Mean**	**SD**	**Effect size**
Demographics																

Table 3C Continued

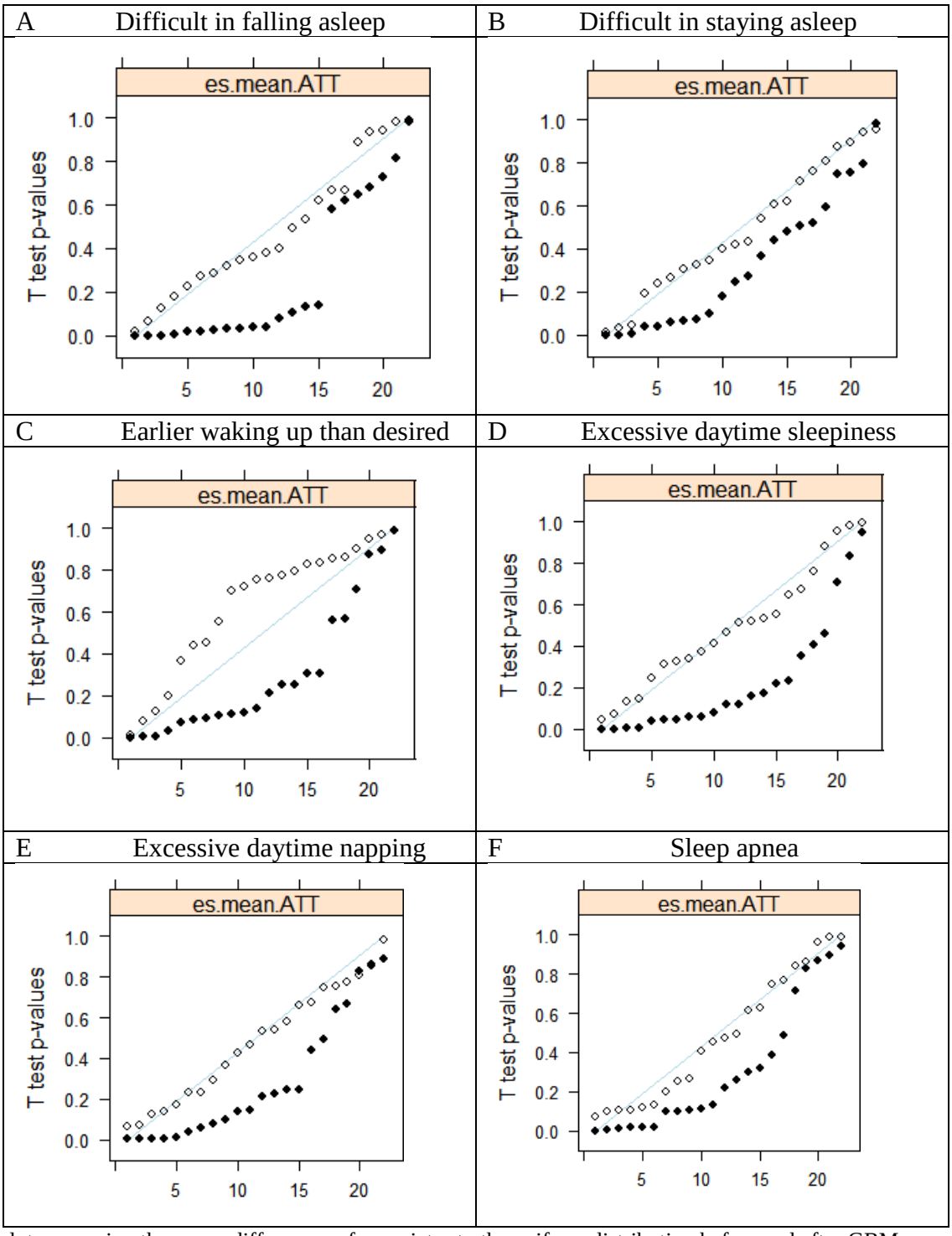
Age(yrs)	**1.711**	**0.671**	**1.766**	**0.671**	**0.085**	**1.744**	**0.656**	**0.034**	**1.750**	**0.659**	**1.562**	**0.645**	**-0.291**	**1.695**	**0.654**	**-0.207**
Female	**0.697**	**0.460**	**0.609**	**0.489**	**-0.180**	**0.666**	**0.472**	**-0.117**	**0.672**	**0.470**	**0.575**	**0.498**	**-0.194**	**0.647**	**0.478**	**-0.144**
Education	**2.285**	**0.676**	**2.308**	**0.700**	**0.033**	**2.278**	**0.681**	**0.042**	**2.294**	**0.676**	**2.288**	**0.772**	**-0.008**	**2.289**	**0.695**	**-0.002**
Self-reported geneal health	**2.283**	**0.672**	**2.199**	**0.712**	**-0.118**	**2.247**	**0.677**	**-0.068**	**2.269**	**0.682**	**2.055**	**0.724**	**-0.296**	**2.196**	**0.684**	**-0.196**
Ever smoking	**0.474**	**0.500**	**0.542**	**0.499**	**0.136**	**0.496**	**0.500**	**0.091**	**0.498**	**0.500**	**0.521**	**0.503**	**0.045**	**0.520**	**0.500**	**0.001**
Ever drinking	**0.878**	**0.328**	**0.859**	**0.349**	**-0.055**	**0.875**	**0.331**	**-0.047**	**0.870**	**0.336**	**0.877**	**0.331**	**0.020**	**0.870**	**0.337**	**0.022**
Exercise	**0.620**	**0.486**	**0.660**	**0.474**	**0.085**	**0.622**	**0.485**	**0.081**	**0.628**	**0.484**	**0.712**	**0.456**	**0.185**	**0.684**	**0.465**	**0.061**
mCESD score >=3	**0.085**	**0.277**	**0.131**	**0.338**	**0.142**	**0.087**	**0.282**	**0.131**	**0.097**	**0.296**	**0.151**	**0.360**	**0.149**	**0.118**	**0.323**	**0.090**
IADL score	**0.059**	**0.236**	**0.128**	**0.335**	**0.207**	**0.090**	**0.286**	**0.114**	**0.079**	**0.270**	**0.151**	**0.360**	**0.198**	**0.084**	**0.277**	**0.186**
No of Rx medications	**0.707**	**0.455**	**0.763**	**0.426**	**0.130**	**0.737**	**0.441**	**0.061**	**0.714**	**0.452**	**0.877**	**0.331**	**0.491**	**0.801**	**0.399**	**0.228**
Subjective memory complaint	**0.602**	**0.490**	**0.708**	**0.455**	**0.234**	**0.667**	**0.472**	**0.091**	**0.635**	**0.482**	**0.726**	**0.449**	**0.203**	**0.653**	**0.476**	**0.162**
Health outcome history																
Cerebro	**0.120**	**0.325**	**0.125**	**0.331**	**0.015**	**0.126**	**0.332**	**-0.002**	**0.118**	**0.322**	**0.164**	**0.373**	**0.125**	**0.126**	**0.332**	**0.103**
Diabetes	**0.193**	**0.395**	**0.285**	**0.452**	**0.204**	**0.222**	**0.416**	**0.139**	**0.212**	**0.409**	**0.397**	**0.493**	**0.376**	**0.297**	**0.457**	**0.203**
Hypertension	**0.691**	**0.463**	**0.731**	**0.444**	**0.089**	**0.720**	**0.449**	**0.024**	**0.695**	**0.461**	**0.822**	**0.385**	**0.330**	**0.767**	**0.423**	**0.142**
Cardiac	**0.350**	**0.477**	**0.455**	**0.499**	**0.212**	**0.403**	**0.491**	**0.105**	**0.378**	**0.485**	**0.521**	**0.503**	**0.284**	**0.417**	**0.493**	**0.205**
Systolic blood pressure	**131.447**	**14.988**	**131.942**	**16.347**	**0.030**	**132.694**	**15.114**	**-0.046**	**131.602**	**15.540**	**132.014**	**15.436**	**0.027**	**131.924**	**15.298**	**0.006**
Diastolic blood pressure	**72.634**	**8.520**	**73.048**	**8.168**	**0.051**	**73.497**	**8.188**	**-0.055**	**72.711**	**8.345**	**73.630**	**8.758**	**0.105**	**73.279**	**8.356**	**0.040**
Cognitive domain score																
Attention	**0.175**	**0.743**	**0.096**	**0.754**	**-0.104**	**0.110**	**0.732**	**-0.018**	**0.138**	**0.743**	**0.206**	**0.793**	**0.085**	**0.176**	**0.787**	**0.038**
Executive	**0.211**	**0.683**	**0.136**	**0.714**	**-0.105**	**0.157**	**0.648**	**-0.030**	**0.168**	**0.688**	**0.317**	**0.763**	**0.196**	**0.245**	**0.706**	**0.095**
Language	**0.170**	**0.713**	**0.178**	**0.689**	**0.011**	**0.163**	**0.692**	**0.022**	**0.165**	**0.701**	**0.254**	**0.726**	**0.122**	**0.207**	**0.709**	**0.064**
Memory	**0.237**	**0.654**	**0.175**	**0.710**	**-0.087**	**0.192**	**0.692**	**-0.025**	**0.212**	**0.679**	**0.222**	**0.654**	**0.017**	**0.206**	**0.630**	**0.025**
Visuospatial	**0.137**	**0.997**	**0.124**	**1.019**	**-0.013**	**0.111**	**1.008**	**0.013**	**0.118**	**0.989**	**0.274**	**1.147**	**0.136**	**0.175**	**1.032**	**0.087**

