Sleep Duration and Memory in the Elderly Chinese: Longitudinal Analysis of the Guangzhou Biobank Cohort Study

Lin Xu, PhD[1]; Chao Qiang Jiang, MD[2]; Tai Hing Lam, MD[1]; Wei Sen Zhang, PhD[2]; Stacey Shawn Cherny, PhD[4]; G. Neil Thomas, PhD[3]; Kar Keung Cheng, PhD[3]

[1]School of Public Health, University of Hong Kong, Hong Kong; [2]Guangzhou No. 12 Hospital, Guangzhou, China; [3]Public Health, Epidemiology, and Biostatistics, University of Birmingham, Birmingham, UK; [4]Department of Psychiatry, University of Hong Kong, Hong Kong

Study Objectives: Previous cross-sectional studies showed that short or long sleep duration was associated with memory impairment (MI), but longitudinal studies are scarce. We examined whether sleep duration was associated with memory decline or development of MI.

Design, Setting, Participants: We conducted a prospective analysis based on the Guangzhou Biobank Cohort Study on 13,888 participants aged 50+ years without MI at baseline and with a follow-up for a mean of 4.1 years.

Measures and Results: Memory decline was assessed using the Delayed 10-Word Recall Test (DWRT), and in a subset (n = 6,020) with the Mini-Mental State Examination (MMSE). Short and long sleep duration was defined as ≤ 5 hours/day and ≥ 9 hours/day, respectively. Data were analyzed both continuously for memory decline and dichotomously for MI (independently defined as DWRT, < 4; MMSE, < 25). After adjusting for multiple potential confounders, both short and long sleep durations were associated with memory decline using DWRT or MMSE score changes (all P < 0.001). Seven percent (n = 980) developed DWRT-defined MI and 4.0% (n = 194) MMSE-defined MI during the follow-up. Only those with a short (≤ 5 h/day) sleep duration had a significantly increased risk of DWRT-defined MI (odds ratio = 1.53 (95% confidence interval; 1.21-1.93); P < 0.001) relative to normal sleepers (7 h/day). The association remained significant after excluding those with poor self-reported health. No associations were observed with MMSE-defined MI for both long and short sleep durations.

Conclusions: This is the largest study to date addressing the association between extremes of sleep duration and memory decline. The observed adverse relationships provide support for an intervention study to examine the potential benefits of normalizing sleep duration in attenuating memory decline.

Keywords: sleep, sleep duration, memory impairment

Citation: Xu L, Jiang CQ, Lam TH, Zhang WS, Cherny SS, Thomas GN, Cheng KK. Sleep duration and memory in the elderly Chinese: longitudinal analysis of the Guangzhou Biobank Cohort Study. *SLEEP* 2014;37(11):1737-1744.

INTRODUCTION

Aging populations are leading to an increasing burden of dementia and cognitive impairment.[1] A clearer understanding of potentially modifiable determinants to enable the design of effective preventative interventions would be helpful in alleviating this increasing individual and societal burden. Our cross-sectional analyses using baseline data from the Guangzhou Biobank Cohort Study (GBCS) found that both short and long sleep duration were associated with memory impairment (MI).[2] To date only two small prospective studies have investigated the impact of sleep duration on MI.[3,4] One, involving 1,844 US nurses found no significant association between sleep duration and decline of cognitive function after very short follow-up of two years.[3] A recent small nested case-control study from the UK (n = 663) found that short sleep duration (≤ 6.5 h/day) was associated with cognitive impairment after 10 years of follow-up.[4] In this longitudinal study, we examined whether sleep duration was associated with the development of MI in a large community-based sample of older Chinese, taking into account multiple potential confounding factors.

MATERIALS AND METHODS

Detailed information describing the GBCS has been reported elsewhere.[5] Briefly, our participants were randomly selected from an unofficial social and welfare organization that is affiliated with the local government: the "Guangzhou Health and Happiness Association for the Respectable Elders" (GHHARE), whose membership is open to people aged > 50 for a nominal monthly fee of 4 RMB (1 USD = 7 RMB). The GHHARE has a city-wide network with more than 150 branches throughout Guangzhou. It has more than 100,000 older Guangzhou permanent residents. All participants of the GBCS were randomly selected from the membership list of the GHHARE. They were permanent Guangzhou residents, ambulatory, and not receiving treatment for dialysis for renal failure, or radiotherapy/chemotherapy for cancer.[6] Finally 30,518 participants included in the present cohort, and the response rate was 93%. Most of the selected GHHARE members were keen to join the GBCS study because they received a free health examination.[6] The study was approved by the Guangzhou Medical Ethics Committee of the Chinese Medical Association in Guangzhou, China. All participants gave written informed consent before participation.

The baseline recruitment from September 2003 to January 2008 included 30,518 participants. Baseline examination involved face-to-face interview using a computer-based questionnaire containing items on demographic characteristics, lifestyle and dietary factors, and disease history. Clinical and laboratory examinations included anthropometry, memory and neurological examination, and fasting blood glucose and lipids. The follow-up repeat examinations started from March 2008 using the same questions for interview and the same clinical and laboratory examinations as the baseline. As of December 31, 2012, vital status has been determined in 30,115 (98.7%) with 1,458 deaths identified and 18,129 participants (62%, 13,192 women and 4,937 men) returned for the repeated measurement.

Exposures

Sleep-related factors including habitual sleep duration per day, frequency of daytime napping and morning tiredness, and insomnia were assessed. We asked the participants how many hours they normally slept per day during the last month (including naps) and the spontaneous answers were rounded to the nearest integer. Participants were also asked whether they had trouble falling asleep, woke up too early and could not sleep again, or needed to take medicine (including herbal or sleeping pills) at least once a week for sleep initiation or maintenance within the last month, and were classified as having insomnia if they reported having any of these sleep problems lasting ≥ 2 weeks.[2] Frequency of daytime napping was assessed and categorized as never to < 1 day/week, 1-3 days/week, 4-6 days/week, and daily.

