Supplementary Results

Results S1. Feature interaction and subgroup analysis based on model explanation

Subgroups were derived by clustering participant-level SHAP values from the AutoGluon models trained with sleep and demographic features in the Greifswald cohort. The resulting clusters therefore reflected similarity in model explanation patterns rather than similarity in the raw input variables. To make these clusters interpretable, SkopeRules was used to derive rule-based descriptions from the original input features.

For Stroop test score prediction, four SHAP-derived subgroups were identified. The rule-based descriptions showed that the subgroup structure was almost entirely organized by age and sex. Cluster 0 included older females, defined by age > 58.1 years and female sex, with 167 participants. Cluster 1 included younger or middle-aged males, defined by age $\leqslant$ 57.6 years and male sex, with 289 participants. Cluster 2 included older males, defined by age > 57.6 years and male sex, with 154 participants. Cluster 3 included younger or middle-aged females, defined by 25.8 < age $\leqslant$ 57.7 years and female sex, with 235 participants (Supplementary table 5). The SHAP-IQ interaction networks indicated that age was the dominant feature across the four Stroop subgroups, with sex representing the main secondary modifier. In older males, age, sex, and the age-by-sex interaction all contributed toward higher predicted Stroop test scores, consistent with longer reaction time. In older females, age also contributed toward higher predicted Stroop test scores, but the age-by-sex interaction contributed in the opposite direction, indicating that it partly offset the age-related prediction pattern. In the younger and middle-aged subgroups, age generally contributed toward lower predicted Stroop test scores, but the direction of the age-by-sex interaction differed between males and females. Across Stroop subgroups, sleep variables contributed less prominently than age and sex and appeared mainly in age-centered interactions, suggesting that sleep measures added context to demographic prediction patterns rather than defining the subgroup structure (Fig 6a, Supplementary Figure 9).

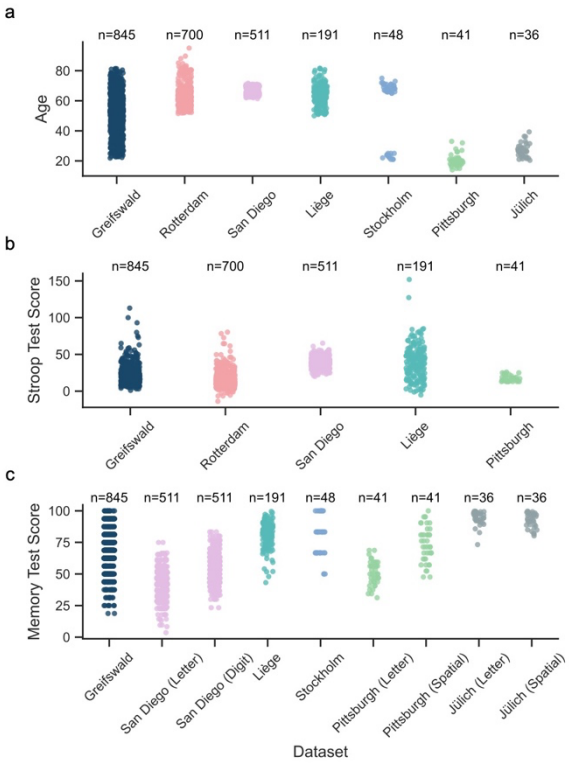
For memory test score prediction, six SHAP-derived subgroups were identified. Compared with Stroop prediction, the rule-based summaries were more differentiated and included PSG-based sleep efficiency, PSG-based sleep duration, and depressive score in addition to age and sex. Cluster 0 