Sleep Efficiency Relates to Hippocampal Integrity Decline in β-Amyloid Positive Adults

Håkon Grydeland, PhD, Donatas Sederevičius, MSc, Yunpeng Wang, PhD, David Bartrés-Faz, PhD, Lars Bertram, MD, Valerija Dobricic, PhD, Sandra Düzel, PhD, Klaus P. Ebmeier, MD, Ulman Lindenberger, PhD, Lars Nyberg, PhD, Sara Pudas, PhD, Claire E. Sexton, DPhil, Cristina Solé-Padullés, PhD, Kristine B. Walhovd, PhD, Anders M. Fjell, PhD

Research Group for Lifespan Changes in Brain and Cognition, Department of Psychology, University of Oslo (H.G., D.S., Y.W., L.B., K.B.W, A.M.F)
Department of Radiology and Nuclear Medicine, University of Oslo (K.B.W, A.M.F)
Departament de Medicina, Facultat de Medicina i Ciències de la Salut, Universitat de Barcelona, Spain (D.B.-F., C.S.-P.)
Lübeck Interdisciplinary Platform for Genome Analytics (LIGA), Institutes of Neurogenetics and Cardiogenetics, University of Lübeck, Lübeck, Germany (L.B., V.D.)
Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany (S.D., U.L.)
Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin, Germany, and London, UK (U.L.)
Department of Psychiatry, University of Oxford, UK (K.P.E., C.E.S.)
Umeå Center for Functional Brain Imaging, Umeå University, Sweden (L.N., S.P.)

Corresponding author: Håkon Grydeland, Department of Psychology, University of Oslo, PO Box 1094 Blindern, 0317 OSLO, Norway, telephone: +47 22845067, fax: +47 22845001, email: hakon.grydeland@psykologi.uio.no

The statistical analysis was conducted by H.G, and A.M.F (University of Oslo).

Supplemental Data file included: Grydeland_etal_supplemental_data.pdf

Word count (paper, abstract): 5559, 250.

Character count title: 86

Number of references, tables, and figures: 79, 1, 3

Study funding: The Lifebrain project is funded by the EU Horizon 2020 Grant: 'Healthy minds 0–100 years: Optimising the use of European brain imaging cohorts ("Lifebrain")'. Grant agreement number: 732592.
LCBC: The European Research Council under grant agreements 283634, 725025 and 313440, as well as the Norwegian Research Council and The National Association for Public Health's dementia research program, Norway. Betula: a scholar grant from the Knut and Alice Wallenberg (KAW) foundation to L.N. Barcelona: Partially supported by a Spanish Ministry of Science, Innovation and Universities (MICIUN) grant to D-BF [grant number RTI2018-095181-B-C21 (AEI/FEDER, UE)], and an ICREA Academia 2019 Award; by the Walnuts and Healthy Aging study (http://www.clinicaltrials.gov; Grant NCT01634841) funded by the California Walnut Commission, Sacramento, California. BASE-II has been supported by the German Federal Ministry of Education and Research under grant numbers

16SV5537/16SV5837/16SV5538/16SV5536K/01UW0808/01UW0706/01GL1716A/01GL1716B, and from the European Research Council under grant agreement 677804. Work on the Whitehall II Imaging Substudy was mainly funded by Lifelong Health and Well- being Programme Grant G1001354 from the UK Medical Research Council ("Predicting MRI Abnormalities with Longitudinal Data of the Whitehall II Substudy").

Search terms

[26] Alzheimer's disease, [244] All Sleep Disorders, [201] Memory, [128] DWI, [122] PET

Disclosure Statement

The authors report no disclosures.

Abstract

Objective: To test the hypothesis that worse self-reported sleep relates to reduced hippocampal integrity as indexed by increased intra-hippocampal water diffusion, and that this relationship is stronger in the presence of β-amyloid (Aβ) accumulation, a marker of Alzheimer's disease (AD) pathology.

Methods. Two-hundred and fifty-one participants, aged 19-81 years, completed the Pittsburgh Sleep Quality Index, and 2 diffusion tensor imaging sessions, on average 3 years apart, allowing estimates of decline in hippocampal microstructural integrity as indexed by increased mean diffusivity (MD). We used the delayed recall from the California Verbal Learning Test to measure memory change. ^{18}F-Flutemetamol PET, in 108 participants above 44 years of age, yielded 23 Aβ positive cases. Genotyping enabled controlling for *APOE* ε4 status, and polygenic scores for sleep efficiency and AD.

Results. Worse global sleep quality and sleep efficiency related to more rapid reduction in hippocampal microstructural integrity over time. Focusing on sleep efficiency, this relationship was stronger in presence of cortical Aβ accumulation. Sleep efficiency also related to memory decline indirectly via hippocampal integrity decline. The results were not explained by genetic risk for sleep efficiency and AD.

Conclusions. Poor self-reported sleep efficiency related to decline in hippocampal integrity, especially in the presence of Aβ accumulation. Poor sleep and hippocampal microstructural decline may partly explain memory decline in older adults with Aβ pathology. The relationships were not explained by genetic risk, and poor self-reported sleep efficiency might constitute a risk factor for AD, although the causal mechanisms driving the of observed associations are unknown.

Individuals with sleep disturbances have increased risk for Alzheimer's disease (AD)[1, 2], and accumulation of β-amyloid (Aβ)[3, 4]. Aβ is modestly related to memory decline[5], and studies have suggested that relationships between Aβ and memory partly depend on sleep[6, 7]. A critical role in this linkage of sleep to Aβ and memory may be played by the integrity of the hippocampus. While a consortium study showed that worse self-reported sleep related to hippocampal atrophy[8], integrity measured by diffusion tensor imaging (DTI) may detect more subtle microstructural decline[9], and hippocampal mean diffusivity (MD) has demonstrated particularly sensitive to memory[10, 11]. Sleep-hippocampal integrity relationships could also reflect effects of the *APOE* ε4 genotype[12], or of common genetic variation affecting sleep and hippocampus[13]. Testing whether worse self-reported sleep relates to memory decline and more rapid reduction of hippocampal integrity while controlling for genetic variation, and whether such relationships are stronger in older adults with pathological levels of Aβ, will help us understand the role of sleep problems in early AD-related pathology.

Here, in 251 cognitively normal participants aged 19-81 years, we asked whether self-reported sleep characteristics were associated with memory-related microstructural (MD) hippocampal integrity reduction over an average of 3 years. We hypothesized worse sleep to be related to stronger degeneration, particularly in individuals with cortical Aβ accumulation, and also when controlling for *APOE* ε4 and polygenic scores for sleep efficiency and AD[14]. To further assess self-reported sleep relations with memory decline, we also performed a meta-analysis using data from the Lifebrain consortium[15].

Methods

Sample. The sample was drawn from projects consisting of 2-6 study waves at the Center for Lifespan Changes in Brain and Cognition, Department of Psychology, University of Oslo, Norway. The Regional Ethical Committee of Southern Norway approved all procedures, and all participants consented in writing prior to commencement. A total of 251 participants were included, all community-dwelling with normal cognition (see **Table 1** for detailed characteristics, see **figure e-1** for attrition of participants). For the first wave of data collection, participants were recruited through advertisements. Recruitment for the follow-up assessments was by written invitation to the original participants. At each time point, participants underwent health interviews to ascertain eligibility. Participants were required to be right-handed, fluent Norwegian speakers, and have normal or corrected to normal vision and hearing. Exclusion criteria were history of injury, disease or psychoactive drug use affecting central nervous system function, including neurological or psychiatric illness or serious head trauma, or being under psychiatric treatment, as well as presence of contraindications to magnetic resonance imaging. On the Mini Mental State Examination (MMS)[16], participants above 40 years scored ≥26, except 2 participants aged 80 years scoring 25. At follow-up, 3 participants aged 42-63 years who scored ≥29 at baseline, were missing MMS. All participants who completed the Beck Depression Inventory (BDI) scored ≤16, except 4 participants, aged 24-45 at follow-up, scoring 18-24, respectively. Ninety-five participants aged above 68 years completed the Geriatric Depression Scale (GDS)[17]. All scored ≤ 9 except for 7 participants. Of these, 5 participants aged 71-74 years scored 11-22 on follow-up. Two participants aged 77 and 73 years, scored 13 and 10, respectively, at baseline, but both scored at non-depression levels on follow-up. A depression score was missing for a total of 16 participants at one time point (14 participants, aged 19-77

years) or both (2 participants, aged 29 and 58 years). To account for potential influences of particularly depression, we undertook sensitivity analyses (see below). General cognitive abilities were assessed by Wechsler Abbreviated Scale of Intelligence (WASI)[18]. The grand mean full-scale IQ (mean across time points for those with full-scale IQ score at both time points) was 119 (SD=9; 3 participants lacked full-scale IQ scores at follow-up). A neuroradiologist deemed all included magnetic resonance imaging (MRI) scans free of significant injuries or pathological conditions.

Figure 1 shows an overview over the study design. Similar to our previous work on self-reported sleep[19], baseline MRI was administered between 2011 and 2016, and follow-up MRI between 2015 and 2018. PSQI was completed once by each participant, between 2012 and 2017, on average 0.6 (SD=0.8) years after baseline MRI (16 participants completed the PSQI on average 5 (SD=3.5) months before baseline MRI, while exact completion date was not available for 34 participants). The memory assessments were performed on average 13 (SD=22) days before the baseline MRI, and 26 (SD=30) days before the follow-up MRI, respectively. PET scanning was performed once in a subset of participants, between 2015 and 2018, on average 0.7 (SD=1) years before the MRI follow-up.

