Review Article

Self-reported sleep duration and cognitive performance in older adults: a systematic review and meta-analysis

June C. Lo *, John A. Groeger b, Grand H. Cheng c, Derk-Jan Dijk d, Michael W. L. Chee a

a Centre for Cognitive Neuroscience, Neuroscience and Behavioral Disorders Program, Duke-NUS Graduate Medical School, Singapore

b Department of Psychology, University of Hull, Hull, United Kingdom

c Counselling & Psychology Department, Hong Kong Shue Yan University, Hong Kong

d Surrey Sleep Research Centre, University of Surrey, Guildford, Surrey, United Kingdom

* Corresponding author at: Centre for Cognitive Neuroscience, Duke-NUS Graduate Medical School, 8 College Road, Level 6, Singapore 169857. Tel: +65 6516 4916; fax: +65 6221 8625.

E-mail address: michael.chee@duke-nus.edu.sg (M.W.L. Chee).
Corresponding author at: Centre for Cognitive Neuroscience, Duke-NUS Graduate Medical School, 8 College Road, Level 6, Singapore 169857. Tel.: +65 6516 3963; fax: +65 6221 8625.

E-mail address: june.lo@duke-nus.edu.sg (J.C. Lo).

Highlights

- Meta-analyses of studies relating sleep duration and cognition in older adults
- Self-reported short and long sleep were associated with poorer cognition
- This was shown in both cross-sectional and prospective studies
- Associations of short and long sleep were found across multiple cognitive domains
- Potential contributors to how sleep duration affects cognitive aging are discussed

ABSTRACT
Sleep is important for optimal cognitive functioning across the lifespan. Among older adults (≥ 55 years), self-reported short and long sleep durations have been repeatedly, albeit inconsistently, reported to elevate the risk for poor cognitive function. This meta-analytic review quantitatively summarizes the risk for poorer cognitive function among short and long sleepers in older adults. Eligible publications were searched online and manually. A total of 35 independent samples (N = 97,264) from 11 cross-sectional and seven prospective cohort studies were included. Pooled odds ratios (OR) with 95% confidence intervals (CI) were derived using random-effects models. Self-reported short and long sleep increased the odds for poor cognitive function by 1.40 (CI = 1.27–1.56) and 1.58 times (CI = 1.43–1.74), respectively. Effect sizes varied across studies and may have been moderated by both study type (cross-sectional and prospective) and cognitive domain assessed. For cross-sectional studies, extreme sleep durations were significantly associated with poorer multiple-domain performance, executive functions, verbal memory, and working memory capacity. Prospective cohort studies revealed significant long-term impact of short and long sleep on multiple-domain performance only. These findings establish self-reported extreme sleep duration as a risk factor for cognitive aging.

Keywords:
1. Introduction

Rapid population aging is of rising concern in many developed and developing nations in part because neurodegenerative diseases increase with age and impose heavy socio-economic burdens. While some lifestyle factors, such as physical exercise, have been identified as moderators of age-related brain and cognitive aging [1], the contribution of sleep is less clear. Previous research has revealed adequate sleep to be vital for optimal cognitive function across the lifespan [2–7]. Although the association between sleep and cognitive function is likely to be bi-directional, it has been suggested that alterations in sleep duration might occur prior to the appearance of cognitive symptoms in Alzheimer’s Disease (AD) [8]. In fact, almost half of older adults report at least one
sleep problem [9], and there is growing concern that sleep complaints and disturbances might have negative effects on cognition [10]. While many studies on older adults show evidence for a negative impact of self-reported short [11–25] and long [11–13,16,17,20,22,23,26–30] sleep on cognitive function, others have failed to find relationships between sleep and everyday functioning in this age group [19,31].

In an attempt to resolve these heterogeneous findings, the present study adopted meta-analytic procedures to review and quantify the risk of self-reported short and long sleep for poor cognitive function in older adults. In addition, risk for poor cognitive function was treated separately for cross-sectional and prospective studies in order to estimate the association between extreme sleep durations and current as well as future cognitive functioning. Since studies vary widely in the cognitive tasks used, we grouped tasks into five types, namely those assessing multiple cognitive domains which provide a general measure of cognitive function (such as the Mini-Mental State Examination, MMSE), executive functions, verbal memory, working memory capacity, and speed of processing. This allowed us to examine whether the impact of self-reported sleep duration varied across study types (cross-sectional vs prospective cohort) and cognitive domains. We also evaluated whether age and gender might account for between-study heterogeneity.
2. Methods

2.1. Literature search

This meta-analysis was performed using the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines. We conducted computer and manual searches to locate studies that examined the associations between self-reported sleep duration and cognitive performance of older adults. We searched three electronic databases, PUBMED, MEDLINE, and PsycINFO, for relevant articles published between 1990 and mid-June 2014 using combinations of one of the sleep-related keywords (‘sleep’ and ‘sleep duration’) and one of the cognition-related keywords (‘cognition’, ‘cognitive performance’, ‘cognitive function’, ‘cognitive decline’, ‘mild cognitive impairment’, ‘dementia’, and ‘Alzheimer’s Disease’). We considered only full-length articles that were written in English. A manual search was conducted on seven journals that were likely to publish relevant studies. These journals were “Sleep”, “Journal of Sleep Research”, “Journal of Clinical Sleep Medicine”, “Sleep Medicine”, “Alzheimer Disease and Associated Disorders”, “International Journal of Geriatric Psychiatry”, and “Journal of American Geriatric Psychiatry”. We also examined the reference lists of relevant articles and review papers to identify other studies not found by the electronic and manual searches.
2.2. Inclusion and exclusion criteria

Studies needed to fulfill the following criteria: (1) original article; (2) either a cross-sectional study or a prospective cohort study with sleep duration assessed at baseline and cognitive performance/status assessed at least 1 year later; (3) assessment of sleep duration in hours with self-reported measures; (4) objective assessment of cognitive performance/status; (5) sleep duration as an exposure variable and cognitive performance/status as a criterion variable; (6) report of descriptive or inferential statistics that could be converted to odds ratios (OR) and 95% confidence intervals (CI); (7) participants aged 55 years and above; (8) participants with no sign of dementia at baseline.