included middle-aged to older females, defined by 47.9 < age $\leqslant$ 67.5 years and female sex, with 240 participants. Cluster 1 included middle-aged to older males, defined by 47.9 < age $\leqslant$ 67.4 years and male sex, with 213 participants. Cluster 2 included younger males with higher PSG-based sleep efficiency and lower depressive score, defined by PSG-based sleep efficiency > 0.608, depressive score $\leqslant$ 39.5, age $\leqslant$ 47.9 years, and male sex, with 168 participants. Cluster 3 included a similar but more selective younger male subgroup, defined by PSG-based sleep efficiency > 0.677, depressive score $\leqslant$ 20.0, age $\leqslant$ 47.8 years, and male sex, with 117 participants. Cluster 4 included older males, defined by PSG-based sleep duration > 3.27 hours, age > 67.4 years, and male sex, with 58 participants. Cluster 5 included older females with longer PSG-based sleep duration and lower depressive score, defined by PSG-based sleep duration > 5.15 hours, depressive score $\leqslant$ 11.9, age > 67.4 years, and female sex, with 49 participants (Supplementary Table 6). The SHAP-IQ interaction networks for memory prediction showed that age remained the dominant organizing feature, but its contribution differed more strongly across subgroups than in Stroop prediction. In the middle-aged to older subgroups, female sex was associated with higher predicted memory scores than male sex