Comparing covariates pretreatment differences between groups before and after GBM propensity score weighting.
Exposure-subjects with sleep problem; Nonexposure-subjects without sleep problem; effect sizes= absolute group mean difference divided by the standard deviation for the exposed group.

3.2.2 GBM propensity score weights caused covariates follow uniform distribution

P-Values from a random experiment have a uniform distribution (45-degree line) due to balanced covariates between groups.

Solid diamonds in Figure 3 indicated the p-values for the t-tests of equal group differences in covariates before weighting. Severe deviation of the p-value below the diagonal suggested statistically significant differences between groups in covariates. After applying the GBM-based propensity score weighting (open squares in Figure 3), the p-values for the t-tests of equal group differences in the covariate means increased suggesting improved balance in covariates.



Q-Q plot comparing the group differences of covariates to the uniform distribution before and after GBM propensity score weights. Solid diamonds depict p-values for t-test of group differences on covariates before weighting. Open squares depict p-values after weighting. Blue 45-degree line depicts uniform distribution of random experiment. Each dot depicts a particular covariate at baseline.

Figure 3. Improved randomness after GBM propensity score weights

3.3 CAUSAL IMPACT OF SLEEP DEPRIVATION ON COGNITIVE DECLINE

Using the honest causal tree method, we examined the causal effect of sleep deprivation (difficulty falling asleep, difficulty staying asleep, waking up earlier than desired, excessive daytime sleepiness, excessive daytime napping, and sleep apnea) on cognitive decline (including attention, executive, memory, language, and visuospatial domains), after adjusting for baseline demographic, health-related conditions, and baseline standardized cognitive domain scores. The results suggested that having difficulty falling asleep, difficulty staying asleep, excessive daytime sleepiness, and excessive daytime napping led to faster cognitive decline in the memory and visuospatial domains. These causal relationship also showed different effects among people with different sociodemigraphic or baseline health conditions, including age, gender, self-reported general health, systolic BP, diastolic BP, exercise, and baseline cognitive domain scores.

3.3.1 Randomly splitting samples into training set and estimation set

Before building a causal tree model, observations were randomly divided into approximately 50% for the training and 50% for the estimation sets. The training set was used to build a causal tree model, and the estimation set was used to evaluate the accuracy of the model built.

As shown in Table 4, no statistically significant difference in baseline covariates (except age and subjective memory complaint) between the training and the estimation samples, suggesting the similarity of baseline characteristics between these two sub-samples.

Table 6. Comparing baseline covariates between training set and estimation set

Baseline variable	**Testing set** (n=389)	**Training set** (n=415)	**P value**
Age in years, n(%)			**0.013**
65-74	146(37.53)	165(39.76)	
75-84	183(47.04)	214(51.57)	
85+	60(15.42)	36(8.67)	
Female gender, n(%)	265(68.12)	268(64.58)	0.288
Education, n(%)			0.123
<HS	56(14.40)	49(11.81)	
HS	182(46.79)	176(42.41)	
>HS	151(38.82)	190(45.78)	
General self-reported health, n(%)			0.521
Poor/Fair	61(15.68)	54(13.01)	
Good	180(46.27)	193(46.51)	
Very good	148(38.05)	168(40.48)	
Ever smoking, n(%)	195(50.13)	207(49.88)	0.944
Ever drinking, n(%)	339(87.15)	361(86.99)	0.947
Exercise, n(%)	244(62.72)	267(64.34)	0.635
mCES-D score >=3, n(%)	44(11.31)	38(9.16)	0.313
IADL score >0, n(%)	39(10.03)	30(7.23)	0.157
No of Rx medications >=3, n(%)	275(70.69)	311(74.94)	0.176
Subjective memory complaint >0, n(%)	267(68.64)	250(60.24)	**0.013**
Health outcome history			
Systolic blood pressure, mean(SD)	131.13(15.16)	132.12(15.85)	0.367
Diastolic blood pressure, mean(SD)	72.40(8.43)	73.16(8.33)	0.200
Stroke, n(%)	17(4.37)	9(2.17)	0.109
TIA, n(%)	43(11.05)	34(8.19)	0.168
Diabetes, n(%)	78(20.05)	106(25.54)	0.064
Hypertension, n(%)	270(69.41)	298(71.81)	0.455
Heart attack, n(%)	55(14.14)	54(13.01)	0.641
Congestive heart failure, n(%)	26(6.68)	27(6.51)	0.583

Table 4 Continued

Irregular heartbeat, n(%)	102(26.22)	131(31.57)	0.151
Cardiacrrest, n(%)	6(1.54)	8(1.93)	0.690
Sleep disturbance			
Difficult in falling sleep, n(%)	163(41.90)	161(38.80)	0.369
Difficult in staying sleep, n(%)	161(41.39)	180(43.37)	0.569
Earlier waking up than desired, n(%)	102(26.22)	119(28.67)	0.436
Excessive daytime sleepiness, n(%)	118(30.33)	126(30.36)	0.993
Excessive daytime napping, n(%)	146(37.53)	166(40.00)	0.473
Sleep apnea, n(%)	33(8.48)	40(9.64)	0.569
Cognitive domain score			
Attention, mean(SD)	0.11(0.72)	0.18(0.77)	0.160
Executive function, mean(SD)	0.14(0.73)	0.22(0.66)	0.115
Language, mean(SD)	0.14(0.74)	0.20(0.66)	0.269
Memory, mean(SD)	0.17(0.73)	0.25(0.63)	0.087
Visuospatial, mean(SD)	0.08(1.02)	0.18(0.99)	0.152

Age- Age at baseline; Self-reported general health-Subjective Health Assessment at baseline; mCES-D score-Modified CES-D depression score; IADL score-Score on the IADL questions at baseline; No of Rx medications-Total number of prescription medications at baseline; Difficult in falling asleep-take longer than a half an hour to fall asleep; Difficult in staying asleep-wake up during the night and find that it takes more than a half hour to go back to sleep; Earlier waking up than desired -wake up earlier than you want to and find that you can't go back to sleep; Excessive daytime sleepiness- fall asleep while actively doing something during the day; Excessive daytime napping -doze off or take a nap most days of the week; Attention-Standardized attention domain score at baseline; Executive-Standardized executive function domain score at baseline; Language-Standardized language domain score at baseline; Memory-Standardized memory domain score at baseline; Visuo-spatial-Standardized visuospatial domain score at baseline.

p-value was calculated using Chi-square test (categorical variables) or t-test or Kruskal-Wallis test (continuous variables); bold P represents statistically significant difference between groups; N=804 total participants.

