Behavioral/Cognitive

Is Short Sleep Bad for the Brain? Brain Structure and Cognitive Function in Short Sleepers

Anders M. Fjell,1,2 Øystein Sørensen,3 Yunpeng Wang,1 Inge K. Amlien,1 William F. C. Baaré,3 David Bartzés-Faz,4 Carl-Johan Boraxbekk,3,5,6,7,21 Andreas M. Brandmaier,4,14,19 Ilja Demuth,8,22,26 Christian A. Drevon,9,10 Klaus P. Ebmeier,11 Paolo Ghisletta,12,23,24 Rogier Kievit,13 Simone Kühn,14,15 Kathrine Skak Madsen,3,16 Lars Nyberg,1,5 Cristina Solé-Padullés,1 Didac Vidal-Piñeiro,1 Gerd Wagner,17 Leiv Otto Watne,18,20,25 and Kristine B. Walhovd1,2

1Center for Lifespan Changes in Brain and Cognition, University of Oslo, 0373 Oslo, Norway, 2Computational Radiology and Artificial Intelligence, Department of Radiology and Nuclear Medicine, Oslo University Hospital, 0424 Oslo, Norway, 3Danish Research Centre for Magnetic Resonance (DRCMR), Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital-Amager and Hvidovre, 2650 Hvidovre, Copenhagen, Denmark, 4Department of Medicine, Facultad de Medicina i Ciències de la Salut, Universitat de Barcelona, 7Institute of Sports Medicine Copenhagen (ISMC), Copenhagen University Hospital Bispebjerg, 2400 Copenhagen, Denmark, 5Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Department of Endocrinology and Metabolic Diseases (including Division of Lipid Metabolism), Biology of Aging working group, Augustenburger Platz 1, 13353 Berlin, Germany, 6Department of Nutrition, Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, 0372 Oslo, Norway, 7Department of Psychiatry, University of Oxford, Oxford OX3 7JX, United Kingdom, 8Faculty of Psychology and Educational Sciences, University of Gent, 2020 Gent, Belgium, 9Cognitive Neuroscience Department, Donostia Institute for Brain, Cognition and Behavior, Basque Country University, 20018 Donostia-San Sebastian, Spain, 10Department for Functional Brain Imaging, Umeå University, 907 36 Umeå, Sweden, 11Department of Radiation Sciences, Diagnostic Radiology, Umeå University, 907 36 Umeå, Sweden, 12Institute of Sports Medicine Copenhagen (ISMC), Copenhagen University Hospital Bispebjerg, 2400 Copenhagen, Denmark, 13Department of Psychiatry and Psychotherapy, Jena University Hospital, 07743 Jena, Germany, 14Oslo Delirium Research Group, Oslo University Hospital, 0424 Oslo, Norway, 15Department of Psychology, MSB Medical School Berlin, Berlin, Germany, 16Department of Geriatric Medicine, Akershus University Hospital, 1478 Lørenskog, Norway, 17Institute for Clinical Medicine, Faculty of Medical and Health Sciences, University of Copenhagen, 2010 Copenhagen, Denmark, 18Berlin Institute of Health at Charité – Universitätsmedizin Berlin, 10178 Berlin, Germany, 19UniDistance Suisse, 3900 Brig, Switzerland, 20Swiss National Centre of Competence in Research LIVES, University of Geneva, 21Institute of Clinical Medicine, Campus Ahus, University of Oslo, 1478, Lørenskog, Norway, and 26BCRT - Berlin Institute of Health Center for Regenerative Therapies, 13353 Berlin, Germany

Many sleep less than recommended without experiencing daytime sleepiness. According to prevailing views, short sleep increases risk of lower brain health and cognitive function. Chronic mild sleep deprivation could cause undetected sleep debt, negatively affecting cognitive function and brain health. However, it is possible that some have less sleep need and are more resistant to negative effects of sleep loss. We investigated this using a cross-sectional and longitudinal sample of 47,029 participants of both sexes (20–89 years) from the Lifebrain consortium, Human Connectome project (HCP), with measures of self-reported sleep, including 51,295 MRIs of the brain and cognitive tests. A total of 740 participants who reported to sleep <6 h did not experience daytime sleepiness or sleep problems/disturbances interfering with

Received Oct. 21, 2022; revised May 1, 2023; accepted May 8, 2023.


The Lifebrain project is supported by the EU Horizon 2020 Grant Agreement Number 732592 (Lifebrain). In addition, the different substudies are supported by different sources. The Center for Lifespan Changes in Brain and Cognition (CLCBI) is supported by European Research Council Grant Agreements 802634 and 720525 (to A.M.F.) and 313440 (to A.M.F.) as well as the Norwegian Research Council (to A.M.F., K.B.W.) and The National Association for Public Health’s Dementia Research Program, Norway (to A.M.F.). Betula is supported by a scholar grant from the Knut and Alice Wallenberg (KAW) Foundation (J.L.). Barcelona is partially supported by a Spanish Ministry of Economy and Competitiveness (MINECO) grant and ICREA Academia 2019 grants to D.B.-F. (grant number PS2015-66227-R (AEI/ FEDER, UE)) and by the Walnuts and Healthy Aging Study (http://www.clinicaltrials.gov; Grant NCT07616841) funded by the California Walnut Commission, Sacramento, California. BASE II has been supported by the German Federal Ministry of Education and Research Grant Numbers 16SH8327, 16SH8337, 16SH8348, 16SH8348, 01UMW0706, 01G11716A, and 01G11716B and the European Research Council Grant Agreement 677804 (to S.K.). Work on the Whitehall II Imaging Substudy was mainly funded by the Lifelong Health and Well-Being Program Grant G01001354 from the United Kingdom Medical Research Council (“Predicting MIR Abnormalities with Longitudinal Data of the Whitehall II Substudy” to K.P.E.). The Wellcome Centre for Integrative Neuroimaging is supported by core funding from the Wellcome Trust Award 201102/Z/62/Z. Data were provided in part by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; Y444HMNR1657577) funded by the 16 National Institutes of Health (NIH) and Centers that support the NIH Blueprint for Neuroscience Research, and by the McConnell Center for Systems Neuroscience at Washington University. Part of the research was conducted using the United Kingdom Biobank resource under application number 32048.