within a similar age range. In the younger male subgroups, higher PSG-based sleep efficiency and lower depressive score characterized the rules, and PSG-based sleep efficiency interacted with age in the direction of higher predicted memory scores. The rule-based summaries for Clusters 2 and 3 overlapped substantially, indicating that these rules captured the dominant descriptive profile of each cluster rather than exact mutually exclusive boundaries. In the oldest male subgroup, age and sex contributed toward lower predicted memory scores, whereas PSG-based sleep duration and efficiency showed positive age-centered interaction contributions. In the oldest female subgroup, age contributed toward lower predicted memory scores, but female sex and the age-by-sex interaction contributed in the opposite direction. Across memory subgroups, PSG-based and self-reported sleep measures showed different attribution patterns, suggesting that they contributed differently to model predictions (Fig 7a, Supplementary Figure 10).

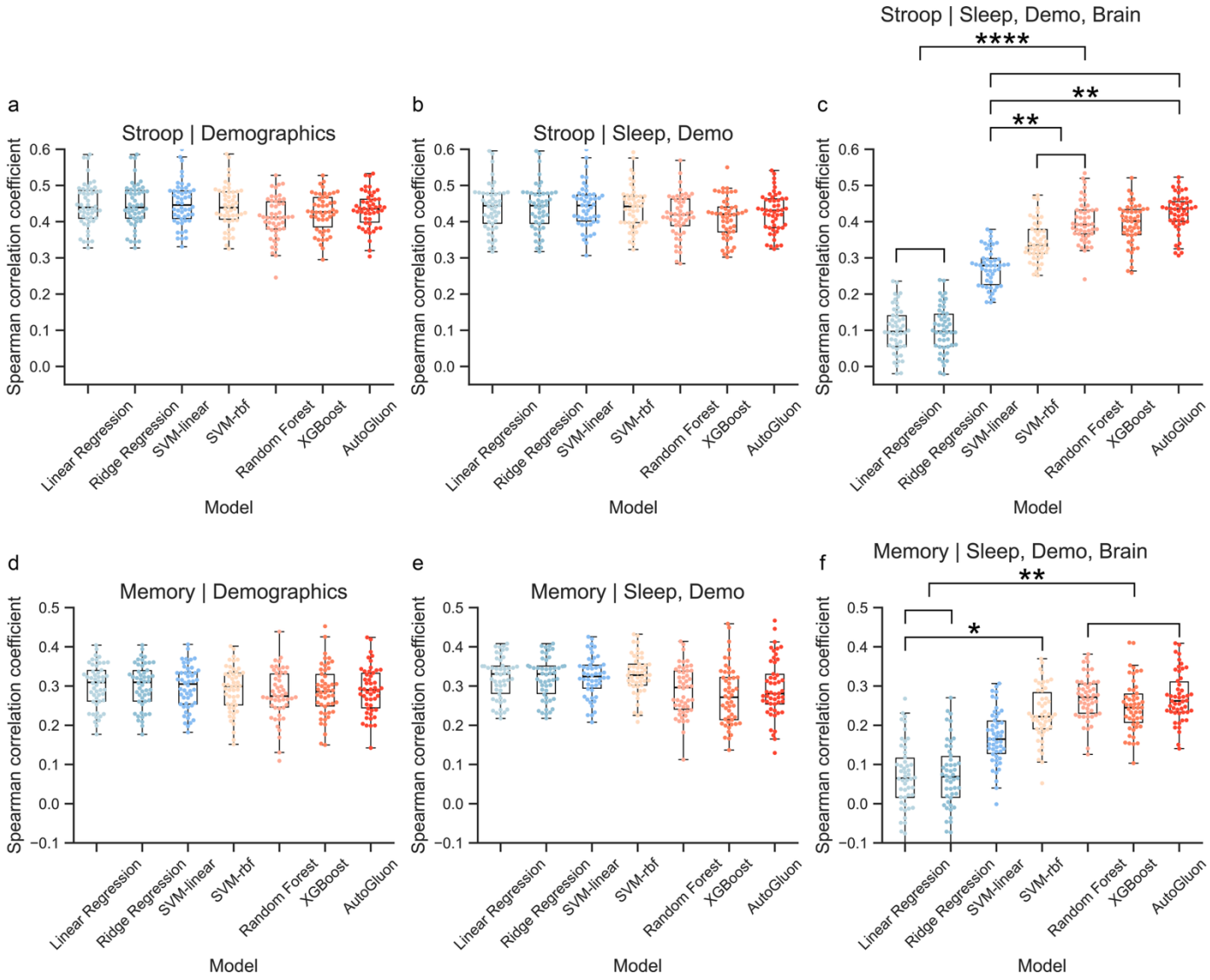