Information on socioeconomic position and lifestyle was assessed by a standardized questionnaire.[5] Status of smoking and alcohol drinking was assessed and categorized into "never, former, and current." Detailed information of physical activity was collected by the Chinese version of International Physical Activity Questionnaire (IPAQ).[7] Subjects were asked to recall the amount of time during the past week spent on physical activity including vigorous activity such as heavy lifting, digging, aerobics or fast bicycling, moderate activity such as carrying light loads, bicycling at a regular pace or doubles tennis, and walking. Frequency and average amount of time per day was asked. A metabolic equivalent value (MET) was assigned to each type of activity according to accepted standards,[8] where 1 MET is a resting metabolic rate obtained during quiet sitting. MET values are 3.3 for walking, 4 for moderate activity, and 8 for vigorous activity. Subjects were classified as physically active, moderately active, or inactive according to the IPAQ. Physically active was defined as weekly vigorous activity ≥ 3 days a week achieving ≥ 1500 METs, or moderate activity daily achieving ≥ 3000 METs. Moderately active was defined as weekly vigorous activity ≥ 3 days achieving 480 METs, or any combination of walking, moderate, or vigorous activities ≥ 5 days achieving a least 600 METs. Those who did not meet the criteria for active or moderately active were considered to be physically inactive.

Anthropometric measurements were performed by well-trained nurses in the Guangzhou 12th Hospital using standard protocols. Participants wore light clothing and no shoes. Body weight was measured to the nearest 0.1 kilogram using a balance beam scale (RGZ-120-RT, China). Waist circumference was measured horizontally around the smallest circumference between ribs and iliac crest, or at the navel, if no natural waistline. Hip circumference was measured at the maximum protuberance of the buttocks, and the waist-to-hip ratio was calculated. Body mass index (BMI) was calculated using measured weight and height as weight in kilograms divided by height in meters squared.

For the reproducibility of the questionnaire, 2 examinations on test-retest reliability were conducted during the baseline examination. For the categorical variables, κ values were used, and for the continuous variables, intra-class correlation coefficient (ICC) was used to indicate the test-retest reliability. The first reliability test in 2003 included 200 subjects with a re-interview after one month. The other reliability test was conducted in another 200 subjects in October 2007 with a 2-week interval. The κ/ICC values for the lifestyle questions including smoking, drinking, tea consumption, and physical activity ranged from 0.6 to 0.96 in the 2003 reproducibility test and 0.6 to 0.98 in the 2007 test.[5]

Outcomes

The modified Consortium to Establish a Registry for Alzheimer's Disease Delayed Word Recall Test (DWRT) and the Mini-Mental State Examination (MMSE) were used as outcome variables of cognitive function due to their reliability and accessibility.[9-11] Both tests are simple, time-saving, and need only about 10 minutes to complete. A standardized protocol has been developed for each of the tests. All interviewers were well trained and were required to strictly follow the protocol during the examination to avoid "interviewer bias."[12] The DWRT was used to assess MI in all participants at baseline and follow-up examination. Ten simple words were used for the test, 4 of which were retained from the original English language test:[9] "arm," "letter," "ticket," and "grass." The other 6—"pole," "shore," "cabin," and "engine" were replaced with "corner," "stone," "book," and "stick"—as in the adapted Consortium 10-word list learning task, and "butter" and "queen" were replaced by "soy sauce" and "chairman," as these are more culturally appropriate.[13] During the interview, the 10 words were read out to participants one by one, pausing for 1 sec between each. Participants were asked to recall the words they heard immediately after the last word. This procedure was repeated 3 times; then after 5 min of answering other questions, the participants were asked to recall as many of the words as possible. Participants were given a score of one for each successfully recalled word, with a maximum score of 10, which was used as an outcome variable in this test. Interviewers scored correctly recalled words each time after finishing the procedure, recording the total number of words recalled. The DWRT has been found to be a very efficient instrument for discriminating normal cognition and patients with mild dementia (mainly memory), with an accuracy value of 95.2%.[9] The DWRT was designed specifically to be used in population-based epidemiological studies or in screening examinations.[13,14] MI was diagnosed by a DWRT score < 4 (of 10), corresponding to 1 standard deviation (SD) below the mean (mean ± SD: 5.5 ± 1.8), or physician-diagnosed dementia or Alzheimer disease.[15] Mini-Mental State Examination (MMSE) was added from September 2006 at baseline to assess cognitive function. Thus 6,020 subjects (4,473 women and 1,547 men) who had both the MMSE and DWRT at both baseline and follow-up were included. The MMSE was scored from 0 to 30, with a score < 25 being an appropriate indicator in community-based studies for poor cognitive function.[16]

Of the participants who returned for the follow-up examination, 17,554 (97%) completed the DWRT test and provided information on sleep habits. As the MMSE test was added since September 2006, 6,020 participants (4,473 women and 1,547 men) who had the MMSE test at both baseline and follow-up were also included. Participants were excluded from the data analysis if they reported (1) neurological diseases (n = 2,382) or mental illness including depression, confused speech, schizophrenia, dementia, or Alzheimer disease (n = 171), or (2) extremely short or long sleep duration (< 3 h or > 15 h/day; n = 52)

Table 1—Baseline demographic and clinical characteristics and sleep-related factors by sleep duration per day.