[Insert Figure 1 about here]

Sleep assessment. To assess sleep, we used the Pittsburgh Sleep Quality Index (PSQI)[20]. This self-reported index surveys sleep habits and quality of the last month. The PSQI yields one global sleep quality score, which is the sum of the score of 7

components: 1) quality, 2) latency, 3) duration, 4) habitual efficiency, 5) disturbance, 6) use of sleeping medication, and 7) daytime dysfunction. Higher scores indicate lower sleep quality. In PSQI, efficiency is calculated as sleep duration (hours slept) divided by the number of hours spent in bed, times 100, and then given score of 0-3 for >85%, 75-84%, 65-74%, and <65%, respectively[20]. We did not evaluate the sixth component as use of medication was an exclusion criterion (the component was included in calculation of the global score for consistency with previous studies). Although the PSQI asks about sleep patterns of the last month, here, as in our previous longitudinal work[19], we take the PSQI to reflect relatively stable sleep patterns, an inference for which there is support in adults above 38 years[21, 22]. In line with this premise, the PSQI self-report was not necessarily completed in close proximity to the baseline MRI scan (see **Table 1** for details). As shown in **Table1**, complete PSQI scores at one time point, and valid DTI scans at two time points, were available for 251 participants (61% female, mean age = 54, range: 19-81, standard deviation (SD) = 20 years).

[Insert **Table 1** about here]

MRI acquisition. Diffusion tensor imaging scans were acquired at two Siemens scanners (Siemens Medical Solutions, Erlangen, Germany), a 1.5 T Avanto (n=64, 70% female, mean age (SD, min-max) = 51 (13, 24-77) years), TR/TE=8200/81 ms, FOV=128, 60 diffusion-sensitizing gradients at a b-value of 700 s mm^{-2} and 2 volumes without diffusion weighting (b-value = 0)), and 3T Skyra scanner (n=187, 58% female, mean age (SD, min-max) = 55 (22, 19-81) years), TR/TE=9200/87 ms, FOV=130, 64 diffusion-sensitizing gradients at a b-value of 1000 s mm^{-2} and 1

volume without diffusion weighting. The sequences and scanner were the same across the two time points for each participant.

Preprocessing. The diffusion-weighted data were analyzed using the FMRIB Software Library, and included susceptibility-induced field correction with *topup* (Andersson et al., 2003), and correction for head motion, signal dropout, eddy current-induced fields using *eddy* [23, 24]. After removing nonbrain tissue, eigenvalue maps were computed. We defined mean diffusivity (MD) as the mean of the three eigenvalues. We employed a DTI-derived measure as results indicate that DTI can detect subtle effects in cellular microstructure, which has previously proven sensitive to age-related lifespan changes[25], and particularly hippocampal MD has been shown to provide increased sensitivity to memory[10, 11].

Hippocampus segmentation and DTI registration. The T1-weighted image was automatically processed with FreeSurfer software suite (version 6.0.0), independently for each time point (as no co-registration across time points was necessary), yielding segmentation of left and right hippocampus[26]. To extract MD from the hippocampi in native DTI space for each participant, a B=0 volume from the diffusion data was registered to the T1-weighted image in FreeSurfer space with a within-subject, cross-modal registration using a boundary-based cost function constrained to be 6 degrees of freedom[27]. The resulting registration matrix was inverted, and applied to the segmentation of the left and right hippocampus, yielding hippocampus masks in native diffusion space. The masks were binarized using *mri_binarize* at a minimum voxel threshold of 1, for the most restricted masks compared with lower thresholds.

To reduce the number of tests, we calculated the average hippocampal MD based on the left and right hippocampus at each time point.

Memory change. The participants underwent neuropsychological testing including memory assessment via the California Verbal Learning Test, second edition (CVLT-II)[28]. In an effort to minimize practice effects due to repeated testing, we administered alternative versions containing different words and categories at follow-up. From this word learning test, we chose the arguably most sensitive measure of hippocampus-dependent memory, namely long delay free recall, that is, the number of correctly recalled words after an approximately 30-minute delay (during which other cognitive tests were performed). All but 7 participants had valid scores at baseline and follow-up on this delayed, free recall measure.

Symmetrized percent change (SPC). As in our previous longitudinal sleep work[19], we calculated symmetrized percent change (SPC), as symmetrized measures have been shown to be more robust, and with equal or greater statistical power[29]. For the average hippocampus value at baseline and follow-up (AH1 and AH2), the SPC was obtained by the following formula: SPC = 100 * (AH2 - AH1)/(AH2 + AH1). The same formula was used to obtain SPC measure for hippocampal volume and memory change.

PET acquisition. A total of 108 participants (mean age (SD, min-max)=68.0 (8.7, 44.4-80.8) years) underwent ^{18}F-flutemetamol-PET scan, sensitive to Aβ accumulation[30]. Images were acquired on a General Electric Discovery PET/CT 690 scanner at Aleris Hospital and Radiology, Oslo, Norway. A low-dose computerized

tomography scan was first performed for subsequent attenuation correction of the PET scan. Participants were injected with 200±20 MBq ^{18}F-flutemetamol as a bolus and examined 90 minutes later. Three-dimensional dynamic data were acquired in list mode for 20 minutes, with the following parameters: 47 image planes, voxel size = 1.33 mm x 1.33 mm x 3.27 mm, field of view = 256 mm. The images were reconstructed using the VUEPoint HD Sharp iterative reconstruction algorithm. This algorithm adds resolution recovery in an iterative reconstruction loop by incorporating information about the PET detector response which improves resolution and contrast recovery compared with traditional analytic methods[31]. We used 4 iterations, 16 subsets, time of flight, and a full width at half maximum Gaussian post-filter of 3 mm. As we were interested in the gross tracer uptake, we binned the data into a single frame, and submitted this static PET image to further pre-processing and value extraction.

Genetic data. A subsample of 180 participants (66% females, mean age (SD, min-max) = 53.9 (20.5, 20.1-80.9) years had genome-wide single nucleotide polymorphism (SNP) and manual *APOE* ε4 genotypes available. Buccal swab and saliva samples were collected for DNA extraction followed by genome-wide genotyping using the “Global Screening Array” (Illumina, Inc.). A*POE* ε4 (rs429358) status was determined using TaqMan (Thermo Fisher Scientific, Inc.) chemistry. Detailed information on DNA collection, quality control, genotyping, and imputation has been reported elsewhere[32]. The polygenic scores of sleep efficiency and AD were computed using summary statistics from previously published genome-wide association studies (GWAS) [33, 34]. These statistics were based on SNPs with p-values <0.01 in the respective GWAS, except for variants located in the extended

MHC region (build hg19; chr6:25,652,429-33,368,333), where we included the most significant SNP. After removing the *APOE* gene region (build hg19; chr19:44,909,011-45,462,650) for which we used the manually derived ε4 (rs429358) genotypes instead, we used the software PLINK[35] to implement the following steps: (i) clumping of the GWAS summary statistics by the –clump option with parameters --clump-p1 1.0 –clump-p2 1.0 –clump-kb 500 –clump-r2 0.1. The linkage disequilibrium (LD) structure was based on the European subpopulation from the 1000 Genomes Project Phase3[36]. (ii) Deriving polygenic scores for our sample using the –score function. To control for population substructures, we computed the genetic ancestry factors using principal components methods[37], and included only participants of European ancestry in the genetic subsample analysis. The polygenic score (PGS) for sleep efficiency was based on a genome-wide association study using accelerometer-derived mean sleep efficiency (calculated as proportion of sleep period time-window classified as sleep)[34], and in our sample a higher PGS reflected a higher genetic propensity towards more efficient sleep. The AD PGS was based on a genome-wide meta-analysis of clinically diagnosed AD and AD-by-proxy (based on parental diagnoses)[33], and in our sample a higher PGS reflected a higher AD risk. To test for the effect of *APOE* separately from the common genetic variation reflected by the polygenic scores, we estimated *APOE* ε4 counts by determining the haplotypes of the two SNPs rs7412 and rs429358[38, 39], coded as 0, 1, or 2 copies of the ε4 allele, and binarized to ε4-non-carrier or ε4-carrier.

PET pre-processing. We used *PetSurfer*, a set of tools within the FreeSurfer suite, for partial volume correction. Specifically, for each participant, we registered the static PET image to the anatomical T1-weighted image using boundary-based

registration[27]. This registration was inverted to get a high-resolution segmentation (upsample factor = 2) from the high-resolution MRI space in PET space, and simultaneously perform the partial volume correction with the Symmetric Geometric Transfer Matrix method, as recommended when using regions of interests (instead of vertex-wise) approach[40, 41]. This procedure yielded PET signal for each of the 68 cortical regions in Desikan-Killiany atlas[42]. We used cortical regions as Aβ has been reported to appear first in cortex[43]. The PET signal in each cortical region was divided by the mean signal of the cerebellum cortex to obtain standardized uptake value ratios (SUVR)[44].