C.A.D. is a cofounder, stockholder, board member, and consultant in the contract laboratory Vitas AS, performing personalized analyses of blood biomarkers. All other authors declare no competing financial interests.

Correspondence should be addressed to Anders M. Fjell at andersmf@psykologi.uio.no.

https://doi.org/10.1523/JNEUROSCI.2330-22.2023

Copyright © 2023 the authors
falling or staying asleep. These short sleepers showed significantly larger regional brain volumes than both short sleepers with daytime sleepiness and sleep problems (n = 1742) and participants sleeping the recommended 7–8 h (n = 3886). However, both groups of short sleepers showed slightly lower general cognitive function (GCA), 0.16 and 0.19 SDs, respectively. Analyses using accelerometer-estimated sleep duration confirmed the findings, and the associations remained after controlling for body mass index, depression symptoms, income, and education. The results suggest that some people can cope with less sleep without obvious negative associations with brain morphometry and that sleepiness and sleep problems may be more related to brain structural differences than duration. However, the slightly lower performance on tests of general cognitive abilities warrants closer examination in natural settings.

Key words: brain; cognition; hippocampus; MRI; sleep; sleepiness

Significance Statement
Short habitual sleep is prevalent, with unknown consequences for brain health and cognitive performance. Here, we show that daytime sleepiness and sleep problems are more strongly related to regional brain volumes than sleep duration. However, participants sleeping ≤6 h had slightly lower scores on tests of general cognitive function (GCA). This indicates that sleep need is individual and that sleep duration per se is very weakly if at all related brain health, while daytime sleepiness and sleep problems may show somewhat stronger associations. The association between habitual short sleep and lower scores on tests of general cognitive abilities must be further scrutinized in natural settings.

Introduction
Approximately 50% of people sleep less than the recommended 7–9 h ( Hirshkowitz et al., 2015; Watson et al., 2015 ), with 6.5% sleeping less than six ( Koceska et al., 2021b ), many of whom do not report excessive daytime sleepiness. Shorter than recommended sleep is believed to reduce brain health and cognitive performance ( Zamore and Veasey, 2022 ); but guidelines do not take into account two factors. First, partly because of genetic differences ( Dashti et al., 2019 ), sleep need varies. Even vulnerability to sleep loss is a heritable and stable trait, and several genetic polymorphisms are identified ( Casale and Goel, 2021 ). Second, like other physiological systems, sleep need may not be fixed, but is expected to have the ability to adapt to changes in external circumstances ( Horne, 2011 ). This may allow short sleep without sleepiness, reduced brain health and cognitive deficits. Crucially, however, the relationship between sleep duration, daytime sleepiness and neurocognitive function is complex ( Horne, 2010 ). A subset of healthy adults sleeping 7–8 h without daytime sleepiness can still fall asleep within 6 min in the multiple sleep latency test ( MSLT; Roehrs et al., 1996 ). This indicates a level of sleepiness similar to patients with primary sleep disorders, and it was suggested that this could represent accumulated sleep debt from chronically mild insufficient sleep ( Roehrs et al., 1996 ). Lack of daytime sleepiness may result from insensitivity to sleep drive or renormalization in response to chronic sleep deprivation preventing feelings of sleepiness, rather than reflecting lower sleep need ( Mander et al., 2017 ). With extended sleep deprivation, subjective sleepiness can return to baseline levels well before deficits in psychomotor vigilance tasks are normalized ( Van Dongen et al., 2003 ). If such short sleepers are indeed suffering undetected sleep debt despite low levels of daytime sleepiness, neurocognitive deficits may be possible to detect by cognitive tests and brain MRIs. In contrast, if short sleep is due lower sleep need, then such participants are not likely to have poorer cognitive function or brain health.

Duration per se may be a less important indicator of insufficient sleep than daytime sleepiness and problems such as frequently having trouble falling or staying asleep. Short duration ( ≤6 h ) combined with problems and insomnia was associated with higher risk of hypertension, whereas sleeping as short as ≤5 h without problems was not ( Vgontzas et al., 2009 ). Daytime sleepiness was associated with thinner cortex ( Carvalho et al., 2017 ), but sleep duration was not addressed.

Here, we compared cognitive function and brain volumes between short sleepers ( ≤6 h ) with or without sleep problems and daytime sleepiness, and in participants sleeping the recommended 7–8 h. Important aspects of brain health can be measured by structural MRI, which is sensitive to aging ( Walhovd et al., 2016 ) and disease ( Fjell et al., 2014 ). Global brain volume is consistently related to higher general cognitive function ( GCA ; Walhovd et al., 2022 ), and atrophy in specific regions has been associated with reduced cognitive function ( Gorbach et al., 2020 ). Brain morphometry and habitual sleep duration form an inverse U-shaped relationship, peaking at ~7 h ( Spira et al., 2016; Fjell et al., 2020b, 2021 ). We combined data from the Lifebrain consortium ( Walhovd et al., 2018 ), UK Biobank ( UKB; Miller et al., 2016 ), and the Human Connectome Project ( HCP; Van Essen et al., 2013 ) in a mixed cross-sectional and longitudinal design, allowing us to target short ( ≤6 h ) and normal (7–8 h) sleepers with or without daytime sleepiness and sleep problems.