3.3.2 Causal effect of difficulty falling asleep on cognitive decline

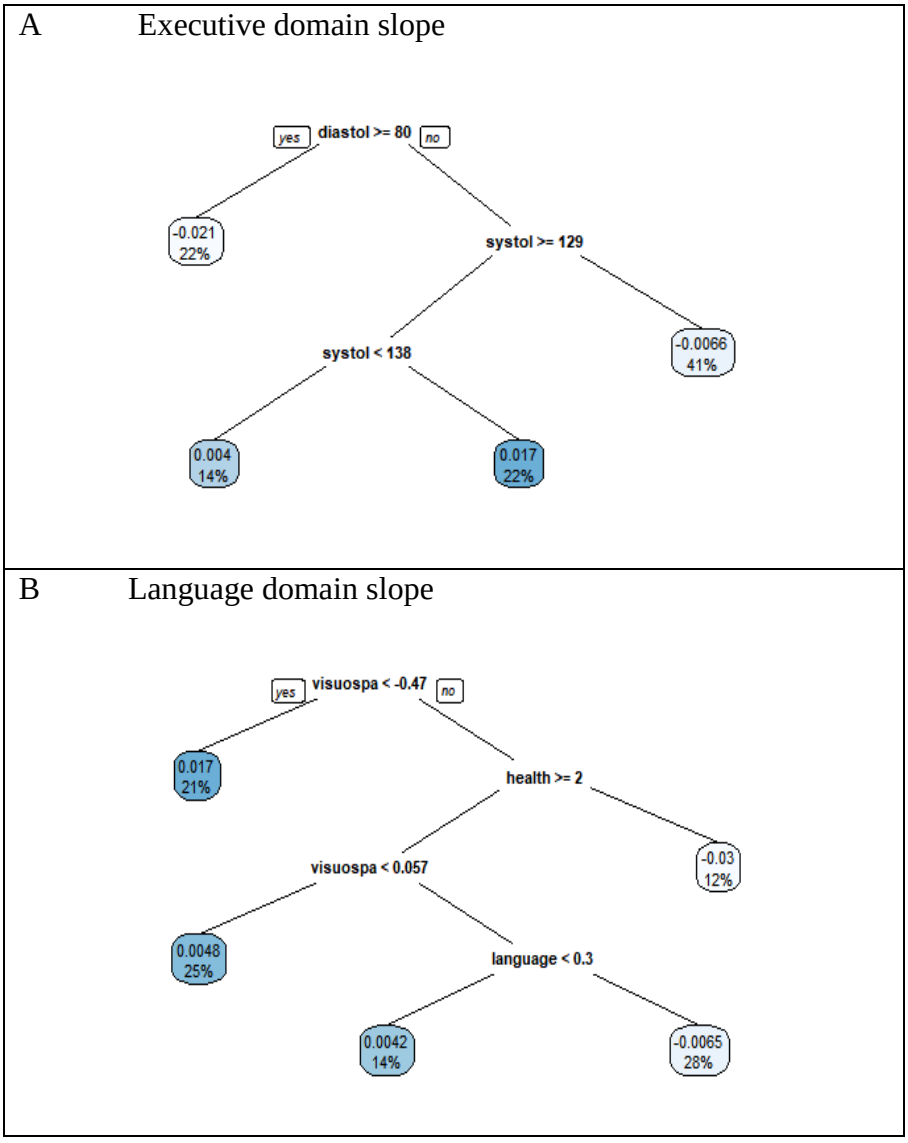
Figure 4 showed heterogeneous effect of difficulty falling asleep on cognitive performance that having difficulty falling asleep led to cognitive deficits in executive, language, memory and visuospatial functions during the 9-year follow up. For each node, a positive number indicates a faster decline among those who had difficulty falling asleep than that among those who did not have difficulty falling asleep. A negative number indicates a slower decline among those who had difficulty falling asleep.

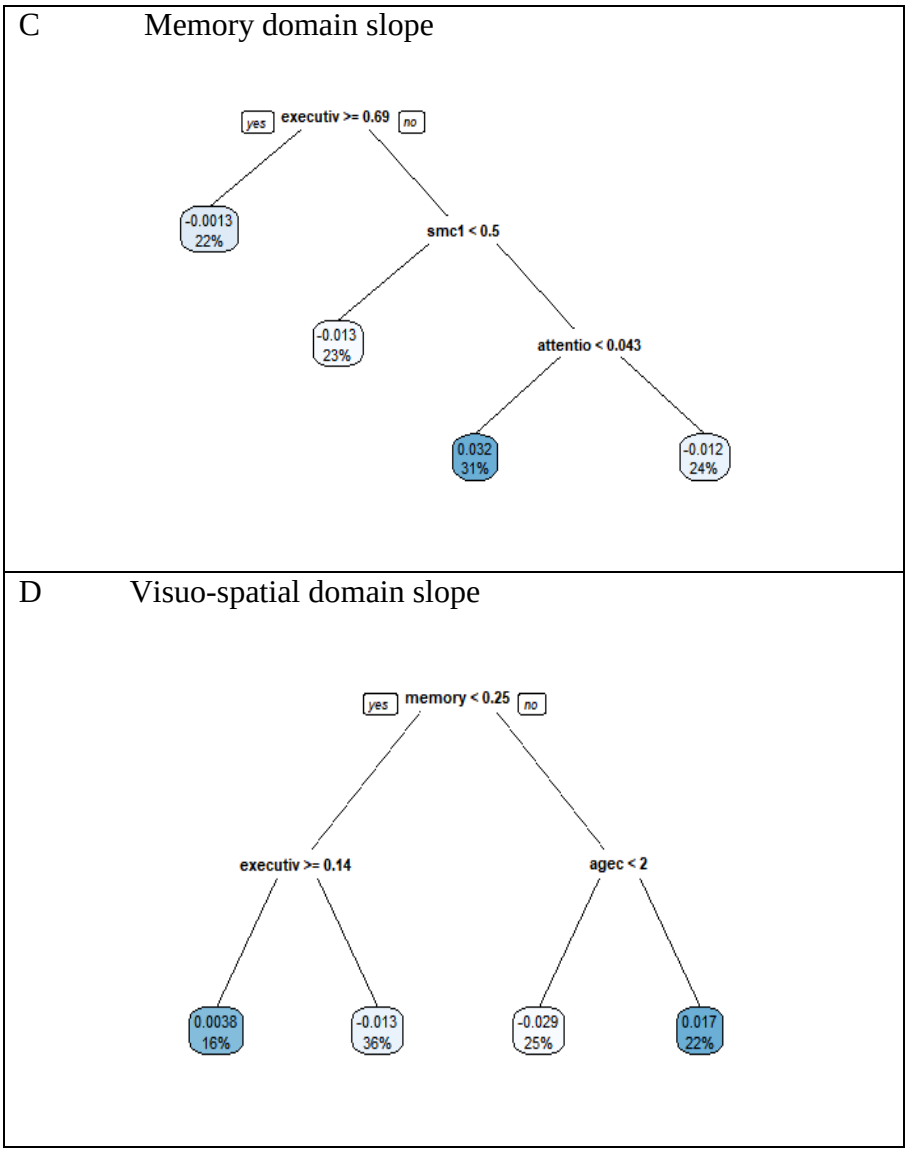
Figure 4A showed that having difficulty falling asleep caused greater executive function decline. Comparing the slopes of 9-yaer executive function domain scores between those who had and did not have difficulty falling asleep, executive function declined faster among individuals who had higher diastolic BP (≥80).

Figure 4B showed that having difficulty falling asleep caused language cognitive decline in older adults over time. Comparing the slopes of language function domain scores between those who had and did not have difficulty falling asleep, faster language decay was associated with those who had higher visuospatial domain score (≥ -0.47) and reported poor or fair general health conditions.