Aβ status. As common in the literature[44], we dichotomized the SUVR into high or low Aβ groups using a data-driven approach. We ran a principal component analysis on SUVR from the 68 cortical regions using the *prcomp* function (R package *stats* v3.6.1, values were zero-centered and scaled to have unit variance), and extracted the first component (which explained 66.7% of the variance, while, for comparison, the second component explained 7%). The cut-off between groups was determined using Gaussian mixture modeling (R package *mclust* v5.2). We fitted 18 models, ranging from 1 to 9 mixtures, allowing for either equal or unequal variance, and selected the model with the lowest Bayesian information criterion value. As previously reported in healthy older participants[44], the optimal model consisted of a 2-distribution model with unequal variance. Participants with a >.5 probability of belonging to the high Aβ distribution were classified as *Aβ positive*, and the remaining as *Aβ negative* (**Figure e-2**).

Meta-analysis of self-reported sleep and memory change. To test the relationship between the relevant sleep variable (see below for selection criteria) and memory change, we also included data from the Lifebrain consortium (http://www.lifebrain.uio.no/)[15], an EU-funded (H2020) project including participants from several major European brain studies: Berlin Study of Aging-II (BASE-II)[45, 46], the BETULA project[47], University of Barcelona brain studies[48-50], and Whitehall-II[51], yielding a total of 1196 participants. The samples and procedures used are described in detail elsewhere[8]. The data available in all projects were (i) self-reported sleep scores from one time point, and (ii) memory change score between two time points. All subsamples used the PSQI for sleep evaluation, except the Betula sample, which used the Karolinska Sleep Inventory (for details of conversion to PSQI scores, see[8]). The following memory tests were used: 30-minutes delayed free recall from the Verbal Learning and Memory Test (BASE-II), an immediate free recall of sentences (Betula), 30-minutes delayed recall from the Rey Auditory Verbal Learning Test (Barcelona), a short-term 20 word free recall test (Whitehall-II)[52].

Statistics. Our main question of a relationship between sleep and microstructural hippocampus change was addressed by multiple regression models testing each of the 7 PSQI variables (1 global and 6 components) versus hippocampal MD change. To correct for the multiple tests in this analysis, we adjusted the 7 resulting p-values by applying false discovery rate (FDR)[53] correction and the *p.adjust* function (R *stats* version 3.6.1). Head movement is a potential important confound in brain imaging studies. As a proxy measure of head movement during MRI diffusion scanning, we calculated temporal signal-to-noise ratio (tSNR) from the scans[54]. As expected, head movement increased with age (R^2=0.40, $p<0.001$), and we included

tSNR in all hippocampal analyses to account for movement-related artifacts (see **Figure 1B** for overview of main regression models and corresponding covariates). We included interval between baseline and follow-up as covariate of no interest, in addition to age, and sex. As participants were drawn from various waves, we included number of prior visits as a covariate to account for potential learning effects on the memory task (please note that, as mentioned above, different versions were used at each visit). In the analyses of hippocampal MD change, we also included as covariates as no interest hippocampus volume at baseline MRI, and difference in movement and hippocampal volume between baseline and follow-up MRI. These covariates were included to (i) assess microstructural effects specifically, and (ii) to correct for volume differences potentially leading to differences in partial volume effects. For the one hippocampal volume analysis run to compare with previous studies, we also included estimated intracranial volume[55]. To test whether the relationships between sleep and hippocampal MD change was similar across the adult lifespan, we assessed the interaction between the PSQI measure and age. To test for mediation of hippocampal MD change, we performed a mediation analysis across 10000 bootstrapped samples (R package *mediation* v4.5.0)[56, 57]. To test for the relationship between sleep and memory change in the Lifebrain consortium data, partial correlations between sleep and memory change were calculated for each sample, correcting for age, sex, and interval between memory tests. We submitted the resulting correlations and corresponding sample sizes to a meta-analysis (R package *meta* v4.9-8). To illustrate the individual data points, and to provide a general measure of effect size, we extracted hippocampal MD SPC values and the PSQI measure of interest, removed the effects of the nuisance regressors, and plotted the resulting residuals, and reported their R^2. For the analyses including

PGSs, the first 3 principal components of the genetic ancestry factors were included as covariates to correct for population substructures. To account for potential influences of depression and cognitive impairment, we undertook sensitivity analyses by excluding the 11 participants with a depression score exceeding either 16 (BDI) or 9 (GDS), and the 2 participants with an MMS score of 25, and assessed the similarities with the results based on the full sample.

Data Availability Policy

The data supporting the results of the current study are available from the corresponding author on reasonable request, given appropriate ethical and data protection approvals. Requests for data included in the Lifebrain meta-analysis can be submitted to the relevant principal investigators of each study. Contact information can be obtained from the corresponding author.

Results

Sleep and age. Summary measures of the PSQI variables can be found in **Table** 1, together with the correlations between PSQI variables, and between PSQI variables and age. The overall global score, duration, efficiency, disturbance, and daytime dysfunction, but not quality and latency, showed significant relationships with age.

Microstructural Hippocampal Change and Memory change. As shown in **Figure 2A**, hippocampal microstructural change related to memory change ($p=0.0032$, $R^2=0.035$) after accounting for covariates. As hypothesized, higher hippocampal MD change values, interpreted as reduced structural integrity[58], related to more memory decline. As seen in **Figure e-3**, excluding potential outliers (defined as a score

deviating from the median by >3 interquartile ranges) yielded a numerically weaker effect ($p=0.052$, $R^2=0.016$), which was stronger ($p= 0.038$) when allowing older adults to have a different slope by accounting for the interaction between memory and age, although the interaction term itself showed no effect ($p= 0.372$).

[Insert **Figure 2** about here]

Sleep and Microstructural Hippocampal Change. We found a relationship between hippocampal MD change and 2 of the 7 self-reported sleep measures, namely the global PSQI score (FDR-corrected p (p_{FDR})<0.05, $R^2=0.030$, uncorrected p (p_{uncorr})=0.007), and sleep efficiency ($p_{FDR}<0.05$, $R^2=0.025$, $p_{uncorr}=0.014$). As hypothesized, the relationships, shown in **Figure e-4** and **Figure 2B**, respectively, revealed that participants with poorer sleep (higher scores) showed more increase in hippocampal MD. The relationship did not differ across the age range (for the interaction term sleep × age, the lowest uncorrected p-value was 0.37). For comparison with previous studies, we performed 2 additional analyses: (i) we re-ran the same analysis with hippocampal volume change (as opposed to hippocampal MD change in the main analyses) as the dependent variable. This analysis showed no relationships, but the lowest p-value was seen for sleep efficiency ($p_{FDR}=0.471$, $p_{uncorr}=0.067$). (ii) We tested whether any of the sleep variables related to microstructural properties of hippocampus at baseline only. This analysis did not reveal any relationships (lowest $p_{uncorr} = 0.214$). Together, these analyses showed that global sleep quality and sleep efficiency related to microstructural hippocampal change, independently of hippocampal volume and hippocampal volume change. As

the sleep efficiency measure conveys more specific information regarding sleep than the global sleep score, we selected this measure for further analyses.

Sleep Efficiency and Memory change. As shown in **Figure 2C**, poor sleep efficiency was not strongly related to more memory decline (partial r=-0.11, p=0.073, R^2=0.013). To test if this result accurately reflected the true relationship, we performed a meta-analysis of partial correlations obtained from sleep efficiency-memory change testing in 5 samples from the Lifebrain consortium (see Methods for details). This analysis yielded a correlation of -0.078 (95% confidence intervals (CI) [-0.13, -0.02]), Z=-2.70, p=0.007). The partial correlations obtained in the main sample was within this confidence interval, suggesting that a relationship between sleep efficiency and memory change may exist, but that it is of modest strength, and that a larger sample is needed to detect it. For comparison with previous studies, we tested in the main sample whether sleep efficiency related to memory at baseline. No such relation was found (partial r=-0.013, p=0.84).

Sleep Efficiency, Microstructural Hippocampal Change, and Memory Change. To test whether hippocampal MD change mediated the relationships between sleep efficiency and memory change, we performed two additional analyses. First, we tested the relationship between hippocampal MD change and memory change, including sleep efficiency as a covariate. The relation was moderate (p=0.016). Second, we ran a mediation analysis (**Figure 2D**). The unstandardized indirect effect on memory change from sleep efficiency via hippocampal MD change was 0.42 × -0.84 = -0.35. The median bootstrapped unstandardized indirect effect was also -0.35 (p= 0.028, 95% CI [-0.83, -0.02], ρ at which the effect equals 0 was -0.16). The

median direct effect estimate (from sleep efficiency to memory change controlling hippocampal MD change) was -1.26 (p= 0.179). This result suggested that hippocampal MD change partly mediated the relationship between sleep efficiency and memory change.