	Sleep duration, hours per day						P value
	3-5	6	7	8	9	≥ 10	
Number of participants	1,661	3,337	4,316	3,419	759	396	–
Age, year	62.5 ± 7.0	61.5 ± 6.9	60.0 ± 6.6	60.3 ± 6.4	60.3 ± 6.3	59.9 ± 6.3	< 0.001
Sex, % men	21.6	25.4	26.7	31.4	33.1	37.1	< 0.001
Education, % primary or below	47.4	41.1	32.1	31.7	33.6	31.8	< 0.001
Longest occupation, % manual	66.3	60.7	57.5	55.8	59.1	58.6	< 0.001
Smoking, % current	6.2	8.2	9.7	10.9	10.5	12.7	< 0.001
Drinking, % current	31.0	31.2	33.7	32.7	30.3	34.9	0.34
Physical activity, % active	54.7	53.4	52.2	53.6	52.7	54.8	0.38
Poor self-rated health, %	22.8	16.3	12.9	12.4	14.0	14.3	< 0.001
Poor health status, %	30.6	26.3	24.4	24.8	27.1	25.0	< 0.001
Waist circumference, cm	78.3 ± 8.9	78.7 ± 9	77.9 ± 8.7	78.5 ± 8.7	78.8 ± 8.8	79.7 ± 9.1	< 0.001
Hip circumference, cm	90.5 ± 6.3	91.1 ± 6.3	90.7 ± 6.2	90.8 ± 6.2	91.1 ± 6.4	91.5 ± 6.1	0.002
Waist-to-hip ratio	0.86 ± 0.07	0.86 ± 0.07	0.86 ± 0.07	0.86 ± 0.06	0.87 ± 0.06	0.87 ± 0.07	< 0.001
Body mass index, kg/m^2	23.7 ± 3.4	23.9 ± 3.2	23.7 ± 3.1	23.8 ± 3.2	23.7 ± 3.3	24.2 ± 3.7	0.04
Daytime napping, % daily	25.5	29.0	35.8	46.8	60.1	64.6	< 0.001
Morning tiredness, % daily	6.1	3.5	2.5	2.4	2.9	4.3	< 0.001
Insomnia, % yes	42.8	17.9	12.1	9.1	10.4	8.3	< 0.001
Baseline delayed word recalled test score	5.9 ± 1.5	6.0 ± 1.5	6.1 ± 1.5	6.1 ± 1.5	6.0 ± 1.4	5.9 ± 1.4	< 0.001
Baseline MMSE score[†]	27.6 ± 2.3	27.9 ± 2.0	28.2 ± 1.8	28.2 ± 1.8	27.9 ± 2.3	27.9 ± 2.0	< 0.001

Data expressed as percentage (%) or mean ± standard deviation. [†]A total of 6,020 participants with both baseline and follow-up MMSE were included here only.

at baseline. Participants with baseline MI (n = 3,817) or without data on sleep-related factors (n = 474) at baseline were also excluded. Thus the final analysis included 13,888 participants (10,060 women and 3,828 men) with all variables of interest.

Statistical Analysis

Pearson χ^2 test was used to compare categorical variables, and one-way analysis of variance (ANOVA) was used for continuous variables when comparing the baseline demographic characteristics by sleep duration per day. Generalized estimating equations were also used for data analysis to take into account the correlated structure of data from repeated measurements at baseline and follow-up.[17] Logistic regression modeling was used to assess the effect of baseline sleep-related factors on the development of MI with crude and adjusted odds ratio (OR) and 95% confidence interval (CI) reported. We have also tested for interactions between sex and sleep-related factors, but no evidence for sex interactions was identified (all P > 0.50). Thus all analyses were conducted by pooling men and women together. Potential confounders included age, sex, occupation, education level, smoking and drinking status, physical activity, waist-to-hip ratio, and sleep-related factors as appropriate. Potential confounders that were significantly associated with sleep duration at baseline in Table 1 were included in the regression models.

Poor health status was defined if participants (1) regularly used medication for chronic diseases, such as diabetes, hypercholesterolemia, or vascular diseases; (2) had any hospital admission during the past 6 months; or (3) had self-reported cardiovascular disease history. Participants with poor health status were more likely to have poorer cognitive function and might

also not sleep well. Hence sensitivity analysis excluding participants with poor health status at baseline was also performed. Another sensitivity analysis was done by including those with self-reported mental or neurological diseases at follow-up, as this group of people could have concurrent MI, which if ignored, could attenuate the associations towards null. Moreover, we also conducted sensitivity analysis after excluding participants with baseline MI defined by different (3 or 4) cutoff points of DWRT. Data analysis was done using STATA/IC 10.1 (Stata Corp LP, College Station, TX, USA).

RESULTS

Among the 13,888 participants with all baseline information of interest, after a mean follow up of 4.1 ± 1.0 years, 980 developed DWRT-defined MI. The incidence was 17.4 per 1,000 person-years in all participants, and was higher in men (21.7 per 1,000 person-years) than in women (15.8 per 1,000 person-years), probably because the men were older than women (mean age ± SD: 63.2 ± 6.4 years for men and 59.9 ± 6.6 years for women). No change in sleep duration was identified over the follow-up period in participants with or without DWRT- or MMSE-defined MI (Table S1, supplemental material).

Table 1 shows that at baseline, 80% of participants had moderate sleep duration of 6-8 h/day, 12% < 6 h, and 8% > 8 hours. Older age, female gender, lower educational level, non-employment, manual occupation, non-current smoking or drinking, and poor self-rated health were associated with shorter sleep duration. Moreover, participants with sleep duration of 7 h per day were associated with lower waist circumference and less morning tiredness and insomnia than those with shorter (< 6 h) or longer (≥ 9 h) sleep duration.