We did not include cross-sectional studies that compared sleep duration between cognitively intact persons, individuals with mild cognitive impairment (MCI), AD, and dementia because differences in sleep in cognitively impaired and normal persons could be a consequence of clinical conditions. However, we did
include studies which investigated whether sleep duration predicted incident AD [22] and dementia [11]. We also excluded studies which examined cognitive impairment associated with sleep disorders because of the potential co-morbidity with medical conditions, such as mood disorders [32].

A few studies only partially met the inclusion criteria. For example, the minimum age in some studies was below 55 years. Also, in some studies, sleep duration was indicated by a subscale score of a questionnaire, e.g., the Pittsburgh Sleep Quality Index, or the outcome variable was the change in cognitive performance from baseline to follow-up. For these studies, we asked the authors for data or findings which met the inclusion criteria. In the end, there were five studies among which the minimum age was 39, and they were retained in our data set (Table 1). Note that subsequent analyses revealed no influence of the minimum age of the sample on the effect sizes of self-reported short and long sleep (refer to Sections 3.2.1 and 3.3.1 for more details).

Tworoger et al. [21] and Devore et al. [12] both reported findings from the Nurses’ Healthy Study. We used the former for quantifying the cross-sectional associations between sleep duration and cognitive function; we used the latter for quantifying the impact of sleep duration at baseline on cognitive function in the future. Meta-analysis requires independence of effect sizes. In the analyses where both the cross-sectional and the prospective cohort studies were considered, we included Devore et al. [12] rather than Tworoger et al. [21] for a more balanced sample number of the two types of studies. Note that
prospective cohort studies are less frequent than cross-sectional studies. Similarly, for articles that reported both cross-sectional and longitudinal associations between sleep duration and cognitive function [12,15,19], we only included the effect size representing the contribution of sleep duration at baseline on cognition years later.

Two studies [16,17] used data from the Health 2000 Survey. We included only the study that used more comprehensive cognitive assessment [16]. Keage et al. [15] reported the risk for cognitive impairment for short and long sleepers at the 2- and the 10-year follow-ups. We included the OR at the 10-year follow-up only.

2.3. Data extraction

Data were extracted independently by two trained technicians, and differences were resolved by discussion with J.C.L. Data extracted included the first author’s surname, year of publication, country of origin and name of the cohort studied, study type (cross-sectional or prospective cohort), baseline year, duration of follow-up (0 for cross-sectional studies), sample size, age at baseline (mean and range), gender, reference sleep duration category, shortest sleep
duration category, longest sleep duration category, the cognitive tasks used, the relevant descriptive or inferential statistics for the sleep duration groups, and covariates adjusted in the statistical analysis. For studies that did not report OR, we converted the descriptive or inferential statistics to OR using standard formulas [33]. For example, for studies that used continuous measures for cognitive performance, we first converted the descriptive statistics of the reference, short-, and long-sleep-duration groups into Cohen’s $d$ and then OR [33].

2.4. Definition of ‘short’ and ‘long’ sleep duration

While reference sleep durations ranged from 5 to 9 h across studies, 7 h, 8 h, and 7–8 h were most commonly used (Table 1). For studies that used multiple short-sleep-duration groups, we selected the most extreme group for quantifying the effects of short sleep [34,35], the mode being 5 h or less (Table 1). The same applied to long sleep, with the most common category being 9 h or more (Table 1).

2.5. Cognitive domains
The cognitive tasks used differed across studies (Table 1). To classify these tasks into distinct cognitive domains, we categorized tasks according to two handbooks on cognitive neuroscience—Neuropsychological Assessment [36] and Cognitive Neuroscience: The Biology of the Mind [37]. We assumed diagnoses of AD and dementia to imply the existence of impairment in multiple cognitive domains. To allow reliable estimates of effect sizes, a minimum of four independent samples (k) were required for each cognitive domain [38]. With these considerations, we derived the pooled OR of self-reported short and long sleep for poor performance in five types of cognitive tasks: (1) multiple cognitive domains; (2) executive functions; (3) verbal memory; (4) working memory capacity; and (5) speed of processing (refer to Table S1 for the tasks in each domain). One limitation of this approach was that separating studies into cross-sectional and prospective studies could cause k to drop below four. When this occurred, findings should be best treated as preliminary.

2.6. Study quality assessment

The quality of the studies included in the meta-analyses was assessed using a modified version of the Downs & Black Quality Index score system [39]. The original scale was developed for assessing the quality of both randomized and non-randomized studies. Thus, some of the items were not applicable to the
current meta-analyses (Table S2). The modified scale has five subscales covering reporting, external validity, bias, confounding, and power, and has a maximum score of 16 for cross-sectional studies and 19 for prospective cohort studies. All the studies were of satisfactory quality (score for cross-sectional studies: 13–16; prospective studies: 13–19; Table 1).

2.7. Statistical analyses

Random-effects meta-analyses were performed using Comprehensive Meta-Analysis Version 2.2 (Biostat, Englewood, NJ). The effect size reported here was OR. Since a ‘pure’ cognitive task rarely exists, cognitive tasks categorized into different domains may still recruit similar cognitive processes, e.g., attention, and their measures may be correlated. Thus, for studies which used multiple cognitive measures, we performed intra-study meta-analyses [40,41] to combine all the ORs to form an independent effect size of the overall impact of self-reported short or long sleep on cognitive performance. For studies that used multiple tasks for the same cognitive domain, we meta-analyzed the results to generate an independent effect size for this domain. For those studies that
used multiple statistical models, results from the most restricted model, i.e. the largest number of covariates, were used to provide an estimate of the unique contribution of sleep to cognitive functions.