Figure 4C showed that having difficulty falling asleep led to a worsening memory in elderly over time. Comparing the slopes of memory function domain scores between those who had and did not have difficulty falling asleep, rapid memory loss was associated with those who had lower baseline executive score (<0.69) and no subjective memory complaint, or associated with elderly having subjective memory complaint and baseline attention score ≥0.043.

Figure 4D showed that having difficulty falling asleep caused visuospatial cognitive decline. Comparing the slopes of visuospatial function domain scores between those who had and did not have difficulty falling asleep, visuospatial function declines faster among individuals who either had baseline memory domain score <0.25 and executive function score <0.14 or had baseline memory domain score ≥0.25 but were younger (65-74 years old).





Causal tree algorithm to examine the heterogeneity of difficulty falling asleep on cognitive domain scores decline. Within each node, the top number represents the mean difference in the slopes of cognitive domain scores between those who had and did not have difficult in falling asleep, and the bottom number shows the percentage of total affected participants. Positive number (dark blue) indicates faster slope changes of cognitive function domain scores among those who had difficulty falling asleep than that among those without sleep disorder. Negative numbers (light blue) means slower slope changes of cognitive function domain scores among those who had difficulty falling asleep than slope changes among those without sleep disorder.

Abbreviations: diastole-Diastolic blood pressure; systol-Systolic blood pressure; visuospa- Standardized visuo-spatial domain scores at baseline; language- Standardized language domain scores at baseline; exectiv- Standardized executive domain scores at baseline; attentio- Standardized attention domain scores at baseline; memory- Standardized memory domain scores at baseline; health-Self-reported general health condition at baseline; smc1- Score on the subjective memory complaint scale at baseline; agec-Age at baseline in years.

Figure 4. Causal effect of falling asleep on cognitive domain slope change

3.3.3 Causal effect of difficulty staying asleep on cognitive decline

Figure 5 showed heterogenous effect of difficulty staying asleep on cognitive performance. Having difficult in staying asleep resulted in cognitive decline in attention domain, executive domain, memory domain, and visuo-spatial domain in older adults during the 9-year follow up. This causal relationship was moderated by baseline cognitive domain scores, baseline systolic BP, and gender.

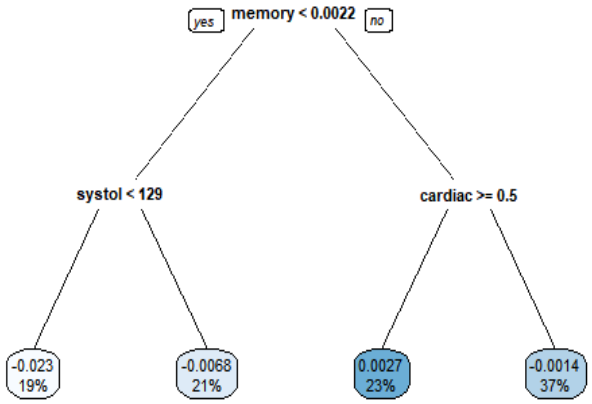
Figure 5A showed that having difficulty staying asleep caused attention cognitive decline. Comparing the slopes of attention function domain scores between those who had and did not have difficulty staying asleep, rapid attention decline was associated with those who had lower baseline memory score (<0.0022) and systolic BP<129.

Figure 5B showed that having difficult in staying asleep impaired executive performance over time. Comparing the slopes of executive function domain scores between those who had and did not have difficulty staying asleep, rapid executive impairment was associated with those who had higher baseline language domain scores (≥-0.2).

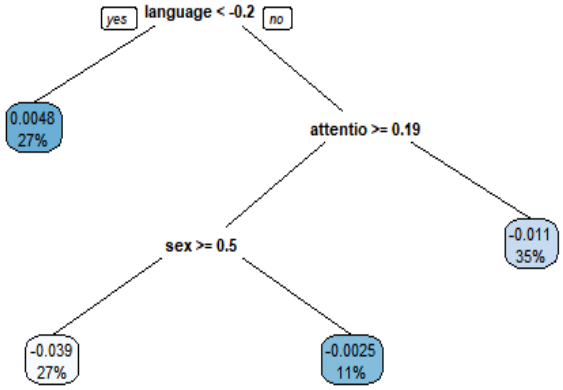
As shown in Figure 5C, comparing the slopes of memory function domain scores between those who had and did not have difficulty staying asleep, faster memory loss was associated with those who had higher baseline executive domain score ≥0.79.

Figure 5D showed that having difficulty staying asleep impaired visuospatial function. Comparing the slopes of visuospatial function domain scores between those who had and did not have difficulty staying asleep, faster visuospatial degeneration was associated with those who had higher baseline memory score (≥0.68).

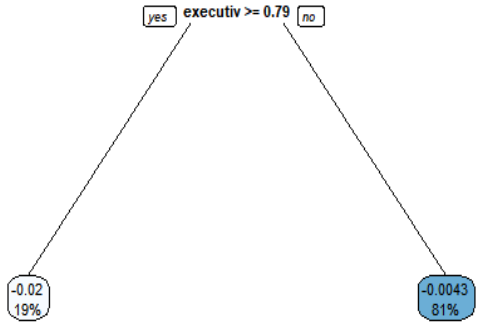
A Attention domain slope

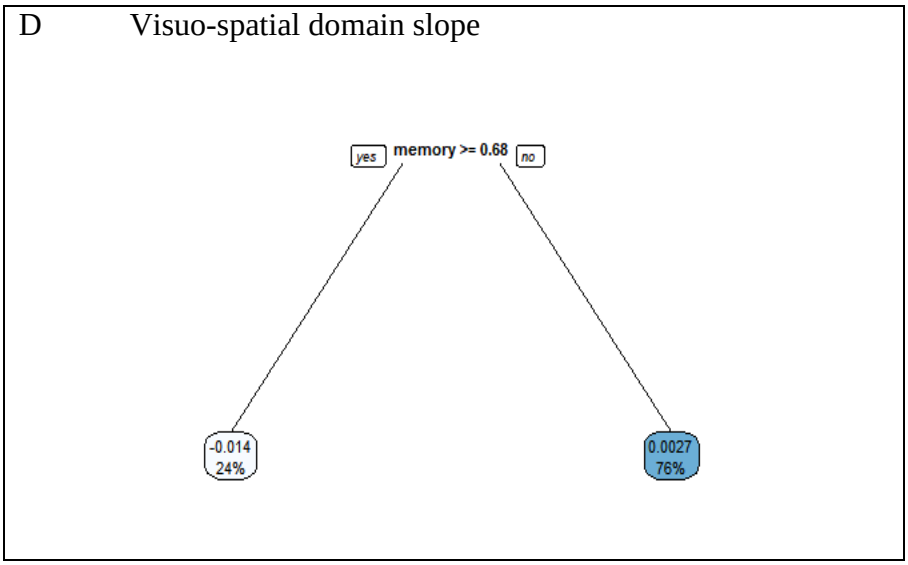


B Executive domain slope



C Memory domain slope





Causal tree algorithm to examine the heterogeneity of difficulty staying asleep on cognitive domain slope decline. Within each node, the top number represents the mean difference in the slope of cognitive domain scores between those who had and did not have difficult in staying asleep, and the bottom shows the percentage of total affected participants. Positive number indicates faster slope changes of cognitive function domain scores among those had difficulty falling asleep than slope changes among those without sleep disorder. Negative numbers mean slower slope changes of cognitive function domain scores among those had difficulty falling asleep than slope changes among those without sleep disorder.
Abbreviations: systol-Systolic blood pressure; language- Standardized language domain scores at baseline; exectiv-Standardized executive domain scores at baseline; attentio- Standardized attention domain scores at baseline; memory- Standardized memory domain scores at baseline; cardiac-Cardiovascular disease at baseline, including heart attack, congestive heart failure, irregular heartbeat, atrial fibrillation, heart racing, palpitations, cardiac arrhythmia, and cardiac arrest; sex-Gender at baseline.