Sleep Efficiency, Hippocampal Change, and Genetic Effects. A subsample of 180 participants had genotype data available, including *APOE* ε4 status, and two types of PGSs (calculated without variants in the *APOE* region). **Figure e-5A** shows the PGS for sleep efficiency at each level of the PSQI-derived sleep efficiency. This analysis showed that worse self-reported efficiency did not relate to lower genetic propensity for efficient sleep (partial r=-0.037, p=0.62). As shown in **Figure e-5B**, lower genetic propensity for efficient sleep related more strongly, but still very modestly, to hippocampal MD change (partial r=-0.13, p=0.086). Self-reported sleep efficiency still related to hippocampal MD change (p=0.035) when controlling for the sleep efficiency PGS, which showed a negligible unique effect on MD change (p=0.107). A total of 70 participants carried one or two *APOE* ε4 alleles. *APOE* ε4 status was not related to better sleep efficiency (r=-0.050, p=0.507), or hippocampal MD change (r=0.080, p=0.293), and when adding *APOE* ε4 status to the model together with the sleep efficiency PGS, PSQI sleep efficiency still related to hippocampal MD change (p=0.031). **Figure e-6A** shows the AD PGS at each level of the PSQI sleep efficiency, and **Figure e-6B** as a function of hippocampal MD change. Higher genetic risk for AD was not related to worse sleep efficiency, that is, higher PSQI scores (partial correlation r=0.023, p=0.761), or lower hippocampal MD change (partial r=-0.059, p=0.445). When controlling for the AD PGS, sleep efficiency was still related to hippocampal MD change (p=0.019), with no effect of the AD polygenic score

(p=0.147). Also, when adding *APOE* ε4 carrier status to this model including AD PGS, sleep efficiency still related to hippocampal MD change (p=0.017).

Sleep Efficiency, Hippocampal Change, and Aβ. In the subsample of 108 participants with PET data, 23 participants were classified as Aβ positive (see **Figure e-2** for details). As shown in **Figure 3A**, we found support for our hypothesis of a stronger relationship between sleep efficiency and hippocampal MD change in participants classified as Aβ positive (efficiency × Aβ interaction term p=0.021). The Aβ positive participants did not show different sleep efficiency (p=0.722), hippocampal MD decline (p=0.932), or memory decline (p=0.680). When repeating the analysis in the Aβ positive and negative groups separately, we observed a relationship between sleep efficiency and hippocampal MD change only in the Aβ positive (p=0.014), but not in the Aβ negative subgroup (n=85, p=0.414).

[Insert **Figure 3** about here]

Sleep Efficiency, Hippocampal Change, Aβ, and Genetic Effects. A subsample of 76 participants (mean (SD) age=69.3 (8.2) years, min-max 44-81 years) had both Aβ and genotype data. In this subset 24 participants, or 32%, had one or two *APOE* ε4 alleles. First, we included the *APOE* ε4 status and PGSs for sleep efficiency to the initial model. The results demonstrated that sleep efficiency still related to hippocampal MD change differently for Aβ negative and positive participants (t=2.2, p=0.035). Here, we observed an effect of the PGS for sleep efficiency on hippocampal MD change (t=-2.3, p=0.027), with higher propensity of efficient sleep showing less MD hippocampal decline. *APOE* genotype showed no effect (t=0.4,

p=0.678). Second, we included *APOE* ε4 status and the AD PGS to the initial model, which did not alter the interaction of between sleep efficiency and Aβ (t=2.5, p=0.015). The AD PGS showed a very weak effect (t=-1.8, 0.083), while A*POE* ε4 genotype showed no effect (t=0.7, p=0.487). Finally, we included all genetic information, that is, *APOE* ε4 status, and the PGSs for sleep efficiency and AD, respectively, in one model. As shown in **Figure 3B**, this analysis still revealed an interaction between sleep efficiency and Aβ on hippocampal MD change (t=2.1, p=0.015). There was again an effect of the sleep efficiency PGS (t=-2.3, p=0.028), but not of the AD PGS (t=-1.7, p=0.087), or *APOE* genotype (t=0.6, p=0.537).

Sensitivity analyses. To rule out possible influence of extraneous variables on our results, particularly depression, we verified that, when excluding the 11 participants with high depression scores and the two with low MMS scores, the results remained highly similar. That is, sleep efficiency related to hippocampal MD change (p=0.012), and weakly to memory change (again partial r = -0.11), while the relationship between sleep efficiency and hippocampal MD change still differed depending on cortical Aβ accumulation (p=0.012).

Discussion

The results indicate that sleep efficiency and hippocampal microstructural decline are related, particularly in presence of cortical Aβ accumulation. This relationship does not appear to be explained by *APOE* genotype, or polygenic scores for sleep efficiency or AD. Sleep efficiency also related to memory reduction indirectly via hippocampal integrity decline. Although we cannot rule out that these Aβ-related correlations stem from unexplored factors such as tau deposition in the medial

temporal lobes, one possibility is that Aβ accumulation constitutes a vulnerability in the case of reduced hippocampal integrity, leading to both lower sleep efficiency and decline in episodic memory.

As the hippocampal effects observed here were independent of baseline hippocampal volume and volume change, microstructural change in the hippocampus might be a particularly sensitive marker of early decline, complementary to atrophy. In support of this hypothesis, while one study has reported cross-sectional associations in 1201 adults (mean age 21 years) between the right hippocampal MD and sleep quality (but not sleep duration), two previous studies of 147 (overlapping with the current sample)[19] and 66[59] participants, respectively, did not observe relationships between sleep and hippocampal volume or atrophy. These findings also suggest larger samples may be needed to detect the sleep-atrophy relationships. In support of this notion, we showed in 3105 cognitively normal participants aged 18-90 years from the Lifebrain consortium, including participants from the present sample, that poorer sleep efficiency, as well as sleep quality, problems, and daytime tiredness, were related to greater hippocampal volume loss[8]. The current finding supports such a relationship between self-reported sleep and hippocampal change but extends previous knowledge by revealing independent intra-hippocampal reductions in microstructural integrity longitudinally.

The underlying neurobiology of the microstructural hippocampal decline remains unclear, but may be due to decay of the dendritic architecture. In mice, hippocampal dendritic spine densities have been shown to be reduced both in aging[60], and after sleep deprivation[61]. Hippocampal dendritic decay might also underlie the relations

observed here between microstructural hippocampus decline and memory reductions. In mice, hippocampal dendritic spine loss has been related with memory defects[62]. Over time, loss of spines and synapses might promote larger dendritic disruptions, which, in mice, has been detected via intra-hippocampal DTI, and linked with memory impairments[63]. Whether whole-hippocampal MD in humans also is influenced by such dendritic effects remain unknown, but these speculations could be tested using ultrahigh-resolution microstructural DTI[64].

The relationship between sleep and Aβ appears reciprocal, as Aβ increase after sleep deprivation[65], while Aβ accumulation in turn increases wakefulness and alters sleep patterns[66]. Here, although Aβ status did not relate to sleep efficiency or hippocampus decline alone, the sleep-hippocampal decline relationship was significantly stronger in the Aβ positive participants. Echoing these result, we recently observed in a separate sample of older adults that both tau and YKL-40, a biomarker of inflammation and astroglial activation, related more strongly to the PSQI global score in Aβ positive than in Aβ negative participants[67]. These results raise the possibility that Aβ accumulation co-occurring with other adverse signs such as hippocampal decline or inflammation, signals sleep problems not observed with Aβ accumulation alone.

As we observed that the effect of sleep efficiency on memory decline was mediated by higher hippocampal diffusivity, we hypothesize that hippocampal decline, when concomitant with Aβ accumulation, causes sleep problems, here in the form of poorer sleep efficiency. One potential mechanism might be that the hippocampal decline causes altered brain oscillations affecting sleep[68]. However, the current data does

not allow inferences that rule out the reverse causality, of sleep affecting hippocampal decline. That said, we find such a reverse pattern less likely as effects were specific for sleep efficiency, rather than sleep duration or quality which would be more likely drivers of possible sleep-generated causal effects. Similarly, we cannot rule out that a variable not assessed here can account for the observed correlations[57]. For instance, sleep spindles have been linked, in addition to Aβ, with tau[4], and potentially AD-related tau is seen first in the locus coeruleus[69]. Activity in this region can alter sleep spindles, affecting memory consolidation[70]. To tease out potential causal pathways, studies could include Aβ-negative participants with healthy sleep patterns and no signs of hippocampal decline, and follow them over several years to assess changes in sleep patterns, hippocampal integrity, Aβ status, and memory performance, as well as tau, and measures of neuroinflammation such as YKL-40 or sTREM2. Intervention studies targeting for instance hippocampal-dependent cognition[71], and investigating similar markers could be a less costly, but more practical, future strategy.

The relationships remained after controlling for polygenetic scores for sleep efficiency and AD and the presence of the APOE ε4 allele. APOE ε4 is a well-known risk factor for AD[72], and studies have reported that healthy APOE ε4 carriers show more pronounced longitudinal hippocampal atrophy[73, 74] and worse sleep quality[75]. In contrast, there are also reports of no relationship between hippocampal atrophy in cognitively healthy adults and AD genes from an exploratory GWAS[76], and of lack of APOE effects on hippocampal volume[77]. Both for the *APOE* genotype and the PGSs, we observed relatively weak relationships with sleep efficiency and hippocampal change. This low correspondence must be further investigated and resolved before

we can draw the conclusion that the relationship between sleep efficiency and hippocampal decline is partly independent of genetics.