While the cognitive tasks used varied across studies, a majority used the MMSE or variants of it (Table S1). The MMSE includes tests of executive functions and verbal memory. As tasks may share similar cognitive processes resulting in correlated measures of performance, two main meta-analyses were performed to compute the effect size for the overall impact of self-reported short and long sleep on cognitive function. Between-study heterogeneity in effect sizes was indexed by $Q$ statistic and $I^2$ value. A significant $Q$ value indicates that the effect sizes are heterogeneous. $I^2$ ranges between 0 and 100%, with 25%, 50%, and 75%, respectively, indicating low, moderate, and high levels of heterogeneity [42]. We assessed the possibility of publication bias using three techniques. We generated funnel plots for the effects of self-reported short and long sleep separately. In a funnel plot, the effect sizes (log OR in our case) of all studies are plotted against their standard errors. If there is no publication bias, the plot will be funnel shaped, since studies with larger samples will show less variability in effect sizes, while those with smaller samples will have more variable effect sizes [43,44]. We used Egger’s test of intercept [45] and the ‘trim and fill’ method [43,46] to, respectively, quantify and identify funnel plot asymmetry.
Four moderator analyses were performed to determine whether specific features of a study or the cognitive domain assessed affected the effect sizes of short and long sleep. Firstly, we performed meta-regression using the random-effects model (methods of moments) to determine whether minimum age of the samples moderated the associations between sleep duration and cognitive function. Mean age was not used since this information was missing for five of the 35 samples. Secondly, we used meta-regression to examine the effects of gender by using the percentage of male participants in each sample. In these meta-regressions, the estimated slope indicates the change in log OR for every one-point increase in minimum age or percentage of male participants.

Thirdly, we derived effect sizes of self-reported short and long sleep separately for the cross-sectional and the prospective cohort studies. This allowed us to quantify the cross-sectional associations of short and long sleep with cognitive function, as well as the long-term impact of extreme sleep durations. Finally, to investigate whether certain cognitive domains were more sensitive to the effects of extreme sleep durations, we quantified the effect sizes of short and long sleep on performance in tasks assessing multiple cognitive domains, executive functions, verbal memory, working memory capacity, and speed of processing.

This was done for all the cross-sectional and the prospective cohort studies both separately and combined.
3. Results

3.1. Characteristics of included studies

Eighteen studies met all the selection criteria (Fig. 1). Of these, 11 were cross-sectional, and seven used a prospective cohort design (Table 1). Some studies reported statistics separately for males and females, and one reported data from six countries, resulting in a total of 35 independent samples (26 cross-sectional and nine prospective). Overall, the meta-analysis included 97,624 individuals from 14 countries (five studies from the USA, three from China, two each from the UK, Spain, and Finland, and one each from France, Ghana, India, Mexico, Russia, South Africa, Germany, Singapore, and Canada). For the prospective cohort studies, mean follow-up duration ranged from 1 to 22 years.

3.2. Self-reported short sleep and cognitive function

3.2.1. Overall short sleep effect
ORs of self-reported short sleep for poor overall cognitive function from 35 samples (18 studies; \( N = 97,624 \)) are illustrated in Fig. 2(a). The odds for poor cognitive function were 1.40-times higher among short sleepers than normal sleepers (OR = 1.40, 95% CI = 1.27–1.56; Table 2). There was no evidence for publication bias (Fig. 3(a); Egger’s test: \( p = 0.71 \)). The ‘trim and fill’ method suggested no missing study. There was significant heterogeneity in effect sizes across studies \( (Q = 55.82, p = 0.01, I^2 = 39.09\%) \). Moderator analyses showed that neither age (slope = 0.00; \( p = 0.91 \)) nor gender (slope = 0.00; \( p = 0.49 \)) affected the overall effect size of short sleep.

3.2.2. Overall short sleep effect: by study type

Study type alone did not account for the between-study heterogeneity. In the subgroup analyses where we quantified the effect size of self-reported short sleep on performance for cross-sectional and prospective cohort studies separately (Table 2), cross-sectional studies showed 1.42-times higher odds for poor cognitive function for short relative to normal sleepers \( (OR = 1.42, 95\% CI = 1.28–1.59) \). Prospective cohort studies also revealed that short sleep posed
1.45-times higher odds for poor cognitive function later in life (OR = 1.45, 95% CI = 1.13–1.87). Thus, the cross-sectional associations and the long-term impact of self-reported short sleep on cognitive performance were similar.

Among the prospective cohort studies, the follow-up duration ranged from 1 to 22 years, allowing us to explore whether this factor affected the effect size of short sleep duration. The OR from studies with shorter follow-up periods (< 6 years; k = 6) was 1.60 (95% CI = 1.14–2.24), while the OR from studies with longer follow-up period (≥ 6 years; k = 3) was 1.38 (95% CI = 0.91–2.07). The latter association just missed statistical significance; however, this finding should be treated as preliminary given that the determination of effect size was based on only three independent samples.

3.2.3. Short sleep effect: by cognitive domain

The effect sizes of self-reported short sleep appeared to vary across cognitive domains (Table 2). Short sleep was significantly associated with poorer performance in multiple-domain tasks (OR = 1.28, 95% CI = 1.07–1.53). Since we assumed that impairment in multiple cognitive domains was associated with AD and dementia, to determine whether these clinical conditions might have inflated the risk for poor performance in this domain, we also calculated a
pooled OR excluding the two relevant studies [11,22]. This OR was of a similar magnitude and remained statistically significant (OR = 1.24, 95% CI = 1.02–1.50).

Other than performance in multiple-domain tasks, we also found significant associations between short sleep duration and poor executive function, verbal memory, and working memory capacity, with ORs ranging from 1.33 to 1.35 (Table 2). However, no significant association was found for speed of processing. This is most likely due to the smaller number of studies (k = 6) investigating the contribution of sleep duration to this cognitive domain as the OR for speed of processing was of similar magnitude to the other domains.

3.2.4 Short sleep effect: by cognitive domain and study type
The associations between self-reported short sleep and performance on various cognitive domains were still observed when we considered cross-sectional studies only (Table 2). Self-reported short sleep was associated with higher odds for deficits in multiple-domain tasks, executive functions, verbal memory, and working memory capacity, with ORs ranging from 1.30 to 1.38. A significant association was not present for speed of processing.

In prospective cohort studies, self-reported short sleepers showed elevated risk for poor performance only in multiple-domain tasks (OR = 1.44, 95% CI = 1.02–2.04). No significant association was found in other domains (Table 2), and this was likely due to the small number of prospective cohort studies which assessed performance in these cognitive domains (k = 2–4).