Figure 5. Causal effect of difficulty staying asleep on cognitive domain slope change

3.3.4 Causal effect of excessive daytime sleepiness on cognitive decline

Figure 6 showed the heterogeneous effects of excessive daytime sleepiness on cognitive performance. Having excessive daytime sleepiness caused cognitive decline in attention, language, memory, and visuo-spatial in older adults during the 9-year follow-up. This causal relationship was moderated by baseline characteristics, including cognitive domain scores, diastolic BP, self-reported general health condition, and exercise.

Figure 6A suggested that having excessive daytime sleepiness impaired attention function in older adults over time. This causal relationship was moderated by baseline memory domain scores, baseline diastolic BP, self-reported general health condition, and baseline attention domain scores.

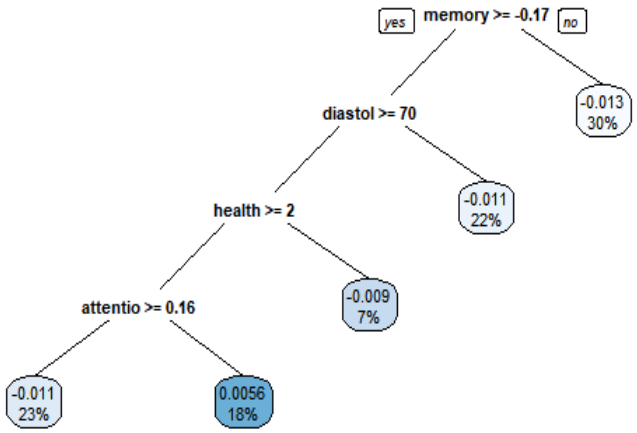
As shown in Figure 6A, having excessive daytime sleepiness induced attention cognitive decline in older adults over time. Comparing the slopes of attention function domain scores between those who had and did not have excessive daytime sleepiness, steeper attention decline was associated with individuals who had lower baseline memory domain scores (<-0.17), or those who had higher memory domain scores (≥-0.17) but lower diastolic BP (<70), or those who had higher baseline memory score (≥-0.17), higher diastolic BP ≥70, higher attention scores (≥ 0.16), and good general health condition.

Figure 6B showed that having excessive daytime sleepiness resulted in language deficit in older adults. Comparing the slopes of language function domain scores between those who had and did not have excessive daytime sleepiness, rapid language decay was associated with those who had less exercise.

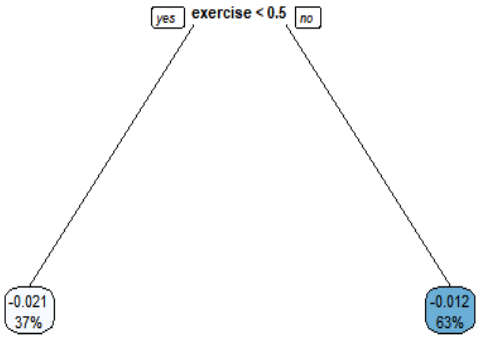
Figure 6C showed that having excessive daytime sleepiness resulted in memory loss. Comparing the slopes of memory function domain scores between those who had and did not have excessive daytime sleepiness, faster memory loss was associated with older adults who had lower baseline memory domain score (<-0.17).

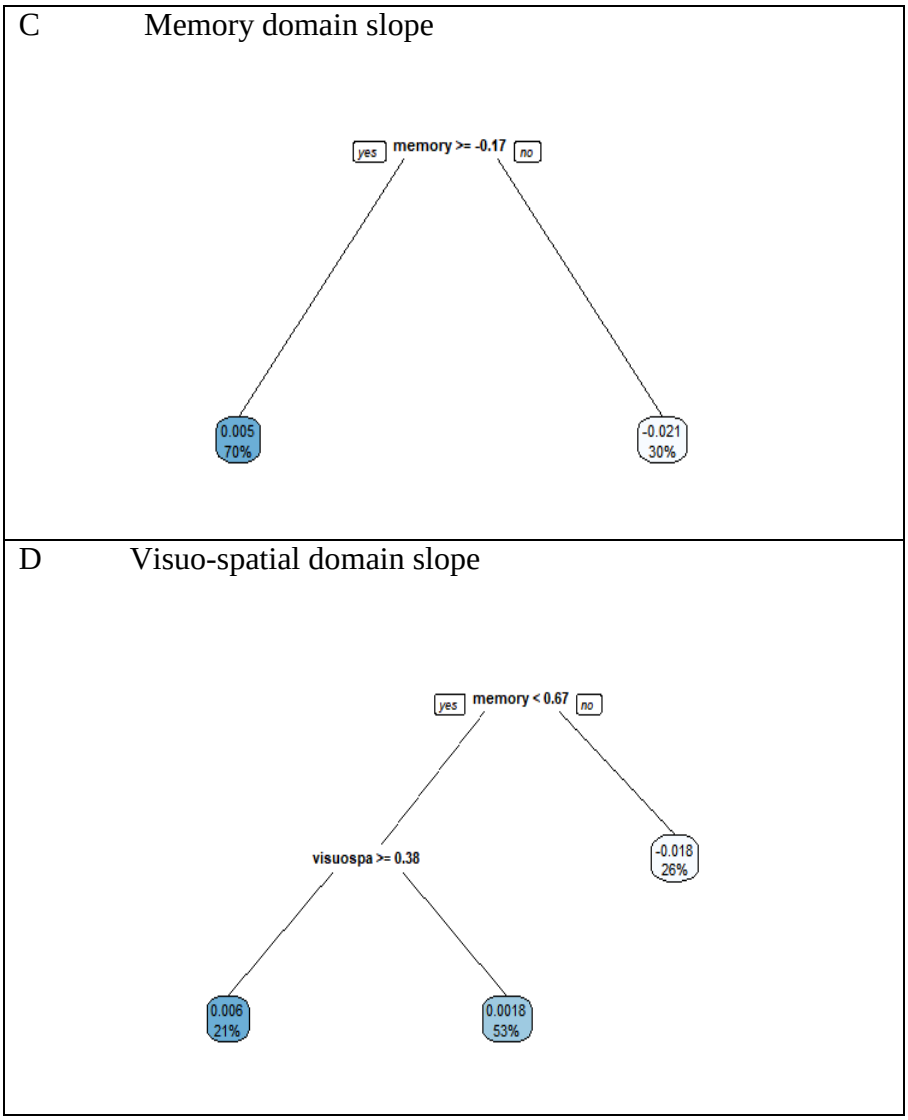
Figure 6D showed that having excessive daytime sleepiness damaged visuospatial function in older adults. Comparing the slopes of visuospatial function domain scores between those who had and did not have excessive daytime sleepiness, faster visuospatial decline was associated with those who had higher baseline memory score (≥0.67).

A Attention domain slope



B Language domain slope





Causal tree algorithm to examine the heterogeneity of excessive daytime sleepiness on cognitive domain slope decline. Within each node, the top number represents the mean difference in the slope of cognitive domain scores between those who had and did not have excessive daytime sleepiness, and the bottom shows the percentage of total affected participants. Positive number indicates faster slope changes of cognitive function domain scores among those who had excessive daytime sleepiness than slope changes among those without sleep disorder. Negative numbers mean slower slope changes of cognitive function domain scores among those had excessive daytime sleepiness than slope changes among those without sleep disorder.
Abbreviations: diastole-Diastolic blood pressure; visuospa- Standardized visuospatial domain scores at baseline; attentio- Standardized attention domain scores at baseline; memory- Standardized memory domain scores at baseline; health-Self-reported general health condition at baseline; exercise-Exercise at baseline.