Materials and Methods
Sample
The sample ( see Table 1 ), described in more detail previously ( Fjell et al., 2023 ), consisted of community-dwelling adults from multiple European countries and the United States. All participants gave written informed consent. The Lifebrain project ( Walhovd et al., 2018 ) was approved by the Regional Committees for Medical and Health Research Ethics South East Norway, and substudies approved by the relevant national review boards. For UKB, ethical approval was obtained from the National Health Service National Research Ethics Service ( Ref 11/NW/0382 ). The full sample consisted of 47,029 participants (20–89 years) with information about sleep duration and MRI of the brain and 8694 with general cognitive ability scores, calculated as a g-factor from the available cognitive tests in each sample ( for details, see Walhovd et al., 2022 ). As the specific cognitive test varied across samples ( see below ), the g-factor was used to reduce the impact of test-specific variance on the results. For 3893 participants, longitudinal MRI examinations were available, yielding a total of 51,295 MRIs (mean follow-up interval 2.5 years, range 0.005–11.2, 26,811 female/24,509 male observations).
Lifebrain
Participants from major European brain studies: Berlin Study of Aging II (BASE II; Bertram et al., 2014; Gerstorf et al., 2016), the BETULA project (Nygberg et al., 2020), the Cambridge Center for Ageing and Neuroscience study (Cam-CAN; Shafie et al., 2014), Whitehall-II (WH-II; Filippini et al., 2014), and Center for Lifespan Changes in Brain and Cognition longitudinal studies (LCBC; Walhovd et al., 2016; Fjell et al., 2018).

UKB
UK Biobank is a national and international health resource open to all bona fide health researchers (https://www.ukbiobank.ac.uk; Guggenheim et al., 2015), and includes MRI data from a subsample (https://www.ukbiobank.ac.uk/imaging-data/; Miller et al., 2016). The dataset released February 2020 was used.

HCP
HCP freely share data from 1143 young adults (ages 22–35) from families with twins and nontwin siblings, with 3T MRI and behavioral testing. The dataset used was the 1200 Subjects Release.

The general cognitive ability score was calculated as a g-factor per sample, using available tests. More details are given in (Walhovd et al., 2022), but the following test scores were included: the Practical Problems, Figural Analogies, and Letter Series tests (Düzel et al., 2016; Whitehall-II); Block Design test 7, a measure of immediate free recall of 16 enacted verb-noun sentences, a 30-item five-alternative forced choice vocabulary test, four measures of verbal fluency measured during 1 min, as well as a 26-item general knowledge test (Betula; Nilsson et al., 1997), the standard form of the Cattell Culture Fair, Scale 2 Form A (Cattell and Cattell, 1973), the Spot The Word task (Baddeley et al., 1993; Cam-CAN), Vocabulary and Matrix reasoning from WASI (Wechsler, 1981), raw fluid intelligence score from the UKB Data-field 20016 (UKB), Practical Problems, Figural Analogies, Letter Series (Düzel et al., 2016; BASE-II), Block design and/or vocabulary from WASI-III (Wechsler, 1997) and National Adult Reading Test (NART; Nelson and Willison, 1991; Barcelona), and for HCP, Flanker, Dimensional Change Card Sort, Picture Sequence Memory, List Sorting and Pattern Comparison, and Picture Vocabulary and Reading Tests.

Classification of participants
Sleep information was available for baseline only for most of the participants. For the small number of participants for whom more than one observation about sleep was available, we used the average value across timepoints. For the HCP and the Lifebrain samples except Betula, sleep characteristics were measured by the Pittsburgh Sleep Quality Index (PSQI; Buysse et al., 1989). For Betula, sleep characteristics were measured by The Karolinska Sleep Questionnaire (KSQ; Nordin et al., 2013; Westerlund et al., 2014). For UKB, sleep was measured through multiple questions.

Participants were classified in four groups, according to criteria specified in Table 2. Groups 1 and 2 consisted of short sleepers (≤6 h). Group 1 reported no daytime sleepiness and no sleep problems (“group 1: Short sleep and no sleep problems/sleepiness”), while group 2 reported daytime sleepiness and/or sleep problems (“group 2: Short sleep and sleep problems/sleepiness”). Group 3 and four consisted of participants sleeping the recommended 7–8 h. Like group 1, group 3 participants did not report daytime sleepiness and sleep problems (“group 3: Recommended sleep and no sleep problems/sleepiness”). Like group 2, group 4 participants reported daytime sleepiness and/or sleep problems (“group 4: Recommended sleep and sleep problems/sleepiness”). Since UKB did not include PSQI, the items used for group assignment were not identical for UKB and Lifebrain participants. Participants not satisfying the criteria were ungrouped and not included in the analyses. 19 Lifebrain participants were short sleepers (≤6 h) and experienced daytime sleepiness without reporting sleep problems. This shows that very few participants report to sleep ≤6 h and feel tired during the day unless they also have sleep problems. We considered this group to be too small for statistical comparisons. Of the total sample, 9611 participants were successfully classified into the predefined sleep groups. Of participants sleeping ≤6 h, 740 (6%) and 1742 (14%) were classified as belonging to group 1 or 2, respectively. Of participants sleeping 7–8 h, 3886 (12%) and 3243 (10%) were classified as belonging to group 3 or 4, respectively (see Table 3 for details).

The same sleep categories were used across the age-range, although recommended sleep duration range is not identical for younger and older adults. This was done because mean sleep duration in this sample is relatively stable across the age-range (Fjell et al., 2023), and the recommended lower limit of sleep is the same across adulthood (Paruthi et al., 2016; Hirshkowitz et al., 2015). This led to the exclusion of young and middle-aged participants sleeping 8–9 h, which is within the recommended duration, and hence represent a risk of reducing the representativity of the “recommended sleep” groups. However, 75% of the sample reported to sleep ≤9 h, and so we believe this choice did not bias the results.