When the Greifswald-derived rules were applied to external validation cohorts, subgroup recovery was cohort-specific. For Stroop prediction, two Greifswald-derived subgroup profiles were recovered in Liège and one was recovered in San Diego (Figs 6b-c). For memory prediction, three Greifswald-derived subgroup profiles were recovered in Stockholm (Fig 7b). The corresponding SHAP and SHAP-IQ profiles showed broadly similar interaction patterns to the matched Greifswald subgroups, although not all Greifswald-derived profiles were represented in the external cohorts. This limited transferability likely reflects cohort differences in age distributions, sex composition, sleep measurements, and cognitive measurements.

Overall, the supplementary subgroup analysis supports the main finding that age was the dominant contributor to the model predictions. Sex modified age-related prediction patterns, particularly for Stroop test score prediction. For memory test score prediction, objective sleep measures and depressive score additionally contributed to subgroup separation and age-centered interaction patterns.

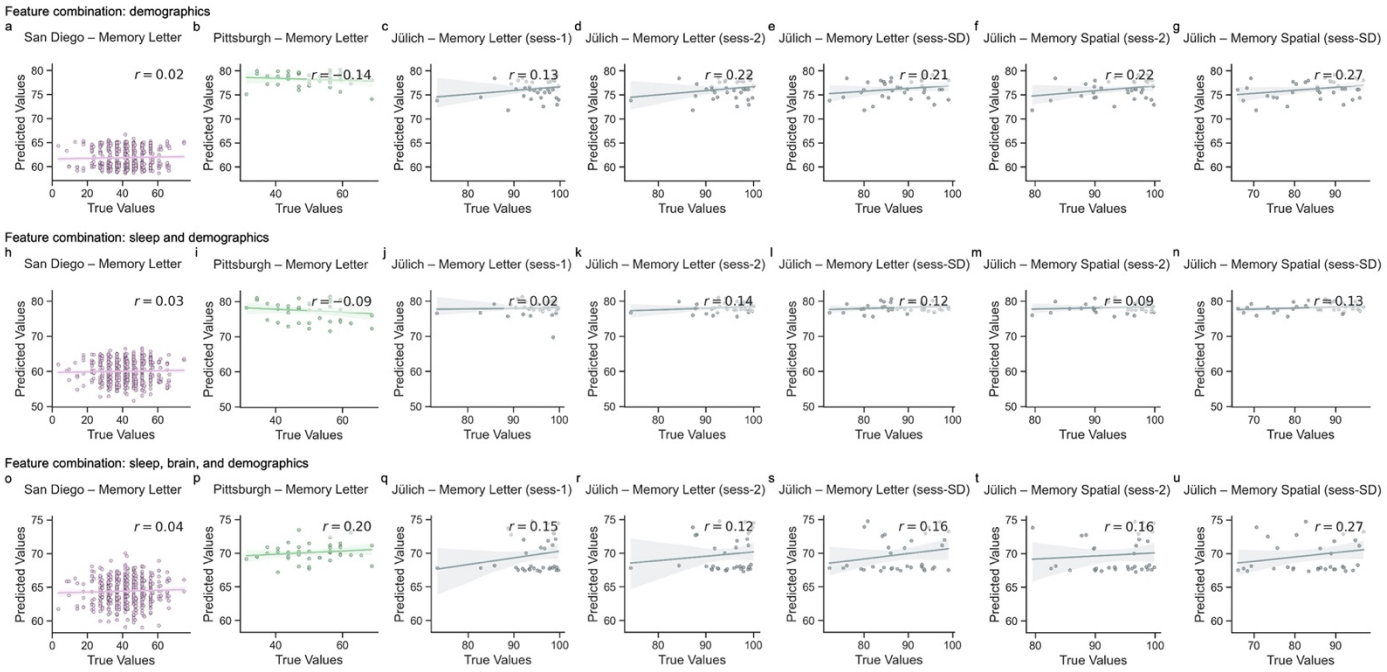
Supplementary figures



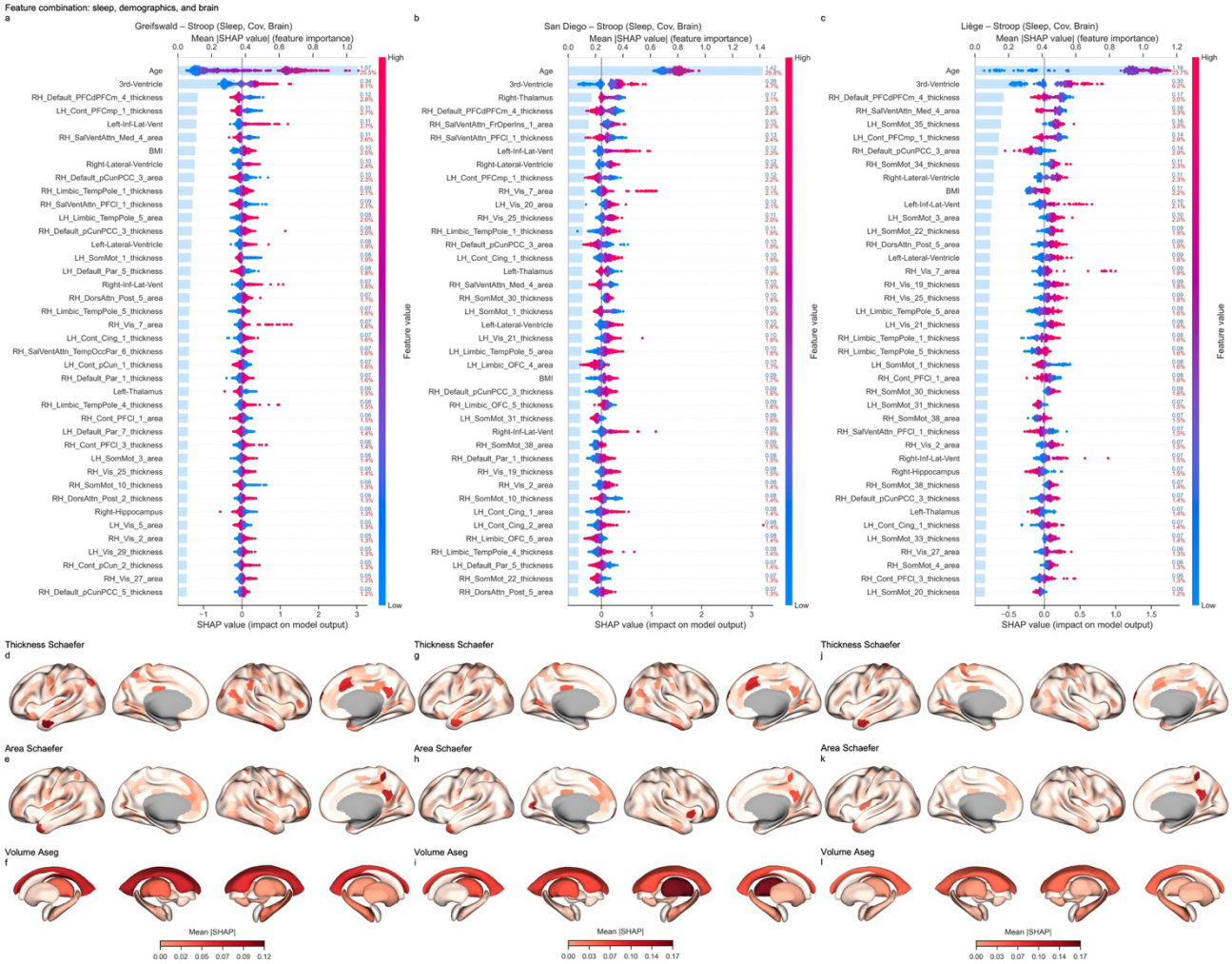