Limitations of this study concern the use of a self-report measure of sleep, at one time point, instead of objective measures such as activity monitors, or polysomnography, collected repeatedly. Although self-reported sleep measures might provide more representative data on sleep than a single-night polysomnography[78], a correlation of 0.47 has been reported between reported and measured sleep duration[79]. The modest nature of this correlation highlights the need for objective measures to assess sleep physiology. In future studies, a likely key is repeated measurements to get a more detailed picture of sleep patterns, as well as the inclusion of other biomarkers such as tau, and markers of neuroinflammation. Inclusion of such markers would also improve analyses of mediation, which here does not establish causality. As activity monitors were used in the GWAS from which the sleep efficiency PGS stems[34], the inclusion of objective sleep measures will also likely shed further light on the relative contribution of sleep genetics and sleep behavior. Although the current sample is relatively large, the potentially complex interplay between Aβ positivity and other biomarkers of relatively low prevalence highlight the need for even larger sample sizes.

Conclusion Section

The results indicate that hippocampal microstructural decline related to sleep efficiency, and episodic memory change in cognitively normal older adults, particularly in Aβ positive participants. This relationship was not readily explained by genetic effects. Poor self-reported sleep efficiency might constitute a risk factor for

AD, and future studies need to address why sleep is related to more hippocampal decline in Aβ positive older adults even without dementia.

Appendix 1: Authors

Name	Location	Contribution
Håkon Grydeland, PhD	University of Oslo, Norway	Design and conceptualized study; analyzed and interpreted the data; drafted the manuscript for intellectual content
Donatas Sederevičius, MSc	University of Oslo, Norway	Analyzed the data; revised the manuscript for intellectual content
Yunpeng Wang, PhD	University of Oslo, Norway	Analyzed the genotype data; revised the manuscript for intellectual content
David Bartrés-Faz, PhD	Universitat de Barcelona, Spain	Analyzed consortium data; revised the manuscript for intellectual content
Lars Bertram, MD, PhD	Max Planck Institute for Molecular Genetics, Germany	Analyzed genotype data; revised the

		manuscript for intellectual content
Valerija Dobricic, PhD	University of Lübeck, Lübeck, Germany	Analyzed the genotype data; revised the manuscript for intellectual content
Sandra Düzel, PhD	Max Planck Institute for Human Development, Germany	Analyzed consortium data; revised the manuscript for intellectual content
Klaus P. Ebmeier, MD	University of Oxford, UK	Analyzed consortium data; revised the manuscript for intellectual content
Ulman Lindenberger, PhD	Max Planck Institute for Human Development, Germany	Analyzed consortium data; revised the manuscript for intellectual content
Lars Nyberg, PhD	Umeå University, Sweden	Analyzed consortium data; revised the manuscript for intellectual content
Sara Pudas, PhD	Umeå University, Sweden	Analyzed consortium data; revised the

		manuscript for intellectual content
Claire E. Sexton, PhD	University of Oxford, UK	Analyzed consortium data; revised the manuscript for intellectual content
Cristina Solé-Padullés, PhD	Universitat de Barcelona, Spain	Analyzed consortium data; revised the manuscript for intellectual content
Kristine B. Walhovd, PhD	University of Oslo, Norway	Design and conceptualized study; interpreted the data; revised the manuscript for intellectual content
Anders M. Fjell, PhD	University of Oslo, Norway	Design and conceptualized study; analyzed and interpreted the data; revised the manuscript for intellectual content

References

1. Yaffe K, Nettiksimmons J, Yesavage J, Byers A. Sleep Quality and Risk of Dementia Among Older Male Veterans. Am J Geriatr Psychiatry 2015;23:651-654.
2. Potvin O, Lorrain D, Forget H, et al. Sleep quality and 1-year incident cognitive impairment in community-dwelling older adults. Sleep 2012;35:491-499.
3. Ju YE, McLeland JS, Toedebusch CD, et al. Sleep quality and preclinical Alzheimer disease. JAMA Neurol 2013;70:587-593.
4. Winer JR, Mander BA, Helfrich RF, et al. Sleep as a Potential Biomarker of Tau and beta-Amyloid Burden in the Human Brain. J Neurosci 2019;39:6315-6324.
5. Bilgel M, An Y, Helphrey J, et al. Effects of amyloid pathology and neurodegeneration on cognitive change in cognitively normal adults. Brain 2018;141:2475-2485.
6. Mander BA, Marks SM, Vogel JW, et al. beta-amyloid disrupts human NREM slow waves and related hippocampus-dependent memory consolidation. Nat Neurosci 2015;18:1051-1057.
7. Wilckens KA, Tudorascu DL, Snitz BE, et al. Sleep moderates the relationship between amyloid beta and memory recall. Neurobiol Aging 2018;71:142-148.
8. Fjell AM, Sorensen O, Amlien IK, et al. Self-reported sleep relates to hippocampal atrophy across the adult lifespan - results from the Lifebrain consortium. Sleep 2019.
9. Le Bihan D. Looking into the functional architecture of the brain with diffusion MRI. Nat Neurosci 2003;4:469-480.
10. van Norden AG, de Laat KF, Fick I, et al. Diffusion tensor imaging of the hippocampus and verbal memory performance: the RUN DMC study. Hum Brain Mapp 2012;33:542-551.
11. den Heijer T, der Lijn F, Vernooij MW, et al. Structural and diffusion MRI measures of the hippocampus and memory performance. Neuroimage 2012;63:1782-1789.
12. Corder EH, Saunders AM, Strittmatter WJ, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 1993;261:921-923.
13. Veatch OJ, Keenan BT, Gehrman PR, Malow BA, Pack AI. Pleiotropic genetic effects influencing sleep and neurological disorders. Lancet Neurol 2017;16:158-170.
14. Dudbridge F. Power and predictive accuracy of polygenic risk scores. PLoS Genet 2013;9:e1003348.
15. Walhovd KB, Fjell AM, Westerhausen R, et al. Healthy minds 0-100 years: Optimising the use of European brain imaging cohorts ("Lifebrain"). Eur Psychiatry 2018;50:47-56.
16. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189-198.
17. Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res 1982;17:37-49.
18. Wechsler D. Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: The Psychological Corporation, 1999.
19. Sexton CE, Storsve AB, Walhovd KB, Johansen-Berg H, Fjell AM. Poor sleep quality is associated with increased cortical atrophy in community-dwelling adults. Neurology 2014;83:967-973.
20. Buysse DJ, Reynolds CF, 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989;28:193-213.

21. Knutson KL, Rathouz PJ, Yan LL, Liu K, Lauderdale DS. Stability of the Pittsburgh Sleep Quality Index and the Epworth Sleepiness Questionnaires over 1 year in early middle-aged adults: the CARDIA study. Sleep 2006;29:1503-1506.
22. Zitser J, Anaturk M, Zsoldos E, et al. Sleep duration over 28 years, cognition, gray matter volume, and white matter microstructure: a prospective cohort study. Sleep 2020.
23. Andersson JLR, Graham MS, Zsoldos E, Sotiropoulos SN. Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. Neuroimage 2016;141:556-572.
24. Andersson JLR, Sotiropoulos SN. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. Neuroimage 2016;125:1063-1078.
25. Westlye LT, Walhovd KB, Dale AM, et al. Life-span changes of the human brain white matter: diffusion tensor imaging (DTI) and volumetry. Cereb Cortex 2010;20:2055-2068.
26. Fischl B, Salat DH, Busa E, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 2002;33:341-355.
27. Greve DN, Fischl B. Accurate and robust brain image alignment using boundary-based registration. Neuroimage 2009;48:63-72.
28. Delis DC, Kramer JH, Kaplan E, Ober BA. California Verbal Learning Test-Second Edition (CVLT-II). San Antonio, TX: Psychological Corporation, 2000.
29. Berry DA, Ayers GD. Symmetrized percent change for treatment comparisons. Am Stat 2006;60:27-31.
30. Wolk DA, Grachev ID, Buckley C, et al. Association between in vivo fluorine 18-labeled flutemetamol amyloid positron emission tomography imaging and in vivo cerebral cortical histopathology. Arch Neurol 2011;68:1398-1403.
31. Ross S, Stearns C. SharpIR White paper2010.
32. Walhovd KB, Fjell AM, Sørensen Ø, et al. Genetic risk for Alzheimer’s disease predicts hippocampal volume through the lifespan. bioRxiv 2019.
33. Jansen IE, Savage JE, Watanabe K, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. Nat Genet 2019;51:404-413.
34. Jones SE, van Hees VT, Mazzotti DR, et al. Genetic studies of accelerometer-based sleep measures yield new insights into human sleep behaviour. Nat Commun 2019;10:1585.
35. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559-575.
36. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic variation. Nature 2015;526:68-74.
37. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS Genet 2006;2:e190.
38. Laws SM, Hone E, Gandy S, Martins RN. Expanding the association between the APOE gene and the risk of Alzheimer's disease: possible roles for APOE promoter polymorphisms and alterations in APOE transcription. J Neurochem 2003;84:1215-1236.
39. Radmanesh F, Devan WJ, Anderson CD, Rosand J, Falcone GJ, Adni. Accuracy of imputation to infer unobserved APOE epsilon alleles in genorne-wide genotyping data. Eur J Hum Genet 2014;22:1239-1242.
40. Greve DN, Salat DH, Bowen SL, et al. Different partial volume correction methods lead to different conclusions: An (18)F-FDG-PET study of aging. Neuroimage 2016;132:334-343.