Accelerometer-derived sleep duration
We ran validation analyses using sleep duration quantified by accelerometer data for the UKB participants. Raw accelerometer data were

---

Table 1. Sample origins of the full sample

<table>
<thead>
<tr>
<th>Study</th>
<th>Observations</th>
<th>Participants</th>
<th>Age (mean)</th>
<th>Age (min–max)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCP</td>
<td>974</td>
<td>974</td>
<td>28.8</td>
<td>22–37</td>
</tr>
<tr>
<td>BASE-II</td>
<td>675</td>
<td>391</td>
<td>63.2</td>
<td>24–83</td>
</tr>
<tr>
<td>Barcelona</td>
<td>113</td>
<td>39</td>
<td>70.9</td>
<td>64–81</td>
</tr>
<tr>
<td>Cam-CAN</td>
<td>884</td>
<td>632</td>
<td>55.1</td>
<td>20–88</td>
</tr>
<tr>
<td>LCBC</td>
<td>1474</td>
<td>803</td>
<td>49.4</td>
<td>20–89</td>
</tr>
<tr>
<td>UK Biobank</td>
<td>45,983</td>
<td>43,137</td>
<td>64.5</td>
<td>45–83</td>
</tr>
<tr>
<td>Betula</td>
<td>423</td>
<td>284</td>
<td>62.3</td>
<td>25–85</td>
</tr>
<tr>
<td>Whitehall-II</td>
<td>769</td>
<td>769</td>
<td>69.8</td>
<td>60–85</td>
</tr>
<tr>
<td>Total</td>
<td>51,295</td>
<td>47,029</td>
<td>63.4</td>
<td>20–89</td>
</tr>
</tbody>
</table>

HCP: Human Connectome Project; BASE-II: Berlin Aging Study II; Barcelona: University of Barcelona brain studies; Cam-CAN: The Cambridge Center for Ageing and Neuroscience; LCBC: Center for Lifespan Changes in Brain and Cognition, University of Oslo.

Table 2. Definition of sleep groups

<table>
<thead>
<tr>
<th>Groups</th>
<th>Sample</th>
<th>Defining criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (≤6 h)</td>
<td>Lifebrain/HCP</td>
<td>Yes to all of the following:</td>
</tr>
<tr>
<td>(1) Usual sleep latency (PSQI2) ≤ 30 min</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Less than once a week (PSQI5a): sleep latency ≥ 30 min</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Less than once a week (PSQI5b): nightly or early morning awakenings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Less than once a week (PSQI8): trouble staying awake during daytime</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (≤6 h)</td>
<td>UKB</td>
<td>Responses to all the following:</td>
</tr>
<tr>
<td>(1) “Very easy/fairly easy” (field 1170): trouble getting up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) “Never/rarely” (field 1190): nap during day</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) “Never/rarely” (field 1200): sleeplessness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) “Never/rarely” (field 1220): daytime dozing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 (≤6 h)</td>
<td>Lifebrain/HCP</td>
<td>Yes on at least two of the following:</td>
</tr>
<tr>
<td>(1) Usual sleep latency (PSQI2) ≥ 30 min</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) More than once a week (PSQI5a): sleep latency ≥ 30 min</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) More than once a week (PSQI5b): nightly or early morning awakenings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) More than once a week (PSQI8): trouble staying awake during daytime</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

UKB Responses to all the following: |
(1) “Not very easy/not at all easy” (field 1170): trouble getting up |
(2) “Sometimes/usually” (field 1190): nap during day |
(3) “Sometimes/usually” (field 1200): sleeplessness |
(4) “Sometimes/often” (field 1220): daytime dozing |

The PSQI response category “Less than once a week” here also includes the response “Not during the past month.”
Because FreeSurfer is almost fully automated, and by use of tools primarily based on FSL (FMRIB Software library, fmrib.ox.ac.uk/fsl/fslwiki), hippocampal volume is likely sufficient to remove the influence of scanner differences. UKB participants were scanned using three identical Siemens 3T Prisma scanners (https://www.fmrib.ox.ac.uk/ukbiobank/protocol/). FreeSurfer outputs (Alfaro-Almagro et al., 2018) and the volumetric scaling from T1 head image to standard space as proxy for Intracranial Volume (ICV) were used in the analyses, generated using publicly available tools, primarily based on FSL (FMRIB Software library, https://fsl.fmrib.ox.ac.uk/fsl/fslwiki). Details of the imaging protocol (http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=2367) and structural image processing are provided on the UK Biobank website (http://biobank.ctsu.ox.ac.uk/crystal/).

Magnetic resonance imaging acquisition and analysis
LifeIaff MRI data originated from seven different scanners (for details, see Fjell et al., 2019; processed with FreeSurfer 6.0, https://surfer.nmr.mgh.harvard.edu/; Dale et al., 1999; Fischl et al., 2002; Reuter et al., 2012; Jovicich et al., 2013). Because FreeSurfer is almost fully automated, to avoid introducing possible site-specific biases, gross quality control measures were imposed and no manual editing was done. To assess the influence of scanner on volumetric estimates, seven participants were scanned on seven scanners across the consortium sites (for details, see Fjell et al., 2019). Using hippocampus as test-region, there was a significant main effect of scanner on volume (F = 4.13, p = 0.046), but the between-participant rank order was close perfectly retained between scanners, with a mean between-scanner Pearson correlation of r = 0.98 (range 0.94–1.00). Similar analyses of cortical regions also revealed close correspondence across scanners (Nyberg et al., 2023). Thus, including site as a random effect covariate in the analyses of hippocampal volume is likely sufficient to remove the influence of scanner differences.

UKB participants were scanned using three identical Siemens 3T Prisma scanners (https://www.fmrib.ox.ac.uk/ukbiobank/protocol/). FreeSurfer outputs (Alfaro-Almagro et al., 2018) and the volumetric scaling from T1 head image to standard space as proxy for Intracranial Volume (ICV) were used in the analyses, generated using publicly available tools, primarily based on FSL (FMRIB Software library, https://fsl.fmrib.ox.ac.uk/fsl/fslwiki). Details of the imaging protocol (http://biobank.ctsu.ox.ac.uk/crystal/). Structural image processing are provided on the UK Biobank website (http://biobank.ctsu.ox.ac.uk/crystal/).