Figure 6. Causal effect of excessive daytime sleepiness on cognitive domain slope change

3.3.5 Causal effect of excessive daytime napping on cognitive decline

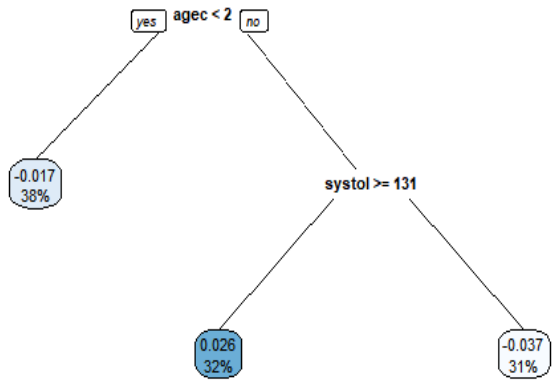
Figure 7 showed heterogeneous effect of excessive daytime napping on cognitive function. Having excessive daytime napping caused cognitive decline in language, memory and visuo-spatial in older people during the 9-year follow-up. The cognitive decline rate was moderated by age, baseline systolic BP, self-reported general health condition, baseline executive domain scores, baseline visuo-spatial domain scores, baseline language domain scores, and gender.

Figure 7A showed that having excessive daytime napping led to language degeneration. Comparing the slopes of language function domain scores between those who had and did not have excessive daytime napping, faster language decay was associated with those who were 65-74 years old or were greater than 75 years old with baseline systolic BP <131.

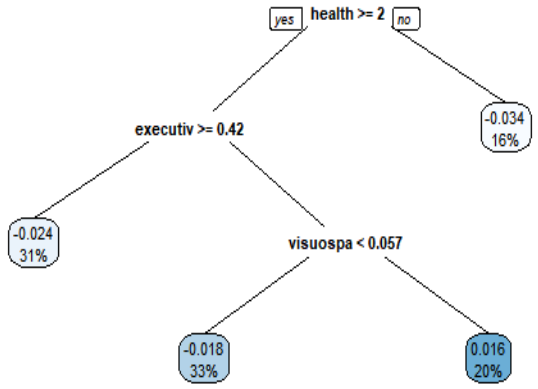
Figure 7B showed that having excessive daytime napping resulted in memory loss in older adults over time. Comparing the slopes of memory function domain scores between those who had and did not have excessive daytime napping, rapid memory loss was associated with those who reported poor/fair health condition, or those who had reported good health condition and had higher baseline executive score (≥0.42), or those who reported good health condition and had lower executive score (<0.42) and lower visuospatial scores (<0.057).

Figure 7C showed that having excessive daytime napping resulted in visuospatial decline in older people with age. Comparing the slopes of visuospatial function domain scores between those who had and did not have excessive daytime napping, faster visuospatial degeneration was associated with women who had higher baseline language domain scores (≥0.23).

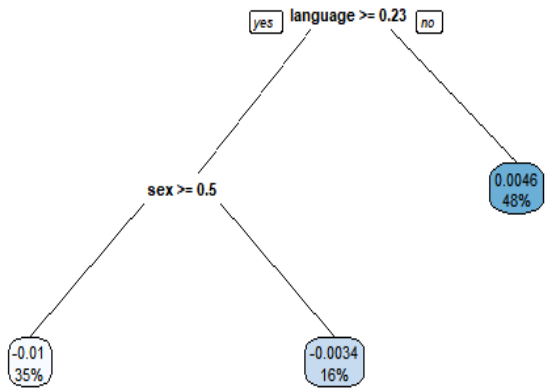
A Language domain slope



B Memory domain slope



C Visuo-spatial domain slope



Causal tree algorithm to examine the heterogeneity of excessive daytime napping on cognitive domain slope decline. Within each node, the top number represents the mean difference in the slope of cognitive domain scores between those who had and did not have excessive daytime napping, and the bottom shows the percentage of total affected participants. Positive number indicates faster slope changes of cognitive function domain scores among those who had excessive daytime napping than slope changes among those without sleep disorder. Negative numbers mean slower slope changes of cognitive function domain scores among those had excessive daytime napping than slope changes among those without sleep disorder.
Abbreviations: systol-Systolic blood pressure; visuospa- Standardized visuospatial domain scores at baseline; language- Standardized language domain scores at baseline; exectiv- Standardized executive domain scores at baseline; health-Self-reported general health condition at baseline; agec-Age at baseline in years; sex-Gender at baseline.

Figure 7. Causal effect of excessive daytime napping on cognitive domain slope change

4.0 DISCUSSION

Aging is accompanied with cognitive dysfunction to some degree, which interferes with older adults' life and functioning [31, 32]. Older people having cognitive impairment often self-reported sleep disturbance. Although the relationship between sleep deprivation and cognitive impairment in the elderly has been explored for decades, the association is conflicting. The inconsistent findings might due to the fact that traditional regression models, which focus on the average exposure effect across the whole sample, fail to detect the variation in this effect among the study individuals. By using the novel causal tree model [26], we were able to examine the causal relationship between sleep disturbance and cognitive impairment in subgroups according to their characteristics. In addition, the modifiable moderators examined in this study (e.g., age, gender, education, self-reported genera health condition, systolic BP, diastolic BP, smoking, alcohol consumption, cerebrovascular disease, cardiac disease, exercise, total number of prescription medications, IADL score, subjective memory complaint scale score, modified CES-D score, hypertension, and baseline cognitive score) can help clinicians and public health practitioners find appropriate prevention and treatments for sleep disturbances.

The following statements and Table 5 are summary of the most important findings of our analysis.

1. There are heterogeneous effects of sleep disorders on various cognitive function domain scores in older adults.

2. Having *difficulty falling asleep* caused faster cognitive declines in all domains except attention. The rate of decline was associated with baseline demographics, health conditions, and cognitive domain scores of the subjects. Our analyses showed that the causal relationship of difficulty falling asleep and faster cognitive decline exists in several group of subjects with such associations: faster executive function decline was associated with those who had higher diastolic BP ; faster language decay was associated with those who had higher baseline visuospatial domain scores (≥-0.47) and reported fair or bad general health condition; rapid memory loss was associated with those who had lower baseline executive function domain scores (<0.69) and no subjective memory complains, OR associated with those who had lower baseline executive function domain scores (<0.69) and subjective memory complains but had higher baseline attention domain scores (≥0.043); faster visuospatial decline was associated with those who either had lower baseline memory domain scores (<0.25) and lower executive function domain scores (<0.14) or had higher baseline memory domain scores (≥0.25) but were younger (65-74 years old).
3. Having *difficulty staying asleep* caused faster cognitive decline in all domains except language. Our analyses showed that the causal relationship exists in several group of subjects with such associations: rapid attention decline was associated with those who had lower baseline memory domain scores (<0.0022), and lower systolic BP (<129); rapid executive impairment was associated with those who had higher baseline language domain scores (≥ -0.2); faster memory loss was associated with those who had higher baseline executive domain scores; and faster visuospatial degeneration was associated with those who had higher baseline memory domain scores (≥0.68).

4. Having *excessive daytime sleepiness* caused faster cognitive decline in all domains except executive function. This causal relationship exists among those had lower baseline memory domain scores (<-0.17), or those who had higher memory domain score (≥-0.17) but lower diastolic BP (<70), or those who had higher baseline memory (≥-0.17) and high attention (≥ 0.16) scores, higher diastolic BP, and good general health condition (which are associated with faster attention decline); those who had less exercise (which is associated with rapid language decay); those who had lower baseline memory domain scores (<-0.17) (which is associated to faster memory loss); and those who had higher baseline memory domain scores (≥0.67) (which is associated to faster visuospatial decline).
5. Having *excessive daytime napping* caused cognitive decline in all domains except attention and executive function domains. Our analyses showed that the causal relationship exists in several group of subjects with such associations: faster language decline was associated with those who were 65-74 years old or were greater than 75 years old with baseline systolic BP <131; faster memory loss was associated with those who reported either poor or fair health condition, or those who had reported good health condition and had higher baseline executive function scores (≥0.42), or those who reported good health condition and had lower executive function (<0.42) and lower visuospatial scores (<0.057); faster visuospatial degeneration was associated with women who had higher baseline language domain scores (≥0.23).

In summary, causal tree method provides another approach to explore heterogeneity effects of sleep disturbance on cognitive decline. Sleep disorders caused faster decline in the memory and visuospatial domains in older adults. These causal relationship also showed different effects among people with different sociodemographic or baseline health conditions.