3.3. Self-reported long sleep and cognitive function

3.3.1. Overall long sleep effect

ORs of self-reported long sleep for poor overall cognitive function from 33 samples (17 studies; N = 97,558) are shown in Fig. 2(b). Long sleep was associated with 1.58-times higher odds for poor cognitive function (OR = 1.58, 95% CI = 1.43–1.74; Table 2). Although the ‘trim and fill’ method identified that
one study on the right side of the funnel plot should be placed on the left side in order to achieve symmetry (adjusted OR = 1.50, 95% CI = 1.41–1.60), the Egger’s test was not significant (Fig. 3(b); Egger’s test: \( p = 0.18 \)). Therefore, we concluded that there was no strong evidence for publication bias. Effect sizes across studies were heterogeneous (\( Q = 61.62, p = 0.001, I^2 = 48.07\% \)). However, neither age (slope = −0.01; \( p = 0.26 \)) nor gender (slope = 0.00; \( p = 0.41 \)) appeared to moderate the relationship between self-reported long sleep and poor cognitive performance.

3.3.2. Overall long sleep effect: by study type

Study type alone could not explain the heterogeneity in the effect sizes of self-reported long sleep across studies. The odds for poor cognitive function among long sleepers were both significant from the cross-sectional (OR = 1.61, 95% CI = 1.48–1.76) and the prospective cohort studies (OR = 1.47, 95% CI = 1.08–2.00). This suggested similarity in the cross-sectional associations and the long-term impact of self-reported long sleep on cognitive performance.
To explore whether the follow-up durations in prospective cohort studies influenced the effect size of long sleep duration, we derived pooled ORs separately for studies with shorter (< 6 years: \( k = 4 \)) and longer (\( \geq 6 \) years: \( k = 3 \)) follow-up. Long sleep duration appeared to have stronger impact on cognitive performance in studies with shorter follow-up (\( OR = 1.81, 95\% CI = 1.14–2.86 \) vs. \( OR = 1.25, 95\% CI = 0.93–1.67 \)). However, it is worth noting that only three independent samples contributed to the pooled ORs for longer follow-up; thus, this association and the apparent difference in findings due to different follow-up durations should be interpreted with caution.

3.3.3 Long sleep effect: by cognitive domain

Cognitive domain seemed to contribute to the heterogeneity in the effect sizes of long sleep on cognitive performance. Subgroup analyses revealed that long sleep increased the odds for poor multiple-domain performance by 1.42-times (95\% CI = 1.13–1.78), and this OR was of a similar magnitude even after studies linking long sleep with incident AD and dementia were removed (\( OR = 1.35, 95\% CI = 1.06–1.73 \)).
We also found significant associations between long sleep and poor executive function, verbal memory, and working memory capacity, with ORs ranging between 1.34 and 1.47 (Table 2). In contrast, long sleep did not significantly affect the odds for slower processing speed (Table 2), although the OR appeared to be greatest among all the domains studied (OR = 1.69, 95% CI = 0.90–3.17). This was likely a result of the small number of independent samples (k = 4) available for analysis.

3.3.4 Long sleep effect: by cognitive domain and study type

The effects of self-reported long sleep also seemed to vary across cognitive domains when we performed subgroup analyses separately for the two types of studies. For cross-sectional studies, self-reported long sleep was associated with poorer executive function, verbal memory, and working memory capacity, with ORs ranging from 1.34 to 1.49 (Table 2). Long sleep did not affect the odds for slower processing speed (Table 2). The association between long sleep and poor multiple-domain performance in cross-sectional studies was marginally non-significant (OR = 1.36, 95% CI = 0.98-1.89), but was similar in magnitude to the significant OR reported in prospective cohort studies (OR = 1.43, 95% CI = 1.04–1.97). Note that no prospective cohort study has reported
the impact of self-reported long sleep on executive functions, working memory capacity, or speed of processing, and only two samples investigated the long-term impact of long sleep on verbal memory.

4. Discussion

The present study quantitatively reviews the last decade of data on the association between self-reported short and long sleep on cognitive performance among older adults. Overall, self-reported short and long sleep elevated the odds for poor cognitive function by 1.40-times and 1.58-times, respectively. These effect sizes are considered in the small range by convention [47]. Although it has been previously shown that age and gender might affect sleep duration [48–50], the findings from this meta-analysis revealed minimal influence of these demographic factors on the association between sleep duration and cognitive function. Extreme sleep durations had broad impact across multiple cognitive domains including executive function, verbal memory, and working memory capacity. Although speed of processing appeared spared, this result is likely a result of the paucity of studies specifically addressing this domain. Most of these significant associations remained significant when only cross-sectional studies were considered. The long-term impact of self-reported short
and long sleep on overall cognitive performance was significant, but was limited to multiple-domain performance probably due to the small number of prospective cohort studies investigating the long-term contribution of extreme sleep durations to specific cognitive domains.

4.1 Potential explanations for the association between sleep duration and cognitive function

The biological mechanisms underlying the association between extreme sleep durations and poorer cognitive function in older adults remain unclear. However, progress has been made in uncovering putative pathways.

4.1.1 Effects of short sleep on the brain

Short sleep has negative effects on brain morphometry, activation, and physiology. In community-dwelling older adults, short sleep has been associated with a higher rate of ventricular expansion and faster decline in global cognitive score [18]. Poor sleep quality has been associated with higher rates of cortical
atrophy in frontal, temporal, and parietal lobes [51]. Relative to short sleepers, older adults who slept for 8 h evidenced greater parahippocampal and inferior frontal activation during a verbal encoding task for optimal performance [52].

Recently, sleep has also been found to be important for the clearance of beta-amyloid from the brain [53]. This clearance can be attenuated following acute total sleep deprivation [53,54]. These experimental findings were buttressed by an epidemiological study that reported a cross-sectional association between self-reported short sleep duration and greater amyloid burden in community-dwelling older adults [55]. Amyloid deposition in brain regions that control sleep could in principle aggravate disruption of sleep, setting up a vicious cycle [56]. Indeed, using a Drosophila model of AD, a recent study demonstrated that accumulation of beta-amyloid results in reduced and fragmented sleep, and sleep deprivation for 1 week increases beta-amyloid burden, suggesting a bi-directional relationship between sleep and cognition [56]. However, the long-term effects of chronic short sleep on amyloid accumulation in humans remain to be clarified.