Supplementary Figure 1. Distributions of age, Stroop test scores, and memory test scores across cohorts. **a,** Age distribution across seven cohorts. **b**, Stroop test score distributions. Stroop test scores for Greifswald, Rotterdam, and Liège cohorts were calculated as the reaction time difference between interference and neutral conditions. San Diego cohort used Stroop interference norm-based T-scores from the Golden version. The Pittsburgh cohort utilized executive function scores based on NAB Mazes. **c**, Memory test score distributions. Different memory tasks were used across cohorts, with scores standardized as accuracy percentages. Greifswald employed the Nuremberg Age Inventory (NAI) word list, Liège, Stockholm, and Jülich cohorts assessed working memory performance through n-back tasks, Pittsburgh evaluated working memory via WMS-III, and San Diego utilized WMS-III tasks including letter-number sequencing and digit span.



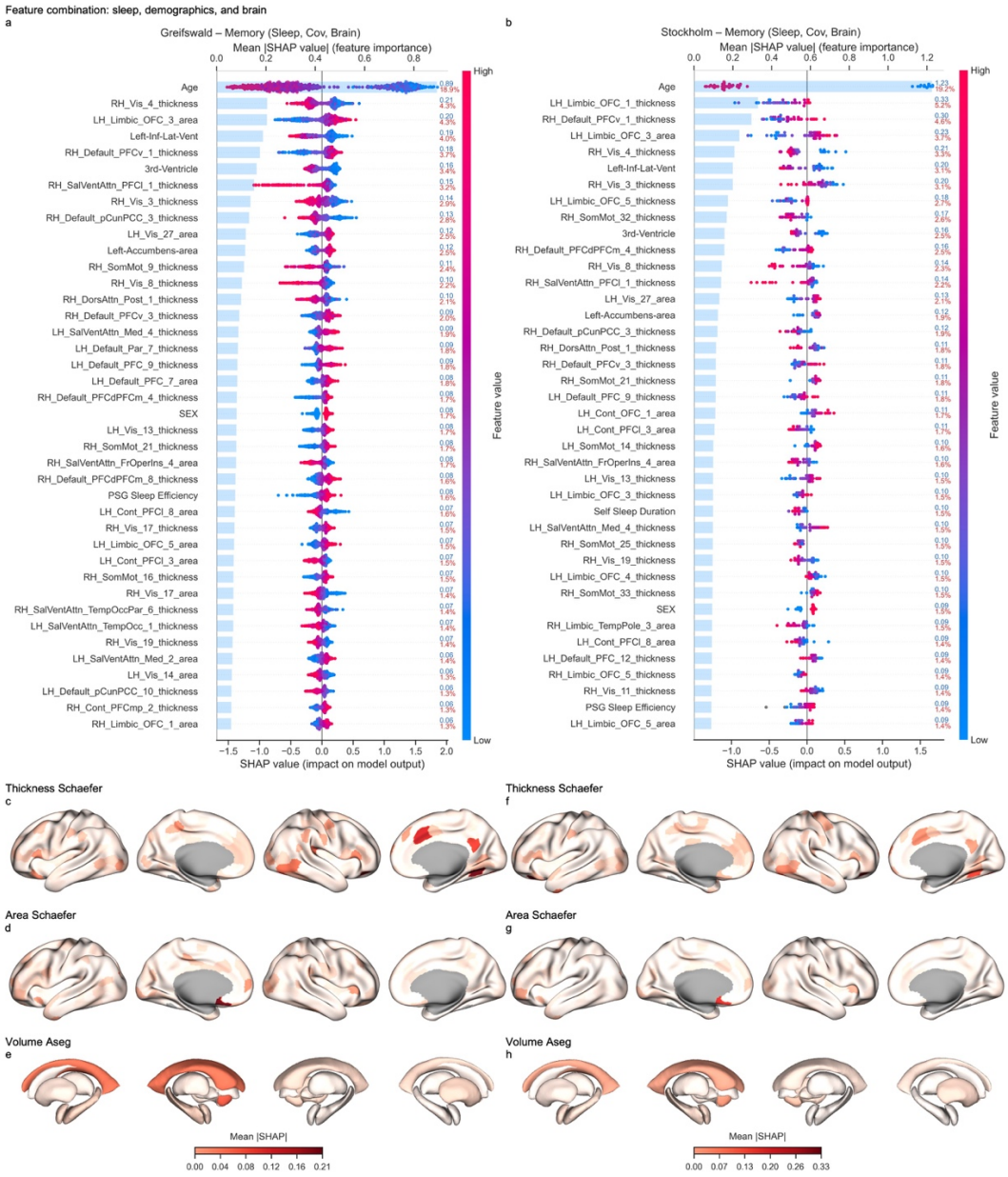
Supplementary Figure 2. Prediction of cognitive performance by different ML model across feature sets in the Greifswald cohort. **a–c,** Stroop test score prediction with demographics data, sleep measurements with demographics; and sleep with demographics and parcel-based brain morphometry data. **d–f,** Memory test score prediction with the same configurations. In each panel, boxplots show the distribution