HCP imaging data were collected and processed (https://www.humanconnectome.org/study/hcp-young-adult) as described previously (Glasser et al., 2013). Imaging data were collected at a customized Siemens 3T “Connectome Skyra” housed at Washington University in St. Louis, using a standard 32-channel Siemens receive head coil and a “body” transmission coil designed by Siemens specifically for the smaller space available using the special gradients of the WU-Minn and MGH-UCLA Connectome scanners. Images were processed using a custom combination of tools from FSL and FreeSurfer (Jenkinson et al., 2002; Fischl, 2012).

An overview of scanning parameters across all subsamples in the present study is given in Table 4.

### Statistical analyses and data availability
Analyses were run in R version 4.0.0 (R Core Team, 2020), by use of Generalized Additive Mixed Models (GAMM) using the packages “gamm4” version 0.2-26 (Wood and Scheipl, 2020) and “mgcv” version 1.8-28 (Wood, 2017). Advantages with GAMM are first that it represents a nonlinear statistical approach which does not require a priori

### Table 3. Description of sleep groups

<table>
<thead>
<tr>
<th>Group</th>
<th>Characteristics</th>
<th>N (%)</th>
<th>Age (SD)</th>
<th>Sex (f/m)</th>
<th>Sleep duration (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Short sleep and no sleep problems/sleepiness</td>
<td>Short sleep (≤6 h)</td>
<td>740 (6%)</td>
<td>55.8 (15.2)</td>
<td>241/499</td>
<td>5.9 (0.4)</td>
</tr>
<tr>
<td></td>
<td>No sleep problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No daytime sleepiness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Short sleep and sleep problems/sleepiness</td>
<td>Short sleep (≤6 h)</td>
<td>1742 (14%)</td>
<td>62.7 (13.1)</td>
<td>848/894</td>
<td>5.8 (0.5)</td>
</tr>
<tr>
<td></td>
<td>Sleep problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Daytime sleepiness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Recommended sleep and no sleep problems/sleepiness</td>
<td>Recommended sleep (7–8 h)</td>
<td>3886 (12%)</td>
<td>58.1 (14.0)</td>
<td>1764/2122</td>
<td>7.6 (0.5)</td>
</tr>
<tr>
<td></td>
<td>No sleep problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>No daytime sleepiness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Recommended sleep and sleep problems/sleepiness</td>
<td>Recommended sleep (7–8 h)</td>
<td>3243 (10%)</td>
<td>65.3 (11.3)</td>
<td>1495/1748</td>
<td>7.5 (0.5)</td>
</tr>
<tr>
<td></td>
<td>Sleep problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Daytime sleepiness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The percentages refer to the total sample with the same sleep duration, i.e., ≤6 h or 7–8 h.

### Table 4. MR Acquisition parameters

<table>
<thead>
<tr>
<th>Sample</th>
<th>Scanner</th>
<th>Field strength (Tesla)</th>
<th>Sequence parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASE-II</td>
<td>Tim Trio Siemens</td>
<td>3.0</td>
<td>TR: 2500 ms, TE: 4.77 ms, TI: 1100 ms, flip angle: 7°, slice thickness: 1.0 mm, FOV 256 × 256 mm, 176 slices</td>
</tr>
<tr>
<td>Betula</td>
<td>Discovery GE</td>
<td>3.0</td>
<td>TR: 8.19 ms, TE: 3.2 ms, TI: 450 ms, flip angle: 12°, slice thickness: 1 mm, FOV 250 × 250 mm, 180 slices</td>
</tr>
<tr>
<td>Cam-CAN</td>
<td>Tim Trio Siemens</td>
<td>3.0</td>
<td>TR: 2250 ms, TE: 2.98 ms, TI: 900 ms, flip angle: 9°, slice thickness 1 mm, FOV 256 × 240 mm, 192 slices</td>
</tr>
<tr>
<td>LCBC</td>
<td>Avanto Siemens</td>
<td>1.5</td>
<td>TR: 2400 ms, TE: 3.61 ms, TI: 1000 ms, flip angle: 8°, slice thickness: 1.2 mm, FOV: 240 × 240 m, 160 slices, iPAT = 2</td>
</tr>
<tr>
<td>Avanto Siemens</td>
<td>1.5</td>
<td>TR: 2400 ms, TE: 3.79 ms, TI: 1000 ms, flip angle: 8, slice thickness: 1.2 mm, FOV: 240 × 240 m, 160 slices</td>
<td></td>
</tr>
<tr>
<td>Skrya Siemens</td>
<td>3.0</td>
<td>TR: 2400 ms, TE: 2.98 ms, TI: 850 ms, flip angle: 8°, slice thickness: 1 mm, FOV: 256 × 256 mm, 176 slices</td>
<td></td>
</tr>
<tr>
<td>Prisma Siemens</td>
<td>3.0</td>
<td>TR: 2400 ms, TE: 2.22 ms, TI: 1000 ms, flip angle: 8°, slice thickness: 0.8 mm, FOV: 256 × 240 mm, 208 slices, iPAT = 2</td>
<td></td>
</tr>
<tr>
<td>UB</td>
<td>Tim Trio Siemens</td>
<td>3.0</td>
<td>TR: 2300 ms, TE: 2.98, TI: 900 ms, slice thickness 1 mm, flip angle: 9°, FOV 256 × 256 mm, 240 slices</td>
</tr>
<tr>
<td>WH-II</td>
<td>Verio Siemens</td>
<td>3.0</td>
<td>TR: 2530 ms, TE: 1.79/3.65/5.51/7.37 ms, TI: 1380 ms, flip angle: 7°, slice thickness: 1.0 mm, FOV: 256 × 256 mm</td>
</tr>
<tr>
<td>HCP</td>
<td>Connectome Skyra Siemens</td>
<td>3.0</td>
<td>TR: 2400 ms, TE: 2.14 ms, TI: 1000 ms, flip angle: 8°, slice thickness: 0.7 mm, FOV: 224 mm, 256 slices, GRAPPA = 2</td>
</tr>
<tr>
<td>UKB</td>
<td>Skrya Siemens</td>
<td>3.0</td>
<td>TR: 2000 ms, TE: 880 ms, slice thickness: 1 mm, FOV: 208 × 256 mm, 256 slices, iPAT = 2</td>
</tr>
</tbody>
</table>