41. Greve DN, Svarer C, Fisher PM, et al. Cortical surface-based analysis reduces bias and variance in kinetic modeling of brain PET data. Neuroimage 2014;92:225-236.
42. Desikan RS, Segonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 2006;31:968-980.
43. Thal DR, Rub U, Orantes M, Braak H. Phases of A beta-deposition in the human brain and its relevance for the development of AD. Neurology 2002;58:1791-1800.
44. Mormino EC, Betensky RA, Hedden T, et al. Amyloid and APOE epsilon4 interact to influence short-term decline in preclinical Alzheimer disease. Neurology 2014;82:1760-1767.
45. Bertram L, Bockenhoff A, Demuth I, et al. Cohort profile: The Berlin Aging Study II (BASE-II). Int J Epidemiol 2014;43:703-712.
46. Gerstorf D, Bertram L, Lindenberger U, et al. Editorial. Gerontology 2016;62:311-315.
47. Nilsson LG, Backman L, Erngrund K, et al. The Betula prospective cohort study: Memory, health, and aging. Aging, Neuropsychology and Cognition 1997;4:1-32.
48. Abellaneda-Perez K, Vaque-Alcazar L, Vidal-Pineiro D, et al. Age-related differences in default-mode network connectivity in response to intermittent theta-burst stimulation and its relationships with maintained cognition and brain integrity in healthy aging. Neuroimage 2019;188:794-806.
49. Rajaram S, Valls-Pedret C, Cofan M, et al. The Walnuts and Healthy Aging Study (WAHA): Protocol for a Nutritional Intervention Trial with Walnuts on Brain Aging. Front Aging Neurosci 2016;8:333.
50. Vidal-Pineiro D, Martin-Trias P, Arenaza-Urquijo EM, et al. Task-dependent activity and connectivity predict episodic memory network-based responses to brain stimulation in healthy aging. Brain Stimul 2014;7:287-296.
51. Filippini N, Zsoldos E, Haapakoski R, et al. Study protocol: The Whitehall II imaging sub-study. BMC Psychiatry 2014;14:159.
52. Singh-Manoux A, Kivimaki M, Glymour MM, et al. Timing of onset of cognitive decline: results from Whitehall II prospective cohort study. BMJ 2012;344:d7622.
53. Yekutieli D, Benjamini Y. Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. J Stat Plan Infer 1999;82:171-196.
54. Roalf DR, Quarmley M, Elliott MA, et al. The impact of quality assurance assessment on diffusion tensor imaging outcomes in a large-scale population-based cohort. Neuroimage 2016;125:903-919.
55. Buckner RL, Head D, Parker J, et al. A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. Neuroimage 2004;23:724-738.
56. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. mediation: R Package for Causal Mediation Analysis. J Stat Softw 2014;59.
57. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. Psychol Methods 2010;15:309-334.
58. Bennett IJ, Madden DJ, Vaidya CJ, Howard DV, Howard JH, Jr. Age-related differences in multiple measures of white matter integrity: A diffusion tensor imaging study of healthy aging. Hum Brain Mapp 2010;31:378-390.
59. Lo JC, Loh KK, Zheng H, Sim SK, Chee MW. Sleep duration and age-related changes in brain structure and cognitive performance. Sleep 2014;37:1171-1178.

60. von Bohlen und Halbach O, Zacher C, Gass P, Unsicker K. Age-related alterations in hippocampal spines and deficiencies in spatial memory in mice. J Neurosci Res 2006;83:525-531.
61. Havekes R, Park AJ, Tudor JC, et al. Sleep deprivation causes memory deficits by negatively impacting neuronal connectivity in hippocampal area CA1. Elife 2016;5.
62. Chen Y, Rex CS, Rice CJ, et al. Correlated memory defects and hippocampal dendritic spine loss after acute stress involve corticotropin-releasing hormone signaling. Proc Natl Acad Sci U S A 2010;107:13123-13128.
63. Molet J, Maras PM, Kinney-Lang E, et al. MRI uncovers disrupted hippocampal microstructure that underlies memory impairments after early-life adversity. Hippocampus 2016;26:1618-1632.
64. Yassa MA, Muftuler LT, Stark CE. Ultrahigh-resolution microstructural diffusion tensor imaging reveals perforant path degradation in aged humans in vivo. Proc Natl Acad Sci U S A 2010;107:12687-12691.
65. Shokri-Kojori E, Wang GJ, Wiers CE, et al. beta-Amyloid accumulation in the human brain after one night of sleep deprivation. Proc Natl Acad Sci U S A 2018;115:4483-4488.
66. Ju YE, Lucey BP, Holtzman DM. Sleep and Alzheimer disease pathology--a bidirectional relationship. Nat Rev Neurol 2014;10:115-119.
67. Fjell AM, Idland AV, Sala-Llonch R, et al. Neuroinflammation and Tau Interact with Amyloid in Predicting Sleep Problems in Aging Independently of Atrophy. Cereb Cortex 2018;28:2775-2785.
68. Buzsaki G. Theta oscillations in the hippocampus. Neuron 2002;33:325-340.
69. Braak H, Thal DR, Ghebremedhin E, Del Tredici K. Stages of the pathologic process in Alzheimer disease: age categories from 1 to 100 years. J Neuropathol Exp Neurol 2011;70:960-969.
70. Swift KM, Gross BA, Frazer MA, et al. Abnormal Locus Coeruleus Sleep Activity Alters Sleep Signatures of Memory Consolidation and Impairs Place Cell Stability and Spatial Memory. Curr Biol 2018;28:3599-3609 e3594.
71. Lovden M, Schaefer S, Noack H, et al. Spatial navigation training protects the hippocampus against age-related changes during early and late adulthood. Neurobiol Aging 2012;33:620 e629-620 e622.
72. Saunders AM, Strittmatter WJ, Schmechel D, et al. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. Neurology 1993;43:1467-1472.
73. Cohen RM, Small C, Lalonde F, Friz J, Sunderland T. Effect of apolipoprotein E genotype on hippocampal volume loss in aging healthy women. Neurology 2001;57:2223-2228.
74. Li B, Shi J, Gutman BA, et al. Influence of APOE Genotype on Hippocampal Atrophy over Time - An N=1925 Surface-Based ADNI Study. PLoS One 2016;11:e0152901.
75. Drogos LL, Gill SJ, Tyndall AV, et al. Evidence of association between sleep quality and APOE epsilon 4 in healthy older adults A pilot study. Neurology 2016;87:1836-1842.
76. Mather KA, Armstrong NJ, Wen W, et al. Investigating the genetics of hippocampal volume in older adults without dementia. PLoS One 2015;10:e0116920.
77. Lyall DM, Royle NA, Harris SE, et al. Alzheimer's disease susceptibility genes APOE and TOMM40, and hippocampal volumes in the Lothian birth cohort 1936. PLoS One 2013;8:e80513.

78. Unruh ML, Redline S, An MW, et al. Subjective and objective sleep quality and aging in the sleep heart health study. J Am Geriatr Soc 2008;56:1218-1227.
79. Lauderdale DS, Knutson KL, Yan LL, Liu K, Rathouz PJ. Self-reported and measured sleep duration: how similar are they? Epidemiology 2008;19:838-845.

Tables

Table 1. Participants demographics

				Correlation	
	M	SD	Range	*PSQIg*	*Age*
Age, baseline (n=251, females=61%)	53.6	19.8	20-81	0.12*	NA
Sleep					
Global	5.0	2.8	0-14	NA	0.12*
Quality	0.8	0.7	0-3	0.74**	0.04
Latency	1.0	0.8	0-3	0.73**	-0.08
Duration	0.6	0.7	0-3	0.60**	0.14*
Efficiency	0.5	0.8	0-3	0.65**	0.21**
Disturbance	1.1	0.5	0-2	0.48**	0.15*
Daytime dysfunction	0.7	0.6	0-2	0.31**	-0.25**
CVLT, 30-min delayed recall (SPC, n=244)	0.08	10.5	-60-39	-0.12	-0.19**
Interval MRI baseline to MRI follow-up	3.1	1.2	1-6	-0.08	-0.45**
Head movement (tSNR, baseline)	5.9	0.6	4-7	-0.03	-0.63**
Head movement (tSNR, baseline vs follow-up)	0.1	0.4	-1-2	0.01	0.15*
Interval PSQI to follow-up MRI (years, n=216[a])	2.5	1	1-5	-0.14[b]	-0.55**
Interval PET scan to follow-up MRI (years, n=108/9 aged 50-80 yeas)	-0.7	1	-2-1	-0.02	NA

Abbreviations: NA=not applicable; PSQI=Pittsburgh Sleep Quality Inventory; PSQIg= PSQI global score; tSNR=temporal signal to noise ratio. SPC=symmetrized percent change. **=p<0.001; *=p<0.05. [a]missing exact date for 14%. [b] P value=0.138 when controlling for age at baseline.