Our causal tree results suggested that older adults with sleep problems can try to improve general health conditions, e.g., having blood pressure controlled, and having regular physical activities, to avoid abnormal cognitive function decline.

However, one must be aware of the fact that unmeasured confounders could cause the estimated causal effect to be biased. In addition, one has to choose an optimal minimum number of observations in each leaf, knowing that; while a smaller number increases variance, a bigger number reduces heterogeneity. The optimal balance between variance and heterogeneity needs to be further investigated. Although we randomly split samples and balanced the group means on nearly all covariates in the model, selection bias and group differences in pre-exposure variables have not been completely removed in our study.

Table 7. Summary of sleep disturbance-cognitive decline causal relationship

Cognitive domain (decline) affected by the sleep disturbance	**Among which subgroups**
Difficulty falling asleep	
Executive function	had higher DBP
Language	had higher baseline visuospatial domain scores (≥-0.47) and reported fair or bad general health condition
Memory	had lower baseline executive function domain scores (<0.69) and no subjective memory complains; OR had lower baseline executive function domain scores (<0.69) and subjective memory complains but had higher baseline attention domain scores (≥0.043)
Visuospatial	either had lower baseline memory domain scores (<0.25) and lower executive function domain scores (<0.14) OR 65-74 years old who had higher baseline memory domain scores (≥0.25)
Difficulty staying asleep	
Attention	had lower baseline memory domain scores (<0.0022) and lower SBP (<129)
Executive function	had higher baseline language domain scores (≥ -0.2)
Memory	had higher baseline executive domain scores
Visuospatial	had higher baseline memory domain scores (≥0.68)
Excessive daytime sleepiness	
Attention	had lower baseline memory domain scores (<-0.17), OR higher memory domain score (≥-0.17) but lower diastolic BP (<70), OR higher baseline memory (≥-0.17), higher attention scores (≥ 0.16), higher diastolic BP (≥70), and reported good general health condition
Language	had less exercise
Memory	had lower baseline memory domain scores (<-0.17)
Visuospatial	had higher baseline memory domain scores (≥0.67)
Excessive daytime napping	
Language	were 65-74 years old or were greater than 75 years old with baseline systolic BP <131
Memory	reported either poor or fair health condition, or those who had reported good health condition and had higher baseline executive function scores (≥0.42), or those who reported good health condition and had lower executive function (<0.42) and lower visuospatial scores (<0.057)
Visuospatial	women who had higher baseline language domain scores (≥0.23)

APPENDIX: STATISTICAL CODES

A.1 STATA EXAMPLE CODES

A.1.1 Stata example codes for descriptive analysis

```
/*for Table 1 to check the distributions of sleep variables*/
foreach var of varlist b_fallasl b_backasl b_wakearl b_sleepint b_sleepday b_dozenap
b_apnea {
	tab `var'
	}


/*for Table 2*/
/*treatment=fallasl(baseline)*/
** categorical variables
Foreach var of varlist agec sex educ health smoke drink stroke tia diabet hibp hrtatt
chf irghrt mcesd1 iadl1 med1 smc1 exercise {
	tab  `var' b_fallasl, column chi2
	}

tab  cardar b_fallasl, column chi2 exact

** continuous variables
foreach var of varlist systole diastole attention executive language memory
visuospatial {
	tabstat `var',s(n mean SD) c(s)format(%14.2fc) by(b_fallasl)
	kwallis `var', by(b_fallasl)
	}
```

A.1.2 Stata example codes for comparing baseline covariates between the training and estimation subsamples

```
/*for Table 4 */
**t-test for continuous variables
foreach var of varlist systole diastole attention executive language memory
visuospatial {
       ttest `var', by(group)
       }

**chi-square test for categorical variables
foreach var of varlist agec sex educ health b_fallasl b_backasl b_wakearl b_sleepint
b_sleepday b_dozenap b_apnea smoke drink stroke tia diabet hibp hrtatt chf irghrt
cardar iadl1 mcesd1 smc1 med1 exercise {
       tab `var' group, column chi2
       }
```

A.2 R EXAMPLE CODES

A.2.1 R example codes for randomly splitting a sample

```
/* for Table 4*/
data1.ind<-runif(nrow(data1))<0.5
train_set<-data1[data1.ind,]
test_set<-data1[!data1.ind,]

write.csv(train_set ,"C:/Users/GUANJOB/Desktop/biost thesis/final
data/train_set_nonmiss.csv",row.names = T)
write.csv(test_set ,"C:/Users/GUANJOB/Desktop/biost thesis/final
data/test_set_nonmiss.csv",row.names = T)
```

A.2.2 R example codes for GBM propensity score weighting

```
/*for Table 3 and Figures 1-2*/
data1= read.csv("C:/Users/GUANJOB/Desktop/biost thesis/final data/all var without
missing value_n= 804.csv", header=T)
library(twang)
set.seed(123)

#1.treatment=fallasl
```

```
## propensity score (PS) model
ps.data1f<-ps(b_fallasl~educ+agec+sex+health+attention+executive+memory+
                language+visuospatial+systol+diastol+smoke+drink+cerebro+
                diabet+hibp+cardiac+exercise+iadl1+mcesd1+smc1+med1,
              data=data1,
              n.trees=10000,
              interaction.depth=3,
              shrinkage=0.01,
              perm.test.iters=0,
              stop.methods=c("es.mean","ks.max"),
              estimand="ATT",
              verbose=FALSE)

## Evaluate the quality of propensity score (PS) weighting
### Checking convergence
plot(ps.data1f)

###Table 3-check balance
data1f.balance<-bal.table(ps.data1f) #return table information on the pretreatment
covariates before and after weighting
data1f.balance

plot(ps.data1f,plots=3) #Figure 1-standardized imbalance plot
plot(ps.data1f,plots=4) #Figure 2-p-value plot for the t-test of group mean difference
for each covariate

###check overlap
plot(ps.data1f,plots=2)
summary(ps.data1f) #note ess(effective sample size)

##extract PS weights from an object
getwt_f<-get.weights(ps.data1f,stop.method = "es.mean")
```

A.2.3 R example codes for Honest Causal Tree

```
/*for Figure 3 */
library(rpart)
library(rpart.plot)
library(causalTree)

##causal trees
###1.causal tree for attention_fallasl
tree_af<-
honest.causalTree(slope_attention~educ+agec+sex+health+attention+executive+memory+
                              language+visuospatial+systol+diastol+smoke+drink+cerebro+
                              diabet+hibp+cardiac+exercise+iadl1+mcesd1+smc1+med1,
                            data=train_set,
                            treatment = train_set$b_fallasl,
                            weights=train_set$getwt_f,
                            est_data = test_set,
                            est_treatment = test_set$b_fallasl,
                            est_weights=test_set$getwt_f,
                            split.Rule = "CT",
                            cv.option = "CT",
                            split.Honest = T,
                            HonestSampleSize=dim(test_set)[1],
                            cv.Honest = T,
```

```
                            split.Bucket = T,
                            split.alpha = 0.5,
                            cv.alpha = 0.5,
                            xval=10,
                            minsize = 30)

###plot cp
plotcp(tree_af)

## get the optimal cp value for pruning
cptab<-tree_af$cptable
print(cptab)
minerr<-min(which(cptab[,4]==min(cptab[,4])))
minplus1se<-cptab[minerr,4]+cptab[minerr,5]
ocp<-cptab[min(which(cptab[,4]<=minplus1se)),1]
print(ocp)

###prune the tree using the best cp
tree_afp<-prune(tree_af,ocp)

###plots
prp(tree_afp,faclen=0,cex=0.70,extra=100,box.palette = 'auto')

###2.for Figure 3.A-causal tree for executive_fallasl
tree_ef<-
honest.causalTree(slope_executive~educ+agec+sex+health+attention+executive+memory+
                              language+visuospatial+systol+diastol+smoke+drink+cerebro+
                              diabet+hibp+cardiac+exercise+iadl1+mcesd1+smc1+med1,
                            data=train_set,
                            treatment = train_set$b_fallasl,
                            weights=train_set$getwt_f,
                            est_data = test_set,
                            est_treatment = test_set$b_fallasl,
                            est_weights=test_set$getwt_f,
                            split.Rule = "CT",
                            cv.option = "CT",
                            split.Honest = T,
                            HonestSampleSize=dim(test_set)[1],
                            cv.Honest = T,
                            split.Bucket = T,
                            split.alpha = 0.5,
                            cv.alpha = 0.5,
                            xval=10,
                            minsize = 30)

###plot cp
plotcp(tree_ef)

## get the optimal cp value for pruning
cptab<-tree_ef$cptable
print(cptab)
minerr<-min(which(cptab[,4]==min(cptab[,4])))
minplus1se<-cptab[minerr,4]+cptab[minerr,5]
ocp<-cptab[min(which(cptab[,4]<=minplus1se)),1]
print(ocp)

###prune the tree using the best cp
tree_efp<-prune(tree_ef,ocp)

###plots
prp(tree_efp,faclen=0,cex=0.70,extra=100,box.palette = 'auto')
```