Downloaded from UK Biobank bulk data field 90001. The UKB physical activity data includes a large sample of participants that wore wrist-worn accelerometers (Axivity AX3 wrist-worn triaxial accelerometers) for up to 7 consecutive days. Data were processed using the R-package GGIR v2.4.0 (van Hees et al., 2014, 2015, 2022), using the configuration provided previously (Jones et al., 2019). We excluded accelerometer-derived sleep duration data if any of the UKB derived data fields 90002, 90015, 90016, and 90017 indicated data quality issues. Further, we excluded data when the number of data recording errors, number of interrupted recording periods, or the duration of interrupted recording periods exceeded Q3 + 1.5IQR of the sample. The GGIR computed variable SptDurationInSpt (Total sleep duration) was used as the accelerometer derived sleep duration measure.
Table 5. Differences in brain volumes between groups of short sleepers

<table>
<thead>
<tr>
<th>Region</th>
<th>Mean volume</th>
<th>Group 2 – Group 1</th>
<th>Difference mm³ (CI)</th>
<th>% difference (CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accumbens</td>
<td>903</td>
<td></td>
<td>−7 (−19, 6)</td>
<td>−0.8 (−2.2, 0.7)</td>
</tr>
<tr>
<td>Amygdala</td>
<td>3290</td>
<td></td>
<td>4 (−25, 34)</td>
<td>0.1 (−0.8, 1)</td>
</tr>
<tr>
<td>Brain stem</td>
<td>21,935</td>
<td></td>
<td>−181 (−348, −13)</td>
<td>−0.8 (−1.6, −0.1)</td>
</tr>
<tr>
<td>Caudate</td>
<td>6757</td>
<td></td>
<td>−56 (−127, 14)</td>
<td>−0.8 (−1.9, 0.2)</td>
</tr>
<tr>
<td>Corpus callosum</td>
<td>3558</td>
<td></td>
<td>−35 (−77, 7)</td>
<td>−1.0 (−2.2, 0.2)</td>
</tr>
<tr>
<td>Cerebellum cortex</td>
<td>111,496</td>
<td></td>
<td>−704 (−1567, 160)</td>
<td>−0.6 (−1.4, 0.1)</td>
</tr>
<tr>
<td>Cerebellum WM</td>
<td>30,979</td>
<td></td>
<td>−209 (−626, 45)</td>
<td>−0.9 (−2.0, 0.1)</td>
</tr>
<tr>
<td>Cerebral WM</td>
<td>475,095</td>
<td></td>
<td>−105 (−2933, 2744)</td>
<td>−0.0 (−0.6, 0.6)</td>
</tr>
<tr>
<td>ICV</td>
<td>1,545,244</td>
<td></td>
<td>−2674 (−15179, 9949)</td>
<td>−0.2 (−1.0, 0.6)</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>8067</td>
<td></td>
<td>8 (−68, 51)</td>
<td>−0.1 (−0.8, 0.6)</td>
</tr>
<tr>
<td>Pallidum</td>
<td>3994</td>
<td></td>
<td>−36 (−68, 4)</td>
<td>−0.9 (−1.8, −0.1)</td>
</tr>
<tr>
<td>Putamen</td>
<td>9221</td>
<td></td>
<td>−27 (−109, 54)</td>
<td>−0.3 (−1.2, 0.6)</td>
</tr>
<tr>
<td>Thalamus</td>
<td>13,718</td>
<td></td>
<td>−75 (−163, 13)</td>
<td>−0.6 (−1.2, 0.1)</td>
</tr>
<tr>
<td>TGV</td>
<td>662,557</td>
<td></td>
<td>−1799 (−4678, 1081)</td>
<td>−0.3 (−0.7, 0.2)</td>
</tr>
<tr>
<td>Ventricles</td>
<td>31,887</td>
<td></td>
<td>116 (−1039, 1270)</td>
<td>0.4 (−3.3, 4.0)</td>
</tr>
</tbody>
</table>

*Group 2: Short sleep and sleep problems/sleepiness* was compared with *group 1: Short sleep and no sleep problems/sleepiness.* Negative estimates represent smaller volumes in group 2.

The interaction term age x Group. Where appropriate, critical p-values were determined by use of the Benjamini–Hochberg procedure with a 5% false discovery rate to control for multiple comparisons.

As we did not specifically model follow-up time between examinations, the results represent the optimal fit to the longitudinal and cross-sectional observations, and should not be interpreted to signify change in brain volumes per se.

Analyses were run for 12 brain regions, the ventricles, total gray matter volume (TGV) and ICV. We tested for whole-brain effects by computing meta analytic estimates of standardized regression coefficients across the 12 regions, using the R package “metfor” (Viechtbauer, 2010). Data supporting the results of the current study are available from the PI of each study on request, given appropriate ethics and data protection approvals. Contact information can be obtained from the corresponding authors. UK Biobank data requests can be submitted to http://www.ukbiobank.ac.uk.

Results

Sleep duration versus sleepiness

For the Lifebrain and HCP samples, we defined “sleepiness” from the Pittsburgh Sleep Quality Index (PSQI) by using the one to four values from item PSQI8 (“trouble staying awake”). Sleepiness and sleep duration correlated $r = −0.11$ ($p < 0.0001$), meaning that short sleep on average was associated with slightly more sleepiness, with $\sim1\%$ explained variance. Sleep duration and sleepiness hence represents largely independent features.