Table 2—Generalized estimating equation analyses of the crude and adjusted differences (mean, 95% CI) of scores of Delayed Word Recall Test (DWRT) and Mini-mental State Examination (MMSE) by baseline sleep duration from baseline to 4.1 years follow-up.

| | Sleep duration, hours per day | | | | | | | P for linear trend (3 to 7 h) | P for linear trend (7 to ≥ 10 h) |
	3-4	5	6	7	8	9	≥ 10		
DWRT									
Crude model	-0.47 *** (-0.58 to -0.36)	-0.35 *** (-0.43 to -0.28)	-0.18 *** (-0.23 to -0.13)	Ref.	0 (-0.05 to 0.05)	-0.15 ** (-0.25 to -0.06)	-0.44 *** (-0.56 to -0.33)	< 0.001	< 0.001
Model 1	-0.20 *** (-0.3 to -0.09)	-0.11 ** (-0.19 to -0.04)	-0.04 (-0.09 to 0.01)	Ref.	0.01 (-0.05 to 0.06)	-0.10 * (-0.19 to -0.01)	-0.33 *** (-0.45 to -0.22)	< 0.001	< 0.001
Model 2	-0.16 ** (-0.27 to -0.05)	-0.11 ** (-0.18 to -0.03)	-0.04 (-0.09 to 0.02)	Ref.	0.01 (-0.04 to 0.07)	-0.08 (-0.17 to 0.01)	-0.31 *** (-0.43 to -0.20)	< 0.001	< 0.001
MMSE†									
Crude model	-0.87 *** (-1.04 to -0.69)	-0.67 *** (-0.79 to -0.55)	-0.39 *** (-0.48 to -0.3)	Ref.	-0.06 (-0.15 to 0.03)	-0.43 *** (-0.58 to -0.27)	-0.66 *** (-0.86 to -0.47)	< 0.001	< 0.001
Model 1	-0.28 ** (-0.46 to -0.10)	-0.26 *** (-0.38 to -0.13)	-0.14 ** (-0.23 to -0.05)	Ref.	-0.04 (-0.13 to 0.05)	-0.36 *** (-0.51 to -0.21)	-0.53 *** (-0.72 to -0.33)	< 0.001	< 0.001
Model 2	-0.28 ** (-0.47 to -0.10)	-0.29 *** (-0.41 to -0.16)	-0.15 ** (-0.24 to -0.06)	Ref.	-0.03 (-0.12 to 0.06)	-0.33 *** (-0.48 to -0.18)	-0.48 *** (-0.67 to -0.28)	< 0.001	< 0.001

†A total of 6,020 participants with both baseline and follow-up MMSE were included here only. Model 1: adjusting for age and sex, occupation, education, smoking, drinking, physical activity, objective health status, and self-rated health. Model 2: additionally adjusting for waist-to-hip ratio, daytime napping, morning tiredness, and insomnia. * P < 0.05, ** P < 0.01, *** P < 0.001.

Table 3—Crude and adjusted odds ratio (OR, 95% CI) for memory impairment (DWRT < 4) by baseline sleep duration.

| | Sleep duration, hours per day | | | | |
	3-5	6	7	8	≥ 9
Overall					
No. of participants	1,661	3,337	4,316	3,419	1,155
% developed MI	10.1	7.5	5.8	6.7	7.3
Crude model	1.83 (1.49 to 2.24) ***	1.31 (1.09 to 1.57) **	Ref.	1.17 (0.97 to 1.41)	1.28 (0.99 to 1.65)
Model 1	1.50 (1.20 to 1.87) ***	1.13 (0.93 to 1.37)	Ref.	1.20 (0.99 to 1.47)	1.26 (0.96 to 1.66)
Model 2	1.53 (1.21 to 1.93) ***	1.12 (0.91 to 1.36)	Ref.	1.20 (0.98 to 1.46)	1.22 (0.92 to 1.61)
Excluding poor health status					
No. of participants	1,153	2,460	3,264	2,571	850
% developed MI	10.2	7.3	5.5	6.9	6.8
Crude model	1.95 (1.53 to 2.49) ***	1.35 (1.09 to 1.67) **	Ref.	1.27 (1.03 to 1.58) *	1.25 (0.92 to 1.70)
Model 1	1.62 (1.25 to 2.11) ***	1.15 (0.91 to 1.45)	Ref.	1.34 (1.07 to 1.68) *	1.25 (0.90 to 1.73)
Model 2	1.62 (1.23 to 2.14) ***	1.14 (0.9 to 1.44)	Ref.	1.35 (1.07 to 1.7) *	1.22 (0.88 to 1.71)

Model 1: adjusting for age and sex, occupation, education, smoking, drinking, physical activity, objective health status and self-rated health. Model 2: additionally adjusting for baseline DWRT score, waist-to-hip ratio, daytime napping, morning tiredness and insomnia. * P < 0.05, ** P < 0.01, *** P < 0.001.

Table 2 describes continuous data and shows that compared to sleep duration of 7 h/day, both those with sleep duration ≤ 5 h and those with sleep duration ≥ 10 h had significant declines in both DWRT and MMSE scores, after adjusting for age, sex, occupation, education, smoking, drinking, physical activity, objective health status and self-rated health, waist-to-hip ratio, daytime napping, morning tiredness, and insomnia. We also tested for quadratic trends between sleep duration and the declines in both DWRT and MMSE scores and found significant results (P values for quadratic trend ranged from 0.02 to < 0.001), suggesting nonlinear associations of sleep duration with both DWRT and MMSE scores. The largest decrease in

MMSE scores (model 2) was 0.5 points over the 4.1 years of follow-up, suggesting a < 2% decline in performance for those with long sleep relative to 7 h sleep, while the largest decrease in DRWT scores was 0.3, equivalent to a 3% decline. The effect size for the association of short sleep duration with both DWRT and MMSE were smaller, with a < 2% and < 1% decline in performance, respectively, relative to 7 h sleep (Table 2).