4.1.2 Effects of short and long sleep via systemic inflammation
Several cohort studies on older adults have linked chronic, low grade inflammation evidenced by elevated levels of inflammatory markers, such as interleukin-6 (IL-6) and C-reactive protein (CRP), with negative effects on brain structure [57], as well as increased risk for vascular dementia and AD [58]. However, one large study did not find an association between inflammatory markers and cognitive decline [59]. Elevated levels of IL-6 and CRP have been found in sleep-deprived individuals [60,61], as well as long sleepers [62], suggesting a possible link between short and long sleep, increased inflammation, and impaired cognition. However, a recent longitudinal study that recruited relatively healthy older adults did not find any significant associations between high-sensitivity CRP, and sleep duration, brain volumes or performance in five cognitive domains [18].

4.1.3. Medical comorbidities associated with extreme sleep durations

Cardiometabolic disease and depression associated with extreme sleep durations might have adverse effects on cognitive function. Both short and long sleep have been found to adversely affect cardiometabolic health [34,63], which in turn have negative effects on brain aging and cognition [1,64]. Insomnia or hypersomnia are prominent symptoms in major depressive disorder [65] which can by itself give rise to cognitive impairment [66]. In addition, long-standing
depression elevates the risk for later dementia via multiple mechanisms including hippocampal atrophy, amyloid deposition, increased inflammation, and elevated risk of vascular disease [67].

4.1.4. Overlapping neural correlates for decreasing sleep and performance with age

With increasing age, sleep becomes fragmented and this has been attributed to the interaction between attenuated homeostatic sleep drive and weaker circadian sleep-promoting signals in the early morning [68]. Circadian wake-promoting signals in the afternoon and evening also become weaker, leading older adults to fall asleep more readily during the wake maintenance zone, and to report higher levels of sleepiness in the late afternoon and evening than their younger counterparts [69]. Such fragmentation in rest–activity rhythm has been associated with slower processing speed, poorer memory and poorer executive function [70]. Furthermore, sleep fragmentation elevates the risk of incident AD and the rate of cognitive decline in older adults [71].

The locus coeruleus is important for the regulation of arousal [72,73] and there is evidence for its contribution to cognitive performance [74,75], the latter possibly related to the neuronal projection from the locus coeruleus to the prefrontal cortex [PFC; 73]. A recent animal study showed that sleep deprivation
can result in neuronal loss in the locus coeruleus [76], which may lead to negative cognitive consequences. This is a candidate mechanism explaining how self-reported short sleep might lead to poorer cognitive function in older adults.

The high prevalence of sleep problems among individuals with MCI [77], AD, and other dementias [78], sundowning (i.e., increased confusion and restlessness in the evening) among patients with AD [79], and the high conversion rates of patients with REM sleep behaviour disorder to various forms of dementia [80,81] all suggest that common brain areas may underlie disrupted sleep–wake regulation and neurodegenerative diseases [10,82].

4.2. Broad effects of sleep duration on cognitive performance

We showed here that self-reported extreme sleep durations had significant associations with poorer multiple-domain performance, executive functions, verbal memory, and working memory capacity. The significant association with multiple-domain performance was likely due to the effects of sleep duration on the sub-domains assessed in these tasks. Most of the studies included in this meta-analysis used the MMSE or its variant to measure multiple-domain
performance (Table S1). The broad-based effect on multiple cognitive functions is consistent with the multi-regional [51] or non-localized [18] associations between poor sleep quality or shorter sleep duration and poorer cognitive performance in older adults.

Although the hippocampus, which supports declarative memory is vulnerable to multiple insults [1], our meta-analysis on sleep duration and cognition in older adults did not uncover a specific predilection to memory impairment. The small number of studies examining speed of processing may not have provided sufficient information as to whether self-reported short and long sleep contributes to slower processing speed, a cognitive domain that shows strong age-related decline [83].

It has been argued that because sleep loss invokes functional deficits in working memory and executive functions that are supported by the PFC [84], this part of the brain might be particularly sensitive to sleep loss [85]. While the results of the present meta-analysis support contribution of extreme sleep durations to degradation of these cognitive domains, the effect is not disproportionate to verbal memory.

4.3. Limitations and future studies
Our meta-analysis has several limitations. Firstly, sleep duration was based on self-report. Most epidemiological studies evaluated sleep with subjective measures possibly because of the low cost and ease of administration. However, it is still not clear what leads someone to declare not sleeping 7–8 h when prompted with a question as simple as “how many hours of sleep do you get each night?” [86]. To further complicate the situation, response to this simple question may not even reflect the actual amount of sleep as discrepancies between subjective and objective measures of sleep exist [87]. Reasons for self-reported short sleep duration include but are not limited to lower sleep need, busy work schedule, sleep disorders, and medical conditions. Self-reported long sleep duration may indicate health problems. For example, patients with major depressive disorder tend to either over-estimate or underestimate the amount of sleep [88]. Despite these limitations of self-reported sleep duration, the present data may provide some insights into the interactions between sleep and cognition in older adults.

Secondly, we focused on the association between self-reported sleep duration and cognitive performance, although a number of studies have also found poor self-reported sleep quality and sleep complaints to be associated with poor cognitive function in older adults [14,20,22,23,25,26,30,31,89–97]. Only a handful of studies have examined the contribution of sleep macrostructure to older adults’ cognitive function. More stage 1 sleep [98] and less rapid eye movement
(REM) sleep [98,99] are associated with poorer cognitive function. In addition, we and others have shown that less slow wave sleep (SWS) is associated with impaired sleep-dependent memory consolidation [100] and more frequent false memory recollection [101]. The contribution of sleep microstructure to older adults’ cognitive function is also not well characterized. Some recent evidence has suggested an association of lower spindle and slow wave density with poorer performance in executive functions, attention, and verbal memory [99].