Brain volumetric comparisons among short sleepers (group 1 vs 2)

Differences in brain volumes between the groups are shown in Table 5 and Figure 1. We performed a meta-analysis across all regions listed in Table 5 (for the ventricles, the sign of the regression coefficient was reversed). Group 2: Short sleep and sleep problems/sleepiness had significantly smaller brain volumes than group 1: Short sleep and no sleep problems/sleepiness (estimate $= 0.0041$ [CI: $−0.0001$, $0.0001$], $SE = 0.001$, $z = 4.1$, $p < 0.0001$). Regional comparisons revealed that the differences were most evident for the brain stem and pallidum.

Next, we tested whether each of the two groups of short ($<6$ h) sleepers had different volumes from group 3: Recommended sleep and no sleep problems/sleepiness, reporting 7–8 h of sleep. Group 1: Short sleep and no sleep problems/sleepiness had overall larger volumes than group 3 (estimate $= 0.0067$ [CI: $0.0030$, $0.0103$], $SE = 0.0019$, $z = 3.60$, $p < 0.001$), driven by differences in brain stem, caudate, cerebellum white matter and cortex volumes.

Brain volumetric comparisons among participants reporting recommended amount of sleep (group 3 vs 4)

Group 4: Recommended sleep and sleep problems/sleepiness showed smaller volumes for cerebellum WM, corpus callosum
and pallidum, as well as ICV, compared with group 3: Recommended sleep and no sleep problems/sleepiness (Table 6; Fig. 1). However, the meta-analysis did not show a significant overall difference across regions (estimate = −0.0012 [CI: −0.0028, 0.0004], SE = 0.0008, z = −1.51, p = 0.13).

Comparisons among participants with daytime sleepiness (group 2 vs 4)

Finally, we compared group 2: Short sleep and sleep problems/sleepiness with group 4: Recommended sleep and sleep problems/sleepiness. The groups differed by 1.7 h in mean sleep duration, but the meta-analysis did not show a significant difference in brain volumes (estimate = 0.0022 [CI: −0.0002, 0.0046], SE = 0.0012, z = 1.80, p = 0.07).

Age interactions

Age was included as covariate in all analyses, but we also ran additional analyses including an interaction term age × sleep group to formally test whether effect of sleep group differed as a function of age. For no brain region or group contrast did the age × sleep group term survive correction for multiple comparisons.

Cognitive function

General cognitive function (GCA) scores were available for 8694 of the classified participants. GCA was calculated as the principal component of different cognitive scores available for each sample. There were no significant differences between the groups of short sleepers (group 2 vs 1; estimate = −0.077 SD, t = −1.55, p = 0.12, n = 2173). Among those sleeping 7–8 h, group 4: Recommended sleep and sleep problems/sleepiness showed significantly lower GCA than group 3: Recommended sleep and no sleep problems/sleepiness (estimate = −0.078 SD, t = 6.89, p < 5.89e−12, n = 6521). The effect size was however not larger than for the contrast between the two groups of short sleep, so the difference in significance was because of a substantially larger sample for the latter analysis. A significant difference was also seen between the two groups reporting daytime sleepiness (group 2 vs 4; estimate = −0.03 SD, t = −2.97, p = 0.003). Both group 1: Short sleep and no sleep problems/sleepiness (estimate = −0.16 SD, t = −3.80, p < 0.0002) and group 2: Short sleep and sleep problems/
still it is unclear whether vigilance. Desai et al., Watson et al., 2006) and 0.08 Carmelli et al., 2001 Landolt, (2011). This explanation fits previous findings 2011). Such trait variability may partly be Kocevska et al., 2021a Linkowski, 1999). It also concurs with the results from recent meta-analyses of twin studies found 38% H. Van Dongen et al., 2003). Importantly, the genetic overlap recent study showing very modest effects of variations in Vallat, 2017). Individual differences in multiple physiological aspects of sleep, such as sleep homeostasis and duration, which in magnitude have been reported to exceed even effects of 36 h of sleep deprivation (Tucker et al., 2007). Such trait variability may partly be accounted for by genetic differences (Linkowski, 1999; Landolt, 2011; Dashti et al., 2019). This explanation fits previous findings that daytime sleepiness is not solely caused by short sleep (Horne, 2010) and that increased duration does not necessarily cause longer sleep latency and less sleepiness even within individuals (Roehrs et al., 1996). It also concurs with the results from a recent study showing very modest effects of variations in within-participant sleep duration on subjective alertness (Vallat, 2022). Recent meta-analyses of twin studies found 38% (Madrid-Valero et al., 2020) and 46% (Kocevska et al., 2021a) of the variability in self-reported sleep duration to be explained by genetics, with GWAS estimates typically being around 10% (Garfield, 2021). Heritability of daytime sleepiness is reported to be 0.38–0.48 in twin studies (Carmelli et al., 2001; Desai et al., 2004; Watson et al., 2006) and 0.08–0.29 (Gottlieb et al., 2007; Lane et al., 2017) with GWAS. Importantly, the genetic overlap between sleep duration, daytime sleepiness, and vulnerability to sleep loss is modest. One study reported a genetic correlation between sleep duration and daytime sleepiness of 0.22 (Wang et al., 2019), and none of the loci associated with duration were associated with sleepiness. The low duration–sleepiness correlation in the present study is in accordance with this, and fits the interpretation that habitual sleep duration and sleepiness are partly independent, trait-like characteristics with different associations to brain health and cognitive function. Differences in sleep quality may further contribute to reduce the relationship between sleep duration and sleepiness. Without taking the individual differences perspective into account, large natural variability in aspects of sleep may lead to spurious, sample-dependent sleep-brain correlations without functional significance (Landolt, 2011).