In Table 3, model 2 shows that compared to sleep duration of 7 h/day, those with sleep duration ≤ 5 h had increased odds of 53% (95% CI 21% to 93%) for DWRT-defined MI. However, no significant association of long sleep duration with DWRT-defined MI was found. Both long and short sleep duration were

Table 4—Crude and adjusted odds ratio (OR, 95% CI) for lower scores of MMSE (< 25) by baseline sleep duration.

	Sleep duration, hours per day			
	3-6	7	8	≥ 9
Overall				
No. of participants	1,683	1,543	1,191	434
% developed MI	4.6	3.0	3.8	5.8
Crude model	1.58 (1.09 to 2.29) *	Ref.	1.28 (0.84 to 1.94)	1.99 (1.21 to 3.28) **
Model 1	1.40 (0.90 to 2.18)	Ref.	1.27 (0.78 to 2.08)	1.75 (0.96 to 3.21)
Model 2	1.32 (0.84 to 2.08)	Ref.	1.28 (0.78 to 2.1)	1.58 (0.85 to 2.92)
Excluding poor health status				
No. of participants	1,211	1,157	859	315
% developed MI	4.6	2.9	4.0	5.4
Crude model	1.65 (1.07 to 2.56) *	Ref.	1.40 (0.86 to 2.29)	1.94 (1.07 to 3.54) *
Model 1	1.38 (0.82 to 2.33)	Ref.	1.55 (0.89 to 2.71)	1.86 (0.91 to 3.80)
Model 2	1.29 (0.76 to 2.21)	Ref.	1.50 (0.86 to 2.64)	1.65 (0.79 to 3.43)

Model 1: adjusting for age and sex, occupation, education, smoking, drinking, physical activity, objective health status, and self-rated health. Model 2: additionally adjusting for baseline MMSE score, waist-to-hip ratio, daytime napping, morning tiredness, and insomnia. * P < 0.05, ** P < 0.01.

Table 5—Generalized estimating equation analyses of the crude and adjusted differences (mean, 95% CI) of scores of Delayed Word Recall Test (DWRT) and Mini-mental State Examination (MMSE) by baseline daytime napping from baseline to 4.1 years follow-up.

	Daytime napping			
	Never to < 1 day/week	1 to 3 days/week	4 to 6 days/week	Daily
DWRT				
Crude model	-0.10 (-0.16 to -0.03) **	Ref.	-0.23 (-0.31 to -0.15) ***	-0.24 (-0.30 to -0.18) ***
Model 1	-0.05 (-0.11 to 0.01)	Ref.	-0.10 (-0.18 to -0.02) **	-0.11 (-0.17 to -0.05) **
Model 2	-0.06 (-0.12 to 0.003)	Ref.	-0.10 (-0.18 to -0.02) **	-0.11 (-0.17 to -0.05) **
MMSE†				
Crude model	-0.11 (-0.21 to 0.002)	Ref.	-0.26 (-0.40 to -0.12) **	-0.26 (-0.37 to -0.16) **
Model 1	0.07 (-0.03 to 0.17)	Ref.	-0.03 (-0.16 to 0.11)	-0.06 (-0.16 to 0.04)
Model 2	0.04 (-0.06 to 0.14)	Ref.	-0.05 (-0.18 to 0.09)	-0.07 (-0.17 to 0.03)

† A total of 6,020 participants with both baseline and follow-up MMSE were included here only. Model 1: adjusting for age and sex, occupation, education, smoking, drinking, physical activity, objective health status, and self-rated health. Model 2: additionally adjusting for waist-to-hip ratio, sleep duration, morning tiredness, and insomnia. ** P < 0.01, *** P < 0.001.

nonsignificantly associated with a higher risk for MMSE-defined MI (Table 4).

Table 5 shows that compared to a daytime napping frequency of 1-3 days per week, those with daily napping had significant declines in DWRT scores, after adjusting for multiple potential confounders. Although those with daily napping appeared to have more decline in MMSE score, the association was not statistically significant. Moreover, compared to participants with a daytime napping frequency of 1-3 days per week, those with daily napping were associated with a higher risk of memory impairment (adjusted OR = 1.34 [1.06 to 1.69]; Table 6)

Adjustment for BMI rather than waist-to-hip ratio in the models did not change the results. (table not shown). Sensitivity analysis (1) excluding participants with poor health status and (2) using different cutoff points of DWRT in defining baseline MI showed similar results (Table S2, supplemental material). No association of insomnia with MI was found (Table S3, supplemental material).

DISCUSSION

To the best of our knowledge, this is the largest study to date addressing the association between sleep duration and memory decline. We found that both short and long sleep duration were associated with memory decline using DWRT and MMSE score changes. However, when we used DWRT-derived MI as an outcome measure, the association was only significant with short sleep duration. Those with sleep duration of 3 to 5 hours per day at baseline were associated with about 50% higher risk of MI than those with sleep duration of 7 hours per day. Although comorbidities could contribute to both shorter sleep duration and poor cognitive performance, the association between baseline sleep duration and MI remained after excluding participants with poor health status and in other sensitivity analyses.

Most existing studies on the association between sleep duration and cognitive function were cross-sectional and showed an association between short sleep duration and

Table 6—Crude and adjusted odds ratio for memory impairment by frequency of baseline daytime napping.