Thirdly, studies varied widely in the question probing sleep duration, making it impossible to unify the operational definition of sleep duration. While most studies assessed average or habitual sleep duration without specifying the time window this was observed for [11–13,15,16,19,21–23,26,27,30], others specified the previous night [28], two nights [14], 4 weeks [29], or 1 month [18,20,31]. Moreover, some studies measured nocturnal sleep duration only [13–15,18–20,26,28–31], while others assessed sleep duration per day or in a 24-h period and thus, included daytime naps [11,12,17, 21–23,27]. Findings regarding the effects of daytime napping on cognitive performance in older adults are heterogeneous. Cross-sectional studies have shown that both frequent and long naps are associated with higher risks for poor cognitive function [23,102]. In contrast, in a prospective cohort study where cognitive performance was tracked over time, older adults who napped at baseline had lower risk for cognitive decline at the 2- and the 10-year follow-up visits [15]. The equivocal
nature of these results remains to be explained, but may be partly due to differences in the reasons for taking a nap across individuals. While for older adults with disrupted nocturnal sleep, naps are needed to satisfy their sleep needs, others may doze off during the day because of health issues or lack of stimulation.

Other than the differences in how sleep duration was assessed, studies also varied considerably in their definition of old age, the confounding factors considered, the duration of follow-up (for prospective cohort studies), and the sleep duration categories used. For example, while most studies involving older adults screened out participants with signs of dementia, those including younger participants might not, and could have included participants with incipient cognitive impairment. However, findings from our meta-regression analyses did not reveal any significant contribution of the minimum age of the samples to the differences in effect sizes across studies. It is also worth noting that where possible, we used the statistical models that included the largest number of covariates for a better estimate of the unique contribution of self-reported sleep duration to cognitive function. This analytic strategy cannot benefit studies that did not record health variables. We also explored the possible influence of follow-up durations on the sleep–cognition associations in prospective cohort studies. It appears that extreme sleep durations do not contribute to cognitive function beyond 6 years of sleep measurement; however, these findings were
from three studies only. To take into account of the diversity across studies in the sleep duration categories used, we selected the most extreme category for the quantification of effect size. While this might result in inflated effect sizes, setting arbitrary cutoffs for short and long sleep duration, e.g., < 6 or > 9 h, would exclude studies that did not use such a categorization scheme, and might underestimate the impact of extreme sleep durations on cognition.

Finally, the number of independent samples from cross-sectional studies contributing to our meta-analysis was almost three-times more than that from prospective cohort studies (26 vs 9). Most of the prospective studies only used a general measure of cognition, e.g., MMSE [15,20], or diagnosis of AD [22] or dementia [11]. In only three prospective studies was performance in specific cognitive domains assessed [12,18,19], limiting the possibility of quantifying the long-term impact of extreme sleep durations on specific domains. Future studies should adopt a prospective cohort design and use a more comprehensive cognitive test battery.

5. Conclusion
Both short and long sleep, as assessed by self-report, is associated with poorer cognitive performance in older adults. These findings suggest the possibility that having good sleep hygiene and establishing good sleep habits may reduce cognitive deficits associated with aging.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Acknowledgements

We thank Mr Kep Kee Loh and Miss Ruth Li-Fang Leong for their contribution to literature search and data coding. Financial support was provided by National Medical Research Council, Singapore (NMRC/STaR/0004/2008) and Biomedical Research Council, Singapore (BMRC 04/1/36/19/372).

References


Fig. 1. Flow chart indicating the results of the systematic literature search with inclusion and exclusion of studies. PSQI, Pittsburgh Sleep Quality Index.
Fig. 2. Forest plots showing all studies relating self-reported (a) short and (b) long sleep to poor cognitive function as compared to a reference group. Results are expressed as odds ratios (OR) and 95% confidence intervals (CI). Higher ORs indicate greater risks for poor cognitive performance relative to the reference group.

Fig. 3. Funnel plots for meta-analyses of self-reported (a) short and (b) long sleep and poor cognitive function.

Table 1. Description of the studies included in the meta-analysis.