Still, although short sleep was not associated with smaller regional brain volumes, the short sleepers scored lower on tests of general cognitive abilities. The effect sizes were equal to 2.9 and 2.4 IQ-points for the short sleepers with and without sleep need.
problems and daytime sleepiness, respectively. This could indicate that these participants sleep less than optimal, in accordance with experimentally induced sleep deprivation yielding reduced cognitive function (Lowe et al., 2017). However, a meta-analysis did not find an effect of sleep deprivation on intelligence and reasoning measures (Lowe et al., 2017), which are the tests most similar to the present measures of cognitive function. In addition, most sleep deprivation experiments involve rapid and dramatic reductions in sleep duration, unlike natural variations in habitual sleep duration between people. A recent very large observational study found ~7 h of sleep to be associated with the highest cognitive function, and <6 h to be associated with mildly lower performance (Coutrot et al., 2022). This fits the results of the present study. Still, we were surprised to find that there were no differences in general cognitive ability between the short sleepers with and without sleep problems and sleepiness, as we expected to see lower scores primarily in the first group. The direction of causality and the possible influence of third variables cannot be decided based on the present data, and must await experimental testing in naturalistic settings, involving modest changes in sleep duration lasting for prolonged periods to mimic everyday sleep duration–cognition relationships.

A complementary account for why the short sleepers in the present study did not show smaller regional brain volumes is that neurocognitive consequences of short sleep depend on adaptation. Sudden sleep deprivation beyond certain limits has negative effects on cognitive performance (Van Dongen et al., 2003; Killgore, 2010) and brain structure (Liu et al., 2014; Saletin et al., 2016; Voldsbek et al., 2021; Zamore and Veasey, 2022). However, adaptation over time within these limits are unlikely to be harmful to health (Freidmann et al., 1977; Mullaney et al., 1977; Horne, 2011), and may account for the weak relationship between sleep duration and brain morphometry. Reductions in sleep duration can seemingly be obtained without increases in daytime sleepiness or reductions in cognitive performance (Horne and Wilkinson, 1985; Youngstedt et al., 2009). This suggests that sleep duration is adaptable in response to environmental conditions, in line with the present results that variations in sleep duration per se is of less importance for regional brain volume if daytime sleepiness is low. Thus, a combination of adaptation and genetic propensities may create less sleep need and protect from potentially negative consequences of short sleep.

As the present study is correlational and targets brain morphometry, we believe it is not warranted to make strong anatomic interpretations from the results. Still, the regions showing volumetric differences between the short sleepers with versus without sleepiness and sleep problems included the brain stem, which has a critical role in sleep regulation, especially in control of REM sleep (Siegel, 2022). Further, regions of the basal ganglia, especially pallidum and caudate, showed trends toward volumetric differences, and it is suggested that the basal ganglia may play an integral role in the sleep-wake cycle (Hasegawa et al., 2020), with basal ganglia GABAergic neurons possibly representing a functional hub for sleep control (Adamantidis et al., 2021). However, many brain regions are involved in sleep regulation, so we caution about specific neuroanatomical claims based on the present study.

Limitations

(1) Morphometric brain measures and general cognitive ability were used as measures of brain health and cognitive function. Other measures could have yielded different results. e.g., different cognitive domains likely show different relationships to sleep, and the same may be true for other indexes of brain health such as Aβ deposition (Fjell et al., 2020a). (2) The samples were not thoroughly screened for sleep disorders such as sleep apnea, which could affect the observed sleep – brain relationships. (3) The samples are not representative of the populations from which they are drawn, and other sleep-brain patterns may exist in other populations. Further, the majority of participants are white, while sleep loss and sleep problems have been shown to be more prevalent in the United States Black than White population (Jean-Louis et al., 2022). (4) Daytime sleepiness is measured as part of PSQI, but other instruments, like the Epworth Sleepiness Scale, may be more sensitive to excessive daytime sleepiness (Pérez-Carbonell et al., 2022). (5) The strict classification of participants was used to ensure that the groups were as homogeneous as possible with regard to sleep variables, but caused most participants to be unclassified. For instance, participants who responded “once or twice a week” on the question of problems sleeping within 30 min, would not be included in the group without sleep problems. Items such as “nap during the day” may both reflect a cultural phenomenon as well as sleep done for pleasure, not necessarily because of sleepiness or sleep need, but would still be assessed as “daytime sleepiness or sleep problems.” As a consequence, only 20% of participants sleeping ≤6 h were classified. The remaining 80% had very slight to slight sleep problems or daytime sleepiness, hence not fitting either group. (6) Relationships between sleep and the brain may develop over long time. In the present analyses, sleep duration was for most participants measured at a single time point, making it impossible to distinguish stable from changing sleep patterns. Still, we have previously found good stability in self-reported sleep over time (Fjell et al., 2018; Tsigkia et al., 2023), and another study found current self-reported sleep quality to be as tightly related to brain characteristics as longitudinal measures of sleep (Sexton et al., 2017).

In conclusion, some people sleep ≤6 h without showing lower regional brain volumes, despite sleeping within a range where smaller regional brain volumes are expected. Hence, short sleep is not necessarily associated with negative structural brain outcomes. In contrast, short sleepers showed slightly lower general cognitive abilities, although the causality is unclear. At the same time, the present results suggest that there are large differences in sleep need because of genetic and environmental factors, making general recommendations about sleep duration problematic when it comes to brain health and general cognitive function.

References


Cattell RB, Cattell HEP (1973) Measuring intelligence with the culture fair tests. Champaign: The Institute for Personality and Ability Testing.


Fjell AM, et al. (2023) No phenotypic or genotypic evidence for a link between sleep duration and brain atrophy. Nature Human Behavior, in press.


Koceska D, et al. (2021b) Sleep characteristics across the lifespan in 1.1 million people from The Netherlands, United Kingdom and United States: a systematic review and meta-analysis. Nat Hum Behav 5:113–122.


Wood SN, Scheipl F (2020) ’mgcv': generalized additive mixed models using 'mgcv' and 'lme4'. R package version 0.2-6. Available at: https://CRAN.R-project.org/package=mgcv.