Figures

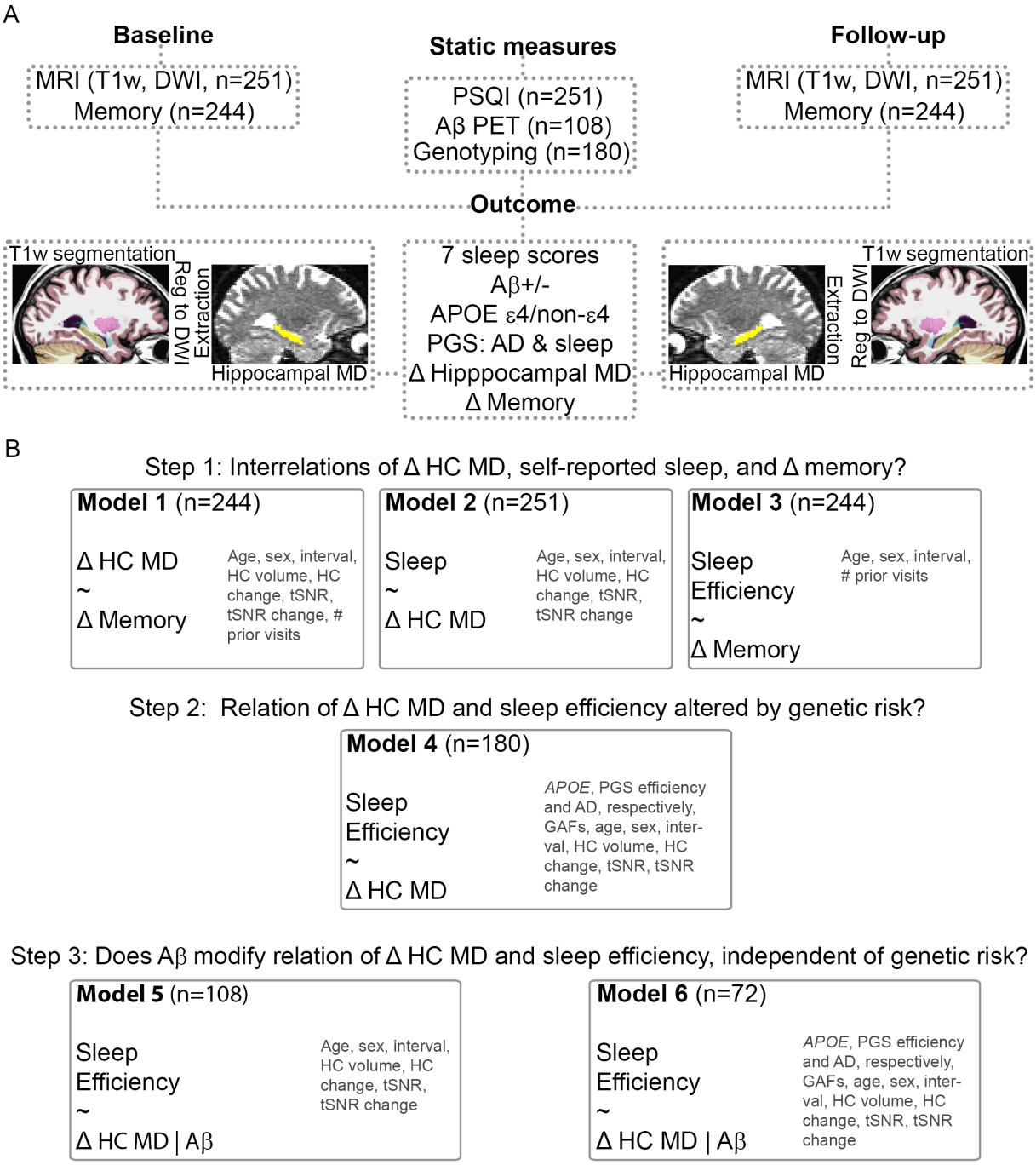

Figure 1. Study overview.

A. Design. **B**. Regression models. Covariates are named in dark gray. Abbreviations: Age=baseline MRI age; HC=hippocampus; HC volume=baseline hippocampal volume; tSNR=temporal signal to noise ratio, derived from DWI scans (see text for details); PSQI=Pittsburgh Sleep Quality Inventory; Aβ=β-amyloid; PGS=polygenic scores; GAF=genetic ancestry factor; # prior visits=number of prior visits in project (see text for details).

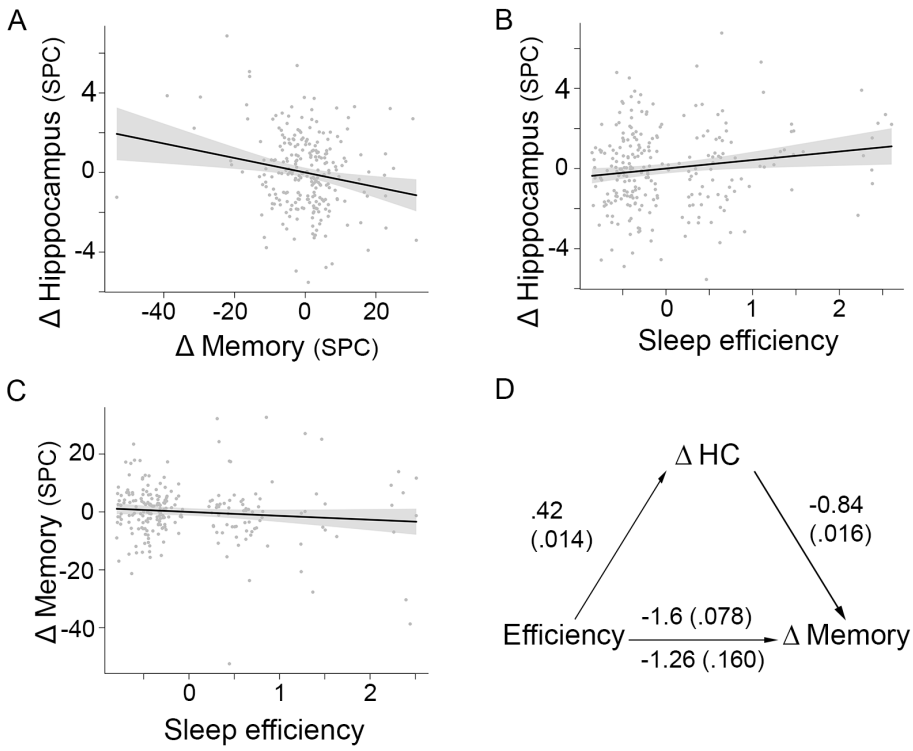

Figure 2. Sleep, and Decline in Microstructural Hippocampal, and Memory.

A. Decline in memory related to MD increase in hippocampus (decline in structural integrity). Values are residualized after regressing out covariates (see **Figure 1B**). **B**. Sleep efficiency related to hippocampal MD change, independently of hippocampal volume and hippocampal volume change, after FDR-correction for multiple comparisons. **C**. Sleep efficiency correlated weakly with memory change (partial r=-0.11, correcting for age at baseline, interval, and sex). **D**. Average causal mediation effect, that is, the indirect effect of sleep on memory via hippocampus, was -0.35 ($p<0.05$).

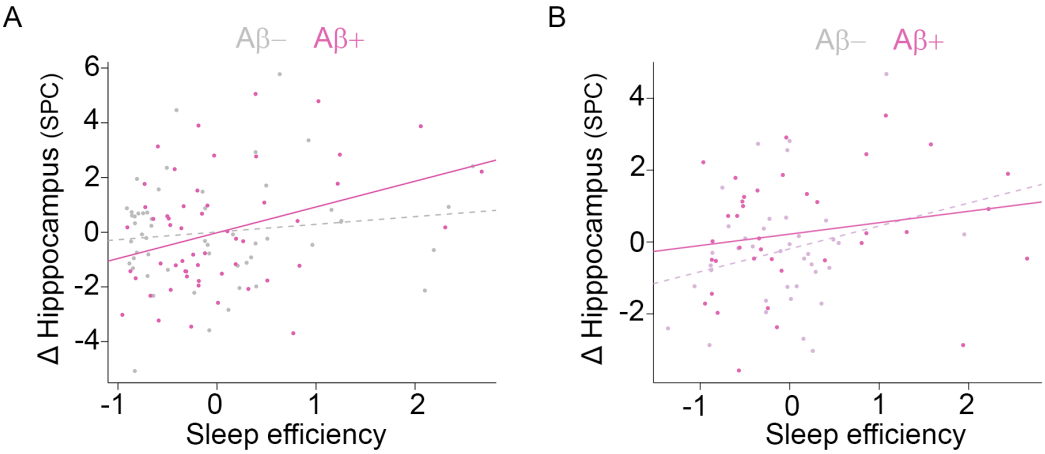

Figure 3. Sleep efficiency, microstructural hippocampal decline, and Aβ accumulation.

A. Efficiency related more strongly to microstructural hippocampal decline in participants with signs of cortical Aβ accumulation. **B**. This relationship remained when controlling for APOE ε4 status and PGS for sleep efficiency and AD, respectively.