	Daytime napping			
	Never to < 1 day/week	1 to 3 days/week	4 to 6 days/week	Daily
Overall				
No. of participants	5,053	2,056	1,516	5,237
% developed MI	6.7	5.5	8.0	7.8
Crude model	1.24 (1.00 to 1.55)	Ref.	1.51 (1.15 to 1.96) **	1.46 (1.18 to 1.81) **
Model 1	1.14 (0.91 to 1.44)	Ref.	1.29 (0.98 to 1.70)	1.28 (1.02 to 1.61) *
Model 2	1.17 (0.93 to 1.49)	Ref.	1.32 (1.00 to 1.75)	1.34 (1.06 to 1.69) *
Excluded poor health status				
No. of participants	3,909	1,543	1,100	3,731
% developed MI	6.5	5.4	8.5	7.6
Crude model	1.21 (0.94 to 1.56)	Ref.	1.60 (1.18 to 2.18) **	1.43 (1.11 to 1.83) **
Model 1	1.13 (0.86 to 1.48)	Ref.	1.34 (0.97 to 1.84)	1.26 (0.96 to 1.64)
Model 2	1.17 (0.89 to 1.54)	Ref.	1.36 (0.98 to 1.89)	1.32 (1.00 to 1.73) *

Model 1: adjusting for age and sex, occupation, education, smoking, drinking, physical activity, objective health status, and self-rated health. Model 2: additionally adjusting for baseline DWRT score, waist-to-hip ratio, sleep duration, morning tiredness, and insomnia. * $P < 0.05$, ** $P < 0.01$.

poorer cognitive function.[18-26] Results from longitudinal studies are scarce. We found only two prospective studies (the US Nurses' Health Study [NHS] cohort and the UK Cognitive Function and Ageing Study [CFAS]) that tested a similar hypothesis. In the NHS cohort, no significant association of long (≥ 9 h/day, OR = 0.81 [95% CI 0.41, 1.62]) or short (≤ 5 h/day, OR = 1.53 [95% CI 0.77-3.06]) baseline sleep duration with MI was observed over the 2-year follow-up period.[3] As this study focused on a sub-cohort of 1,844 nurses aged 70 years or older and had a short follow-up, it was likely underpowered. The CFAS study followed up 663 participants for 10 years and found that short (≤ 6.5 h/day), but not long sleep duration (≥ 8.5 h/day) was associated with a higher risk of cognitive impairment (OR = 2.02 [1.17-3.48] and 1.27 [0.64-2.48], respectively).[4] Two earlier studies examining the association of changes of sleep duration during follow-up with cognitive performance measured at follow-up only showed that both long and short sleep duration were significantly associated with poor cognitive performance.[22,23] However, the absence of baseline data means these results should be interpreted cautiously, as the association could be explained by participants with baseline cognitive impairment being more likely to develop adverse sleep patterns during the follow-up period. In our study, most participants (96%) had repeated measurements of both sleep-related factors and DWRT. We analyzed the association with newly developed MI with participants having MI (DWRT < 4) at baseline being excluded. Sensitivity analysis excluding those with poor health status suggests potential issue of reverse causation was less likely a reason for the observed associations.

In contrast to the model using DWRT as a continuous outcome measure, the trend toward a higher risk of newly developed dichotomously defined MI in those with longer sleep duration was not significant. One possible explanation may be that using a dichotomized outcome instead of a continuous outcome like scores of DWRT is statistically less powerful.[27] In our study, the number of participants with long sleep duration at baseline was relatively small. Support for this comes from a recent study using actigraph as an objective measurement of sleep that found an association of sleep duration with β-amyloid deposition,[28] a key molecule involved in the pathogenesis of Alzheimer disease, was more pronounced in those with short sleep duration than with long sleep duration, suggesting that short sleep duration might be a more important factor for newly developed MI in older people.

The mechanisms that underlie the association between sleep and memory function remain unclear. Possible explanations include those which are common to vascular dementia. For example, compared to sleep duration of seven hours per day, short sleep duration was associated with a higher risk of obesity[29] and insulin resistance.[30] Insulin, which can cross the blood-brain barrier, promotes the accumulation of β-amyloid plaques that lead to the impairment of cognitive function.[30,31] Finally, it is possible that the aging process or underlying preclinical dementia present at baseline, which was not fully excluded by removing those with existing poor DWRT, leads to both shorter sleep duration and MI. In other words, shorter sleep itself may not result in poorer cognitive function but could be a marker of increased risk. The high incidence of MI in this population emphasizes the importance of this condition in older individuals and the huge potential individual and societal burden that this may impart as populations age.

We also found that daily napping was associated with a higher risk of MI. In agreement with a recent cross-sectional study in Hong Kong Chinese reporting an inverted U-shaped association between cognitive function and the frequency of napping,[18] with those who napped 1-2 times per week having the highest score of cognitive function. This result was also in line with one of our earlier studies using the baseline data of GBCS,[2] in which the adjusted DWRT score in those with daytime napping of 1-3 times/week was slightly higher than the other groups. Interestingly, Lam et al. using the baseline date from GBCS showed that the risk of diabetes due to napping was the lowest in those who napped 1-3 times/week

compared with never or frequent nappers, after excluding subjects with comorbidity or family history of diabetes.[6] However, definite mechanisms of how occasional napping (1 to 3 times per week) could be protective of cognitive function are unclear.