of test Spearman's ρ from 5-fold 10-repeat nested CV for seven algorithms: linear regression, ridge regression, SVM with linear kernel, SVM with rbf kernel, random forest, XGBoost, and AutoGluon. Pairwise model comparisons used Nadeau–Bengio corrected resampled paired t-tests on nested-CV scores with Bonferroni adjustment. Statistical significance: * P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001.



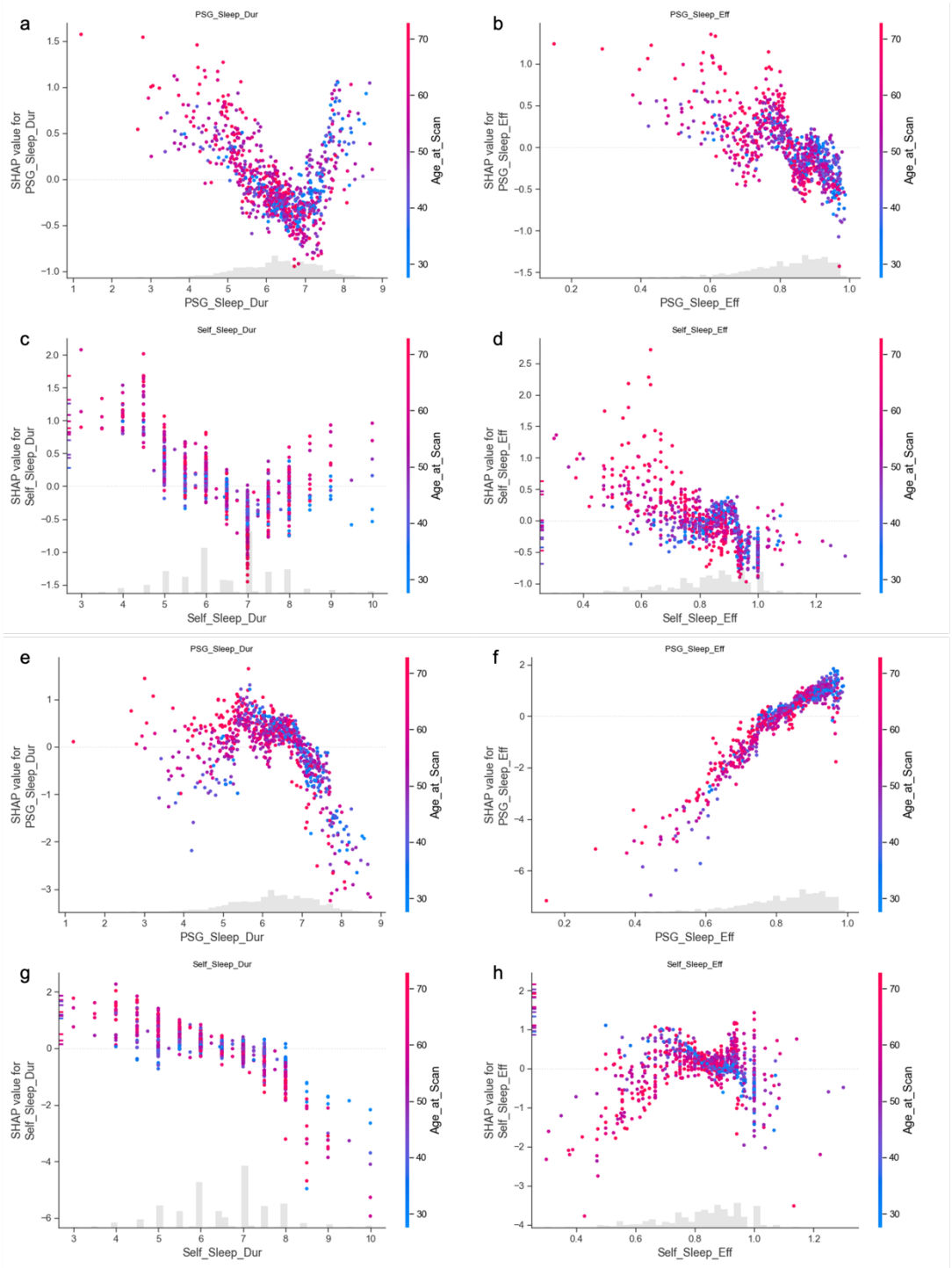
Supplementary Figure 3. Out-of-cohort validation of memory test score (rest) prediction by AutoGluon. AutoGluon models were trained in the Greifswald cohort and evaluated on San Diego, Pittsburgh, and Jülich cohorts for spatial memory test score prediction. **a-e,** Model inputs are demographic features only. **f-j,** Model inputs are sleep measurements and demographics. **k-o,** Model inputs are sleep, demographics and parcel-based brain morphometry data. Within each group the panels from left to right correspond to San Diego, Pittsburgh, Jülich session 1, Jülich session 2, and Jülich session sleep deprivation (SD). Each panel displays predicted versus true Stroop test scores. Significance of the Spearman's rank correlation coefficient is indicated by: *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.