<table>
<thead>
<tr>
<th>First author</th>
<th>Publication year</th>
<th>Country</th>
<th>Cohort</th>
<th>Study type</th>
<th>Baseline year</th>
<th>Follow-up (year)</th>
<th>Sample size</th>
<th>Baseline mean age (range)</th>
<th>Gender</th>
<th>Quality score</th>
<th>Sleep duration (h)</th>
<th>Cognitive task</th>
<th>Adjusted variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auyeung</td>
<td>2013</td>
<td>Hong Kong</td>
<td>MrOs and MsOs (Hong Kong)</td>
<td>Cross-sectional</td>
<td>2002–2004</td>
<td>0</td>
<td>2918</td>
<td>73.9 (≥ 65)</td>
<td>Male, female</td>
<td>14</td>
<td>7 to 7.9</td>
<td>&lt; 4, ≥ 10</td>
<td>Mini-Mental State Examination</td>
</tr>
<tr>
<td>Benito</td>
<td>2009</td>
<td>Spain</td>
<td>NEDICES</td>
<td>Prospective</td>
<td>1994–2002</td>
<td>3.2</td>
<td>3286</td>
<td>73.2</td>
<td>Male, female</td>
<td>18</td>
<td>7</td>
<td>≤ 5, ≥ 9</td>
<td>Dementia, Alzheimer’s Disease, Non-</td>
</tr>
</tbody>
</table>
| First author | Publication year | Country | Cohort | Study type | Baseline year | Follow up (year) | Sample size (±SD) | Baseline mean age (range) | Gender | Quality score | Sleep duration (h) | Cognitive task | Adjusted variables  
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Leon</td>
<td>2014</td>
<td>USA</td>
<td>Nurses' Health Study</td>
<td>Prospective</td>
<td>1995</td>
<td>6</td>
<td>13052 (±65)</td>
<td>61.1 (56–66)</td>
<td>Female</td>
<td>16</td>
<td>7 ≤ 5</td>
<td>Alzheimer's Disease type dementia</td>
<td>25, 29</td>
</tr>
<tr>
<td>Devore</td>
<td>2009</td>
<td>Spain</td>
<td>-</td>
<td>Cross-sectional</td>
<td>2001</td>
<td>0</td>
<td>3210 (±60)</td>
<td>71.6 (56–69)</td>
<td>Male, female</td>
<td>16</td>
<td>7 ≤ 5</td>
<td>Telephone Interview for Cognitive Status, Mini-Mental State Examination, Mini-Examen Cognoscitivo</td>
<td>1, 4, 6</td>
</tr>
<tr>
<td>Faubel</td>
<td>2011</td>
<td>UK</td>
<td>Whitehall II Cohort</td>
<td>Cross-sectional</td>
<td>1997–1999</td>
<td>0</td>
<td>5425 (±50)</td>
<td>66.76 (55–70)</td>
<td>Male, female</td>
<td>14</td>
<td>7 ≤ 5</td>
<td>Mini-Mental State Examination, free recall test, Digit Span Forward, Digit Span Backwards Test</td>
<td>1, 2, 4, 6, 8–12, 15, 16, 18, 24</td>
</tr>
<tr>
<td>Ferrie</td>
<td>2014</td>
<td>USA</td>
<td>Study on Global Ageing and Adult Health</td>
<td>Cross-sectional</td>
<td>2007–2010</td>
<td>0</td>
<td>30,200 (±50)</td>
<td>66.76 (55–70)</td>
<td>Male, female</td>
<td>15</td>
<td>6.6 to &lt; 7.5</td>
<td>Immediate Verbal Recall, Delayed Verbal Recall, Digit Span, Digit Span Backwards Test</td>
<td>1, 4</td>
</tr>
<tr>
<td>Gildner</td>
<td>2012</td>
<td>UK</td>
<td>MRC Cognitive Function and Aging Cohort</td>
<td>Prospective</td>
<td>1991–1993</td>
<td>10</td>
<td>1503 (±65–74)</td>
<td>68.76 (55–70)</td>
<td>Male, female</td>
<td>16</td>
<td>&gt; 6.5 to ≤ 8.5</td>
<td>Microvascular dementia, Alzheimer's disease, dementia</td>
<td>3, 6, 7, 11, 12</td>
</tr>
<tr>
<td>Kronholm</td>
<td>2009</td>
<td>Finland</td>
<td>Health 2000 Survey Cohort</td>
<td>Cross-sectional</td>
<td>2000–2001</td>
<td>0</td>
<td>2232 (±55)</td>
<td>71.6 (55–70)</td>
<td>Male, female</td>
<td>16</td>
<td>≥ 8 to &lt; 10</td>
<td>Verbal Fluency Task, Delayed recall test, Mini-Mental State Examination</td>
<td>1, 4, 8–12, 14, 19, 22, 23, 32–36</td>
</tr>
<tr>
<td>Lambiase</td>
<td>2010</td>
<td>USA</td>
<td>Healthy Women Study Cohort</td>
<td>Cross-sectional</td>
<td>2010–2011</td>
<td>0</td>
<td>121 (±68–77)</td>
<td>73.3 (68–77)</td>
<td>Female</td>
<td>13</td>
<td>7 ≤ 5</td>
<td>Modified Mini-Mental State Examination, Delayed Recall, Digit Span Forward, Digit Span Backwards Test</td>
<td>—</td>
</tr>
<tr>
<td>Loebrooks</td>
<td>2010</td>
<td>Germany</td>
<td>HeIDE Cohort</td>
<td>Prospective</td>
<td>1.0</td>
<td>688</td>
<td>72.3 (±70)</td>
<td>72.3 (±70)</td>
<td>Male, female</td>
<td>13</td>
<td>7 ≤ 6</td>
<td>Telephone Interview for Cognitive Status, Mini-Mental State Examination</td>
<td>1, 4, 8, 9</td>
</tr>
<tr>
<td>Potvin</td>
<td>2012</td>
<td>Canada</td>
<td>Survey on Elders' Health Cohort</td>
<td>Prospective</td>
<td>2005–2006</td>
<td>1</td>
<td>1664 (±73.5)</td>
<td>73.5 (65–80)</td>
<td>Male, female</td>
<td>19</td>
<td>&gt; 5 to ≤ 9</td>
<td>Telephone Interview for Cognitive Status, Mini-Mental State Examination</td>
<td>1, 4, 12</td>
</tr>
<tr>
<td>Ramos</td>
<td>2013</td>
<td>USA</td>
<td>Northern Manhattan Study Cohort</td>
<td>Cross-sectional</td>
<td>2006</td>
<td>0</td>
<td>927 (±45)</td>
<td>79 (±45)</td>
<td>Male, female</td>
<td>15</td>
<td>≥ 6 to ≤ 8.9</td>
<td>Telephone Interview for Cognitive Status, Mini-Mental State Examination</td>
<td>1, 4, 9, 13, 14, 19, 20, 26, 27</td>
</tr>
<tr>
<td>Saint</td>
<td>2012</td>
<td>France</td>
<td>Prognostic Cohort</td>
<td>Cross-sectional</td>
<td>2001–2006</td>
<td>0</td>
<td>272 (±72)</td>
<td>74.8 (±72)</td>
<td>Male</td>
<td>15</td>
<td>≥ 7 to ≤ 5</td>
<td>Telephone Interview for Cognitive Status, Mini-Mental State Examination, Free and Follow-up Longitudinal Study</td>
<td>1, 4, 8, 9</td>
</tr>
</tbody>
</table>