Our study had several limitations. First, assessments of cognitive function and sleep related factors relied on repeated measurements, but not all participants returned for the follow-up examination and interview. Compared to those who participated in the follow-up examination, non-participants tended to be older men with lower socioeconomic position and had unhealthier lifestyle and poorer health status at baseline. They also had a lower DWRT score but similar duration of sleep at baseline. It is possible that some participants who had shorter sleep did not return for the follow-up examination because of cognitive impairment. If this were the case, our estimation of the effect of sleep-related factors on MI could be conservative. Second, information on sleep was self-reported and the reliability of the sleep measures was not assessed, even though the reliabilities of the lifestyle questions were satisfactory.[5] Previous validation studies on the association between self-reported sleep duration and polysomnography, a more objective measurement, found a moderate correlation of 0.6 to 0.7,[32-34] with self-reported data overestimating sleep duration.[35,36] Third, the current study lacks data on sleep apnea. There appears to be some evidence that sleep apnea may be associated with cognitive decline.[37] While adjustment for a measure of obesity, such as waist-to-hip ratio, may partially compensate for the lack of data on sleep apnea, further studies are needed to clarify whether additional control for sleep apnea changes the results. Moreover, participants with Alzheimer disease or other forms of dementia might not be able to report their sleep habits correctly. Even though we excluded participants with neurological or mental disorders at baseline, underdiagnosis could not be completely ruled out. Misclassification of sleep duration in these participants was possible but it would be unlikely that reporting was unidirectional, i.e., only reporting shorter sleep duration. Finally, we used DWRT and MMSE for the assessment of cognitive function, rather than a battery of cognitive tests, which was not feasible in our large population-based study.

In summary, our results have provided the strongest longitudinal evidence to date that both long and short sleep duration at baseline was associated with greater memory decline. It would be interesting to study further if shorter sleep duration leads to MI, or if both of these are features of an underlying aging process. Our results also provide strong evidence for an intervention study to normalize sleep duration, probably by extending sleep duration in shorter sleepers, to clarify the direction of the associations and establish the potential benefits on reducing memory impairment. This would be important to offset the increasing burden of dementia and cognitive impairment being seen in our rapidly aging populations.

ACKNOWLEDGMENTS

DISCLOSURE STATEMENT

REFERENCES

1. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Kokmen E, Tangelos EG. Aging, memory, and mild cognitive impairment. Int Psychogeriatr 1997;9 Suppl 1:65-9.
2. Xu L, Jiang CQ, Lam TH, et al. Short or long sleep duration is associated with memory impairment in older Chinese: the Guangzhou Biobank Cohort Study. Sleep 2011;34:575-80.
3. 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;20:41-8.
4. Keage HA, Banks S, Yang KL, Morgan K, Brayne C, Matthews FE. What sleep characteristics predict cognitive decline in the elderly? Sleep Med 2012;13:886-92.
5. Jiang C, Thomas GN, Lam TH, et al. Cohort profile: The Guangzhou Biobank Cohort Study, a Guangzhou-Hong Kong-Birmingham collaboration. Int J Epidemiol 2006;35:844-52.
6. Lam KB, Jiang CQ, Thomas GN, et al. Napping is associated with increased risk of type 2 diabetes: the Guangzhou Biobank Cohort Study. Sleep 2010;33:402-7.
7. Xu L, Jiang CQ, Lam TH, Zhang WS, Thomas GN, Cheng KK. Dose-response relation between physical activity and cognitive function: Guangzhou biobank cohort study. Ann Epidemiol 2011;21:857-63.
8. The IPAQ group. IPAQ scoring protocol. Available at: https://sites.google.com/site/theipaq/scoring-protocol.
9. Knopman DS, Ryberg S. A verbal memory test with high predictive accuracy for dementia of the Alzheimer type. Arch Neurol 1989;46:141-5.
10. Kucukdeveci AA, Kutlay S, Elhan AH, Tennant A. Preliminary study to evaluate the validity of the mini-mental state examination in a normal population in Turkey. Int J Rehabil Res 2005;28:77-9.
11. Gagnon M, Letenneur L, Dartigues JF, et al. Validity of the Mini-Mental State examination as a screening instrument for cognitive impairment and dementia in French elderly community residents. Neuroepidemiology 1990;9:143-50.
12. Porta MS. A dictionary of epidemiology. 5th ed. Oxford University Press, 2008.
13. Welsh KA, Butters N, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery. Neurology 1994;44:609-14.
14. Prince M, Acosta D, Chiu H, Scazufca M, Varghese M. Dementia diagnosis in developing countries: a cross-cultural validation study. Lancet 2003;361:909-17.
15. Crook T, Bartus RT, Ferris SH, Whitehouse P, Cohen GD, Gershon S. Age-associated memory impairment: proposed diagnostic criteria and measures of clinical change—report of a national institute of mental health work group. Dev Neuropsychol 1986;2:261-76.
16. van Ojen R, Hooijer C, Bezemer D, Jonker C, Lindeboom J, van Tilburg W. Late-life depressive disorder in the community. I. The relationship between MMSE score and depression in subjects with and without psychiatric history. Br J Psychiatry 1995;166:311-5, 319.
17. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986;73:13-22.
18. Auyeung TW, Lee JS, Leung J, et al. Cognitive deficit is associated with phase advance of sleep-wake rhythm, daily napping, and prolonged sleep duration-a cross-sectional study in 2,947 community-dwelling older adults. Age (Dordr) 2013;35:479-86.

19. Kronholm E, Sallinen M, Suutama T, Sulkava R, Era P, Partonen T. Self-reported sleep duration and cognitive functioning in the general population. J Sleep Res 2009;18:436-46.

20. Faubel R, Lopez-Garcia E, Guallar-Castillon P, Graciani A, Banegas JR, Rodriguez-Artalejo F. Usual sleep duration and cognitive function in older adults in Spain. J Sleep Res 2009;18:427-35.

21. Ohayon MM, Vecchierini MF. Normative sleep data, cognitive function and daily living activities in older adults in the community. Sleep 2005;28:981-9.

22. Ferrie JE, Shipley MJ, Akbaraly TN, Marmot MG, Kivimaki M, Singh-Manoux A. Change in sleep duration and cognitive function: findings from the Whitehall II Study. Sleep 2011;34:565-73.

23. Loerbroks A, Debling D, Amelang M, Sturmer T. Nocturnal sleep duration and cognitive impairment in a population-based study of older adults. Int J Geriatr Psychiatry 2010;25:100-9.