BIBLIOGRAPHY

1. Gooneratne NS, Vitiello MV. Sleep in older adults: Normative changes, sleep disorders, and treatment options. *Clin Geriatr Med*. 2014;30:591-627
2. Dzierzewski JM, Dautovich N, Ravyts S. Sleep and cognition in older adults. *Sleep Med Clin*. 2018;13:93-106
3. Nebes RD, Buysse DJ, Halligan EM, Houck PR, Monk TH. Self-reported sleep quality predicts poor cognitive performance in healthy older adults. *J Gerontol B Psychol Sci Soc Sci*. 2009;64:180-187
4. Schmutte T, Harris S, Levin R, Zweig R, Katz M, Lipton R. The relation between cognitive functioning and self-reported sleep complaints in nondemented older adults: Results from the bronx aging study. *Behav Sleep Med*. 2007;5:39-56
5. Blackwell T, Yaffe K, Ancoli-Israel S, Redline S, Ensrud KE, Stefanick ML, et al. Association of sleep characteristics and cognition in older community-dwelling men: The mros sleep study. *Sleep*. 2011;34:1347-1356
6. Jelicic M, Bosma H, Ponds RW, Van Boxtel MP, Houx PJ, Jolles J. Subjective sleep problems in later life as predictors of cognitive decline. Report from the maastricht ageing study (maas). *Int J Geriatr Psychiatry*. 2002;17:73-77
7. Potvin O, Lorrain D, Forget H, Dube M, Grenier S, Preville M, et al. Sleep quality and 1-year incident cognitive impairment in community-dwelling older adults. *Sleep*. 2012;35:491-499
8. Elwood PC, Bayer AJ, Fish M, Pickering J, Mitchell C, Gallacher JE. Sleep disturbance and daytime sleepiness predict vascular dementia. *J Epidemiol Community Health*. 2011;65:820-824
9. Sterniczuk R, Theou O, Rusak B, Rockwood K. Sleep disturbance is associated with incident dementia and mortality. *Curr Alzheimer Res*. 2013;10:767-775
10. Blackwell T, Yaffe K, Ancoli-Israel S, Schneider JL, Cauley JA, Hillier TA, et al. Poor sleep is associated with impaired cognitive function in older women: The study of osteoporotic fractures. *J Gerontol A Biol Sci Med Sci*. 2006;61:405-410
11. Jaussent I, Bouyer J, Ancelin ML, Berr C, Foubert-Samier A, Ritchie K, et al. Excessive sleepiness is predictive of cognitive decline in the elderly. *Sleep*. 2012;35:1201-1207
12. Keage HA, Banks S, Yang KL, Morgan K, Brayne C, Matthews FE. What sleep characteristics predict cognitive decline in the elderly? *Sleep Med*. 2012;13:886-892
13. Foley D, Monjan A, Masaki K, Ross W, Havlik R, White L, et al. Daytime sleepiness is associated with 3-year incident dementia and cognitive decline in older japanese-american men. *J Am Geriatr Soc*. 2001;49:1628-1632
14. Chen PL, Lee WJ, Sun WZ, Oyang YJ, Fuh JL. Risk of dementia in patients with insomnia and long-term use of hypnotics: A population-based retrospective cohort study. *PLoS One*. 2012;7:e49113
15. Cricco M, Simonsick EM, Foley DJ. The impact of insomnia on cognitive functioning in older adults. *J Am Geriatr Soc*. 2001;49:1185-1189

16. Foley DJ, Masaki K, White L, Larkin EK, Monjan A, Redline S. Sleep-disordered breathing and cognitive impairment in elderly japanese-american men. *Sleep*. 2003;26:596-599
17. Beebe DW. Cognitive, behavioral, and functional consequences of inadequate sleep in children and adolescents. *Pediatr Clin North Am*. 2011;58:649-665
18. Bourke R, Anderson V, Yang JS, Jackman AR, Killedar A, Nixon GM, et al. Cognitive and academic functions are impaired in children with all severities of sleep-disordered breathing. *Sleep Med*. 2011;12:489-496
19. Ancoli-Israel S, Kripke DF, Klauber MR, Mason WJ, Fell R, Kaplan O. Sleep-disordered breathing in community-dwelling elderly. *Sleep*. 1991;14:486-495
20. Djonlagic I, Guo M, Matteis P, Carusona A, Stickgold R, Malhotra A. Untreated sleep-disordered breathing: Links to aging-related decline in sleep-dependent memory consolidation. *PLoS One*. 2014;9:e85918
21. Cohen-Zion M, Stepnowsky C, Johnson S, Marler M, Dimsdale JE, Ancoli-Israel S. Cognitive changes and sleep disordered breathing in elderly: Differences in race. *J Psychosom Res*. 2004;56:549-553
22. Merlino G, Piani A, Gigli GL, Cancelli I, Rinaldi A, Baroselli A, et al. Daytime sleepiness is associated with dementia and cognitive decline in older italian adults: A population-based study. *Sleep Med*. 2010;11:372-377
23. Tworoger SS, Lee S, Schernhammer ES, Grodstein F. The association of self-reported sleep duration, difficulty sleeping, and snoring with cognitive function in older women. *Alzheimer Dis Assoc Disord*. 2006;20:41-48
24. Saint Martin M, Sforza E, Barthelemy JC, Thomas-Anterion C, Roche F. Does subjective sleep affect cognitive function in healthy elderly subjects? The proof cohort. *Sleep Med*. 2012;13:1146-1152
25. Blackwell T, Yaffe K, Ancoli-Israel S, Redline S, Ensrud KE, Stefanick ML, et al. Associations between sleep architecture and sleep-disordered breathing and cognition in older community-dwelling men: The osteoporotic fractures in men sleep study. *J Am Geriatr Soc*. 2011;59:2217-2225
26. Athey S, Imbens G. Recursive partitioning for heterogeneous causal effects. *Proc Natl Acad Sci U S A*. 2016;113:7353-7360
27. Ganguli M, Snitz B, Vander Bilt J, Chang CC. How much do depressive symptoms affect cognition at the population level? The monongahela-youghiogheny healthy aging team (myhat) study. *Int J Geriatr Psychiatry*. 2009;24:1277-1284
28. 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;17:2265-2281
29. McCaffrey DF, Griffin BA, Almirall D, Slaughter ME, Ramchand R, Burgette LF. A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Statistics in medicine*. 2013;32:3388-3414
30. McCaffrey DF, Ridgeway G, Morral AR. Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychol Methods*. 2004;9:403-425
31. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Jr., Kawas CH, et al. The diagnosis of dementia due to alzheimer's disease: Recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. *Alzheimers Dement*. 2011;7:263-269

32. Spaan PE, Dolan CV. Cognitive decline in normal ageing and early alzheimer's disease: A continuous or discontinuous transition? *Behav Neurol*. 2010;23:203-206