Notes: SD = standard deviation, Ref. = reference.
<table>
<thead>
<tr>
<th>First author</th>
<th>Publication year</th>
<th>Country</th>
<th>Cohort</th>
<th>Study type</th>
<th>Baseline year</th>
<th>Follow-up year</th>
<th>Sample size</th>
<th>Baseline mean age (range)</th>
<th>Gender</th>
<th>Quality score</th>
<th>Sleep duration (h)</th>
<th>Cognitive task</th>
<th>Adjusted variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Martin</td>
<td>2003</td>
<td>USA</td>
<td>Bronx Aging Study</td>
<td>Cross-sectional</td>
<td>1980–1983</td>
<td>0</td>
<td>222</td>
<td>79.4 (75–85)</td>
<td>Male, female</td>
<td>15</td>
<td>7 to 8 ≤ 5</td>
<td>Cued Selective Reminding Test, Benton Visual Retention Test, Trial Making Test A, Trial Making Test B, Stroop Test, Categorical Fluency Task, Alphabetical Fluency Task, WAIS-III Similarities Test</td>
<td>1, 4, 8, 9, 11, 14, 21, 23, 30</td>
</tr>
<tr>
<td>Schmutte</td>
<td>2007</td>
<td>USA</td>
<td>Bronx Aging Study</td>
<td>Cross-sectional</td>
<td>1980–1983</td>
<td>0</td>
<td>222</td>
<td>79.4 (75–85)</td>
<td>Male, female</td>
<td>15</td>
<td>7 to 8 ≤ 5</td>
<td>Months backward, Digit Span Test, information, vocabulary, similarities, selective reminding, digit symbol, block design, object assembly, Purdue pegboard</td>
<td>1, 4, 8, 9, 11, 14, 21, 23, 30</td>
</tr>
<tr>
<td>Twoegeer</td>
<td>2006</td>
<td>USA</td>
<td>Nurses’ Health Study</td>
<td>Cross-sectional</td>
<td>2000</td>
<td>0</td>
<td>1844</td>
<td>74.1 (70–81)</td>
<td>Male, female</td>
<td>15</td>
<td>7 ≤ 5</td>
<td>Telephone Interview for Cognitive Status, East Boston Memory Test, Category Fluency Task, Digit Span Backwards Test</td>
<td>1, 2, 4, 12, 28</td>
</tr>
<tr>
<td>Vrta</td>
<td>2013</td>
<td>Finland</td>
<td>Finnish Twin Cohort</td>
<td>Prospective</td>
<td>1975–1981</td>
<td>22.1</td>
<td>1323</td>
<td>52.3 (2–89)</td>
<td>Male, female</td>
<td>18</td>
<td>≥ 7 to &lt; 8</td>
<td>Alzheimer’s Disease</td>
<td>1, 2, 4, 12, 28</td>
</tr>
<tr>
<td>Xu</td>
<td>2011</td>
<td>China</td>
<td>Guangzhou Biobank</td>
<td>Cross-sectional</td>
<td>2003–2008</td>
<td>0</td>
<td>28,670</td>
<td>62.0 (50–85)</td>
<td>Male, female</td>
<td>15</td>
<td>7 ≤ 4</td>
<td>Delayed word recall test</td>
<td>1, 2, 4–6, 8–12, 16–18, 24</td>
</tr>
</tbody>
</table>

*1 = age; 2 = sex; 3 = race; 4 = education; 5 = employment; 6 = occupation; 7 = shift work history; 8 = smoking; 9 = alcohol intake; 10 = caffeine intake; 11 = physical activity; 12 = body mass index; 13 = diabetes; 14 = hypertension; 15 = high cholesterol; 16 = systolic blood pressure; 17 = high-density lipoprotein and low-density lipoprotein; 18 = fasting plasma glucose; 19 = medications; 20 = depression; 21 = hypnotic use; 22 = anxiolytic medication; 23 = mental health status; 24 = self-rated health; 25 = life satisfaction; 26 = medical insurance; 27 = risk for sleep-disordered breathing; 28 = cognition at baseline; 29 = apolipoprotein; 30 = living status; 31 = medical comorbidities; 32 = night-time awakening; 33 = number of social ties; 34 = head of family’s work status; 35 = number of chronic diseases
Table 2 Odds ratios (95% confidence interval) of self-reported short and long sleep for poor cognitive performance

<table>
<thead>
<tr>
<th></th>
<th>Cross-sectional &amp; prospective combined</th>
<th>Cross-sectional</th>
<th>Prospective</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>k OR 95% CI</td>
<td>k OR 95% CI</td>
<td>k OR 95% CI</td>
</tr>
<tr>
<td>Overall*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short sleep</td>
<td>35 1.40 1.27 1.56</td>
<td>27 1.42 1.28 1.59</td>
<td>9 1.45 1.13 1.87</td>
</tr>
<tr>
<td>Long sleep</td>
<td>33 1.58 1.43 1.74</td>
<td>27 1.61 1.48 1.76</td>
<td>7 1.47 1.08 2.00</td>
</tr>
<tr>
<td>Multiple cognitive domains*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short sleep</td>
<td>19 1.28 1.07 1.53</td>
<td>11 1.30 1.05 1.59</td>
<td>9 1.44 1.02 2.04</td>
</tr>
<tr>
<td>Long sleep</td>
<td>17 1.42 1.13 1.78</td>
<td>11 1.36 0.98 1.89</td>
<td>7 1.43 1.04 1.97</td>
</tr>
<tr>
<td>Executive functions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short sleep</td>
<td>22 1.33 1.22 1.46</td>
<td>20 1.33 1.21 1.47</td>
<td>2 1.69 0.69 4.11</td>
</tr>
<tr>
<td>Long sleep</td>
<td>20 1.47 1.29 1.68</td>
<td>20 1.47 1.29 1.68</td>
<td>— — — —</td>
</tr>
<tr>
<td>Verbal memory*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short sleep</td>
<td>23 1.33 1.19 1.49</td>
<td>20 1.38 1.25 1.53</td>
<td>4 1.00 0.86 1.16</td>
</tr>
<tr>
<td>Long sleep</td>
<td>21 1.47 1.31 1.66</td>
<td>20 1.49 1.33 1.67</td>
<td>2 1.28 0.76 2.13</td>
</tr>
<tr>
<td>Working memory capacity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short sleep</td>
<td>15 1.35 1.15 1.59</td>
<td>13 1.38 1.17 1.62</td>
<td>2 0.78 0.32 1.88</td>
</tr>
<tr>
<td>Long sleep</td>
<td>13 1.34 1.22 1.47</td>
<td>13 1.34 1.22 1.47</td>
<td>— — — —</td>
</tr>
<tr>
<td>Speed of processing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short sleep</td>
<td>6 1.36 0.88 2.12</td>
<td>4 1.31 0.79 2.19</td>
<td>2 1.53 0.63 3.72</td>
</tr>
<tr>
<td>Long sleep</td>
<td>4 1.69 0.90 3.17</td>
<td>4 1.69 0.90 3.17</td>
<td>— — — —</td>
</tr>
</tbody>
</table>

k, number of independent samples; OR, odds ratio; CI, confidence intervals.

* For these cognitive domains, Tworoger et al. (2006) and Devore et al. (2014) which studied the same cohort were, respectively, included in the meta-analyses of cross-sectional and prospective cohort studies. For the combined analyses, to ensure independence of effect sizes across studies, only Devore et al. (2014) was included since prospective cohort studies were under-represented. Odds ratios in bold were statistically significant at p < 0.05.