24. Ramos AR, Dong C, Elkind MS, et al. Association between sleep duration and the Mini-Mental Score: The Northern Manhattan Study. J Clin Sleep Med 2013;9:669-73.

25. 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;13:1146-52.

26. Blackwell T, Yaffe K, Ancoli-Israel S, et al. Association of sleep characteristics and cognition in older community-dwelling men: the MrOS sleep study. Sleep 2011;34:1347-56.

27. Campbell MJ, Julious SA, Altman DG. Estimating sample sizes for binary, ordered categorical, and continuous outcomes in two group comparisons. BMJ 1995;311:1145-8.

28. Ju YE, McLeland JS, Toedebusch CD, et al. Sleep quality and preclinical Alzheimer disease. JAMA Neurol 2013;70:587-93.

29. Theorell-Haglow J, Berglund L, Janson C, Lindberg E. Sleep duration and central obesity in women–differences between short sleepers and long sleepers. Sleep Med 2012;13:1079-85.

30. Banks WA, Jaspan JB, Huang W, Kastin AJ. Transport of insulin across the blood-brain barrier: saturability at euglycemic doses of insulin. Peptides 1997;18:1423-9.

31. Williamson R, McNeilly A, Sutherland C. Insulin resistance in the brain: an old-age or new-age problem? Biochem Pharmacol 2012;84:737-45.

32. Lockley SW, Skene DJ, Arendt J. Comparison between subjective and actigraphic measurement of sleep and sleep rhythms. J Sleep Res 1999;8:175-83.

33. Armitage R, Trivedi M, Hoffmann R, Rush AJ. Relationship between objective and subjective sleep measures in depressed patients and healthy controls. Depress Anxiety 1997;5:97-102.

34. Patel SR, Ayas NT, Malhotra MR, et al. A prospective study of sleep duration and mortality risk in women. Sleep 2004;27:440-4.

35. Lauderdale DS, Knutson KL, Yan LL, Liu K, Rathouz PJ. Self-reported and measured sleep duration: how similar are they? Epidemiology 2008;19:838-45.

36. Van Den Berg JF, Van Rooij FJ, Vos H, et al. Disagreement between subjective and actigraphic measures of sleep duration in a population-based study of elderly persons. J Sleep Res 2008;17:295-302.

37. Bucks RS, Olaithe M, Eastwood P. Neurocognitive function in obstructive sleep apnoea: a meta-review. Respirology 2013;18:61-70.

Table S1—Sleep duration (hours per day) at baseline and follow-up in those with and without memory impairment (MI)

	Sleep duration, hours per day	
	MI by DWRT (−)	MI by DWRT (+)
Baseline examination	6.97 ± 1.29	6.84 ± 1.39
Follow-up examination	7.03 ± 1.20	6.80 ± 1.34
P value	0.08	0.19
	MI by MMSE (−)	MI by MMSE (+)
Baseline examination	6.98 ± 1.32	7.04 ± 1.40
Follow-up examination	7.04 ± 1.20	6.77 ± 1.42
P value	0.26	0.06

Values presented as mean ± standard deviation.

Table S2—Crude and adjusted odds ratio (OR, 95% CI) for memory impairment using different cutoff points of delayed word recall test (DWRT < 3) by groups of baseline sleep duration.

	Sleep duration, hours per day						P for linear trend (3 to 7 h)	P for linear trend (7 to ≥ 10 h)
	3-5	6	7	8	9	≥ 10		
No. of participants	1,661	3,337	4,316	3,419	759	396	–	–
% memory impairment cases	4.1	3.4	2.5	2.8	3.6	1.8	–	–
Crude model	1.66 ** (1.22 to 2.27)	1.37 * (1.04 to 1.78)	Ref.	1.1 (0.83 to 1.46)	1.44 (0.94 to 2.21)	0.70 (0.32 to 1.52)	0.001	0.73
Model 1	1.36 (0.99 to 1.89)	1.16 (0.87 to 1.53)	Ref.	1.14 (0.85 to 1.53)	1.5 (0.97 to 2.33)	0.73 (0.33 to 1.59)	0.06	0.80
Model 2	1.45 * (1.04 to 2.04)	1.16 (0.87 to 1.55)	Ref.	1.12 (0.83 to 1.51)	1.43 (0.91 to 2.23)	0.68 (0.31 to 1.49)	0.04	0.57

Model 1: adjusting for age and sex, occupation, education, smoking, drinking, physical activity, objective health status, and self-rated health. Model 2: additionally adjusting for baseline DWRT score, waist-to-hip ratio, daytime napping, morning tiredness, and insomnia. * P < 0.05, ** P < 0.01.

Table S3—Crude and adjusted hazards ratio (OR, 95% CI) for memory impairment by baseline insomnia.

	Insomnia		P
	No	**Yes**	
Overall			
No. of participants	11,634	2,254	–
% memory impairment cases	7.0	7.4	
Crude model	Ref.	1.06 (0.90-1.27)	0.99
Model 1	Ref.	1.05 (0.87-1.27)	0.79
Model 2	Ref.	0.96 (0.78-1.18)	0.49
Excluded poor health status			
No. of participants	8,760	1,537	
% memory impairment cases	6.9	7.2	
Crude model	Ref.	1.05 (0.85-1.30)	0.43
Model 1	Ref.	1.08 (0.86-1.36)	0.65
Model 2	Ref.	0.97 (0.76-1.24)	0.96

Model 1: adjusting for age and sex, occupation, education, smoking, drinking, physical activity, objective health status, and self-rated health. Model 2: additionally adjusting for baseline DWRT score, waist-to-hip ratio, sleeping duration, napping, and morning tiredness.