Supplementary Figures

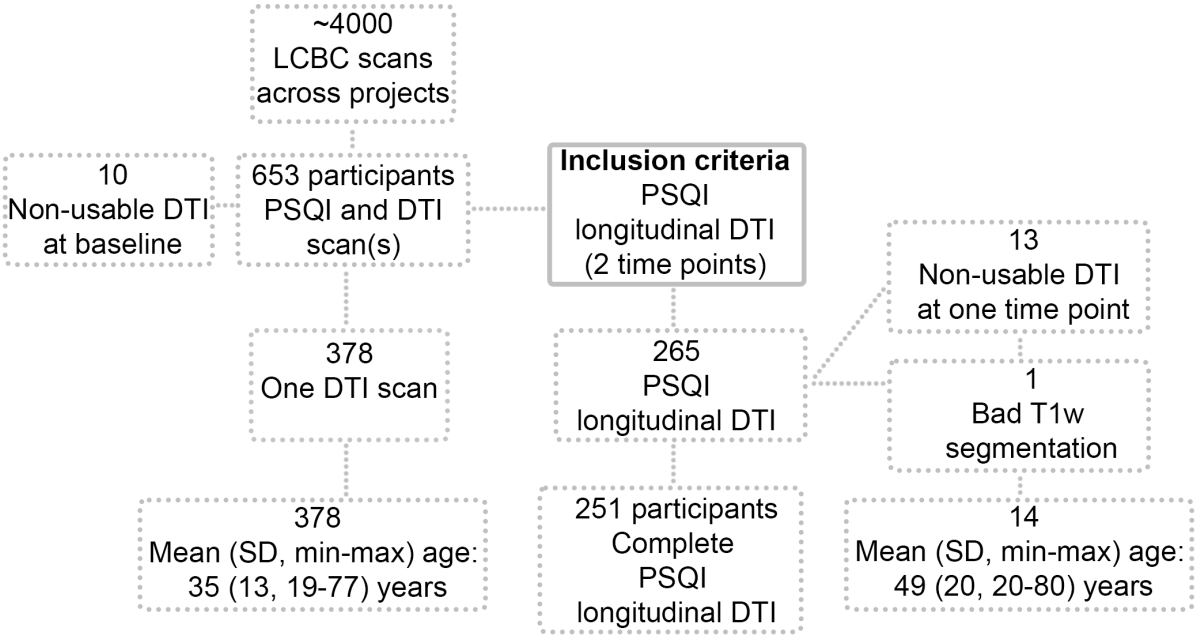

Figure e-1. Attrition.

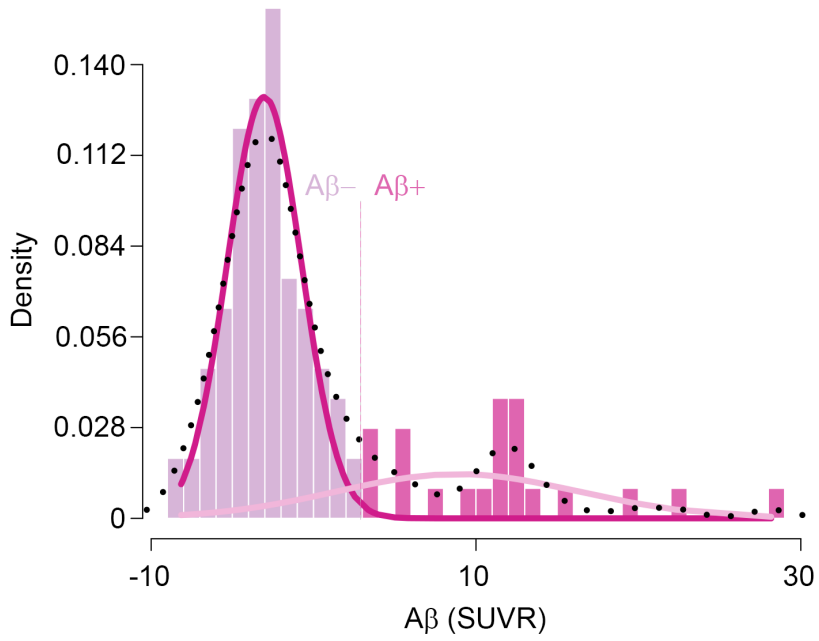

Figure e-2. Classification of Aβ negative and positive.

Distributions of amyloid. The best fit was a two-distribution solution (with unequal variance), represented in different colours. Fit of the distributions overlaid together with actual density (black dotted line). SUVR=standardized uptake value ratios. Aβ negative sample: n=85 (60%F), mean (SD, min-max) age: 67.4 (9.1, 44.4-80.8) years. Aβ positive sample: n=23 (48%F), mean (SD, min-max) age: 70 (6.6, 51.1-78.6) years.

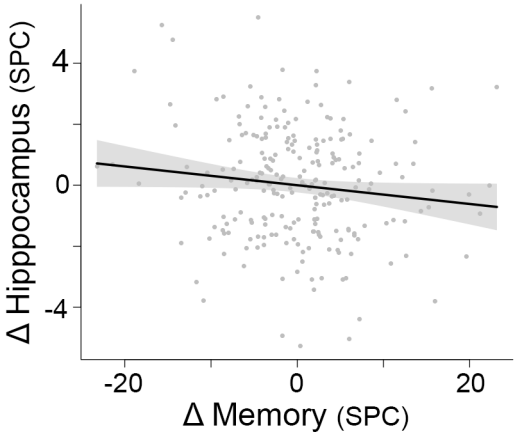

Figure e-3. **Sensitivity analysis**.

Hippocampal MD decline and memory decline when removing extreme (interquartile range*5) memory values.

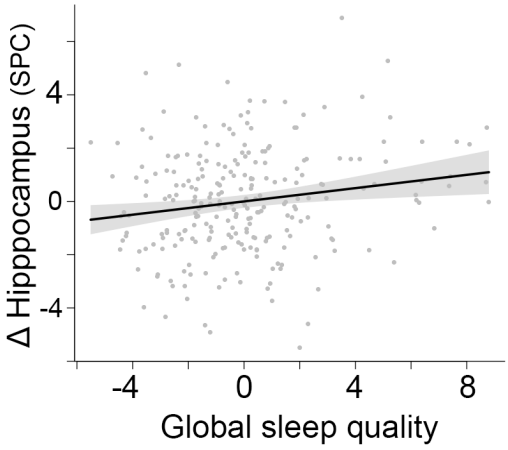

Figure e-4. **Global sleep quality related to hippocampal MD change.**

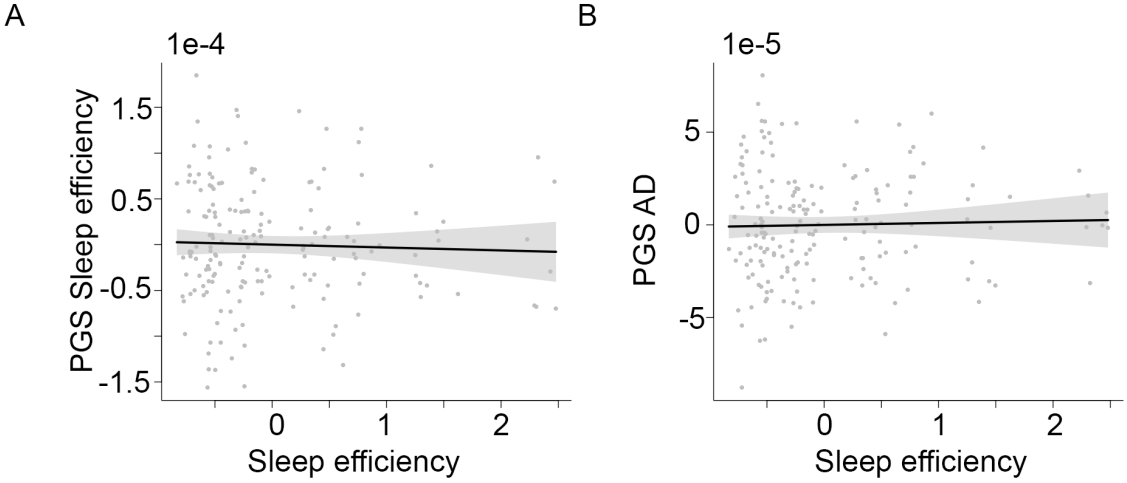

Figure e-5. Self-reported sleep efficiency and polygenic scores (PGSs) for (**A**) sleep efficiency, and (**B**) AD.

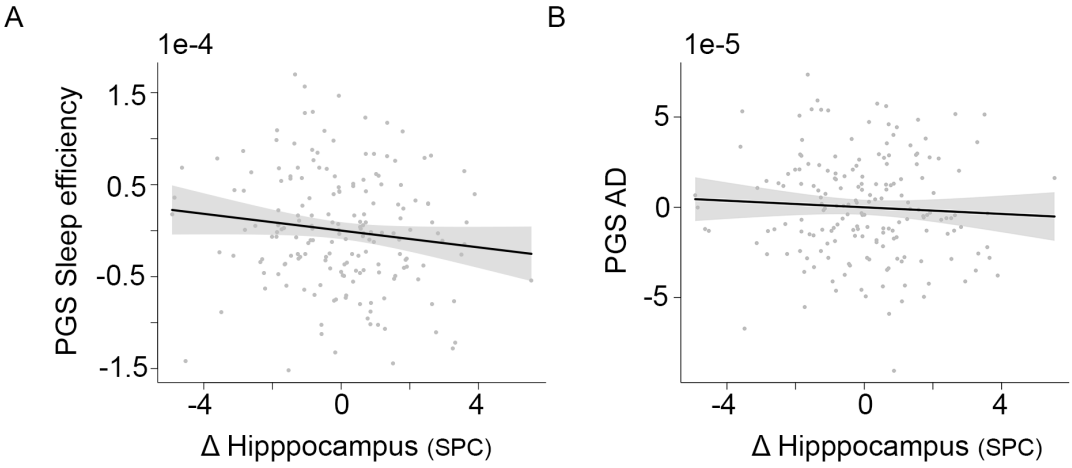

Figure e-6. Hippocampal MD change and polygenic scores (PGSs) for (**A**) sleep efficiency, and (**B**) AD.