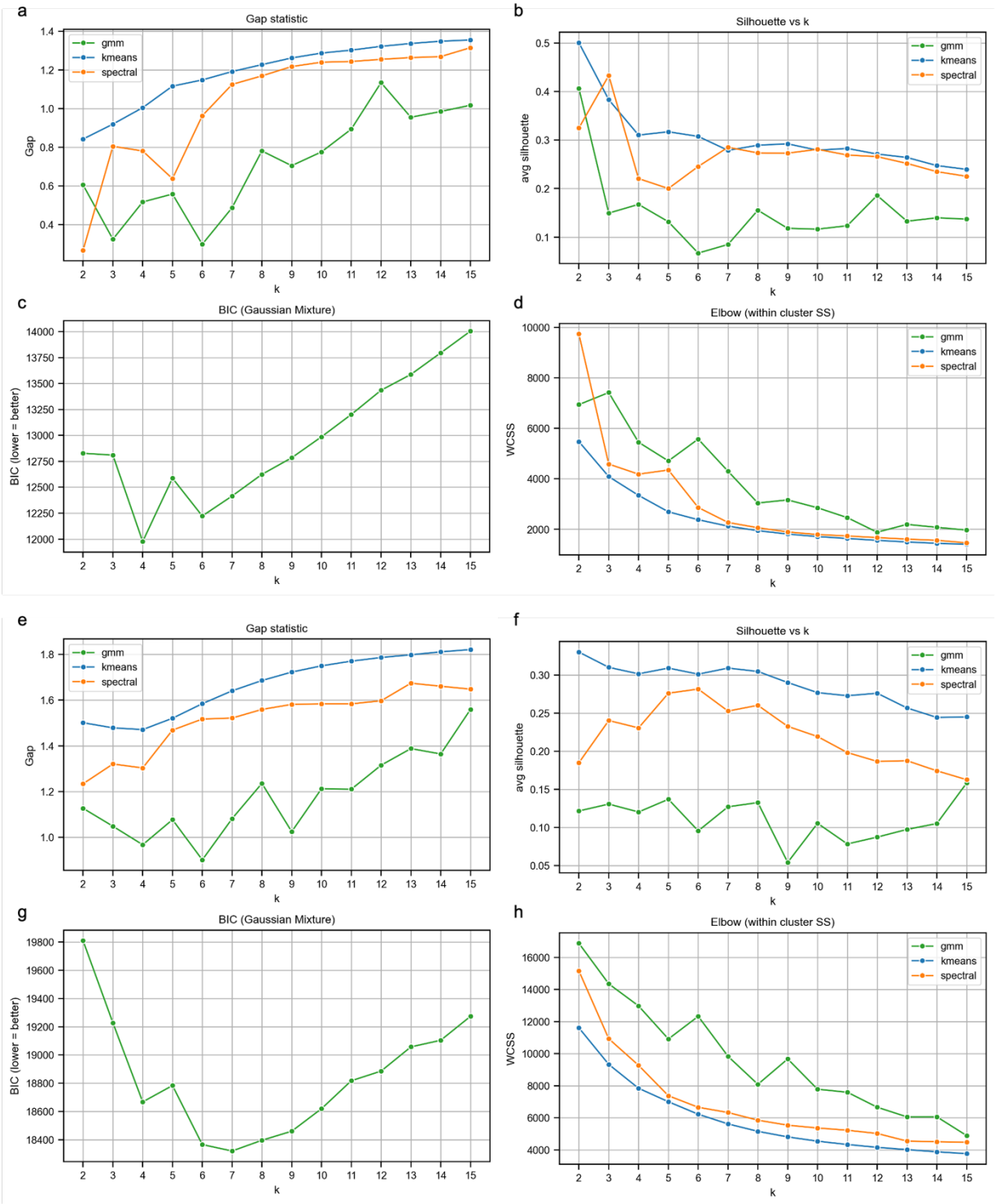
Supplementary Figure 4. Model explanation by SHAP for Stroop test score prediction. **a–c**, SHAP based model explanations for Stroop test score prediction in Greifswald (a), San Diego (b), and Liège (c) using AutoGluon trained with sleep, demographics, and parcel-based brain morphometry. Each panel shows a beeswarm plot of SHAP values for individual features, where each point represents a participant and points color corresponds to the underlying feature value. Each panel also includes a horizontal bar plot that displays the mean absolute SHAP value of each feature across all participants. Blue numbers next to each bar report the mean absolute SHAP value of that feature and red numbers report the percentage contribution of that feature to the total mean absolute SHAP value across all features. For Stroop test scores, lower values indicate faster reaction times and better performance, so negative SHAP values correspond to features that decrease the predicted Stroop score, whereas positive SHAP values correspond to features that increase the predicted score. **d–f**, Cortical and subcortical maps of parcel-wise mean absolute SHAP values in the Greifswald cohort for cortical thickness (d), cortical surface area (e), and subcortical volumes (f). **g–i**, Corresponding maps for the San Diego cohort. **j–l**, Corresponding maps for the Liège cohort. Within each set of maps, warmer colors denote parcels with larger mean absolute SHAP values.



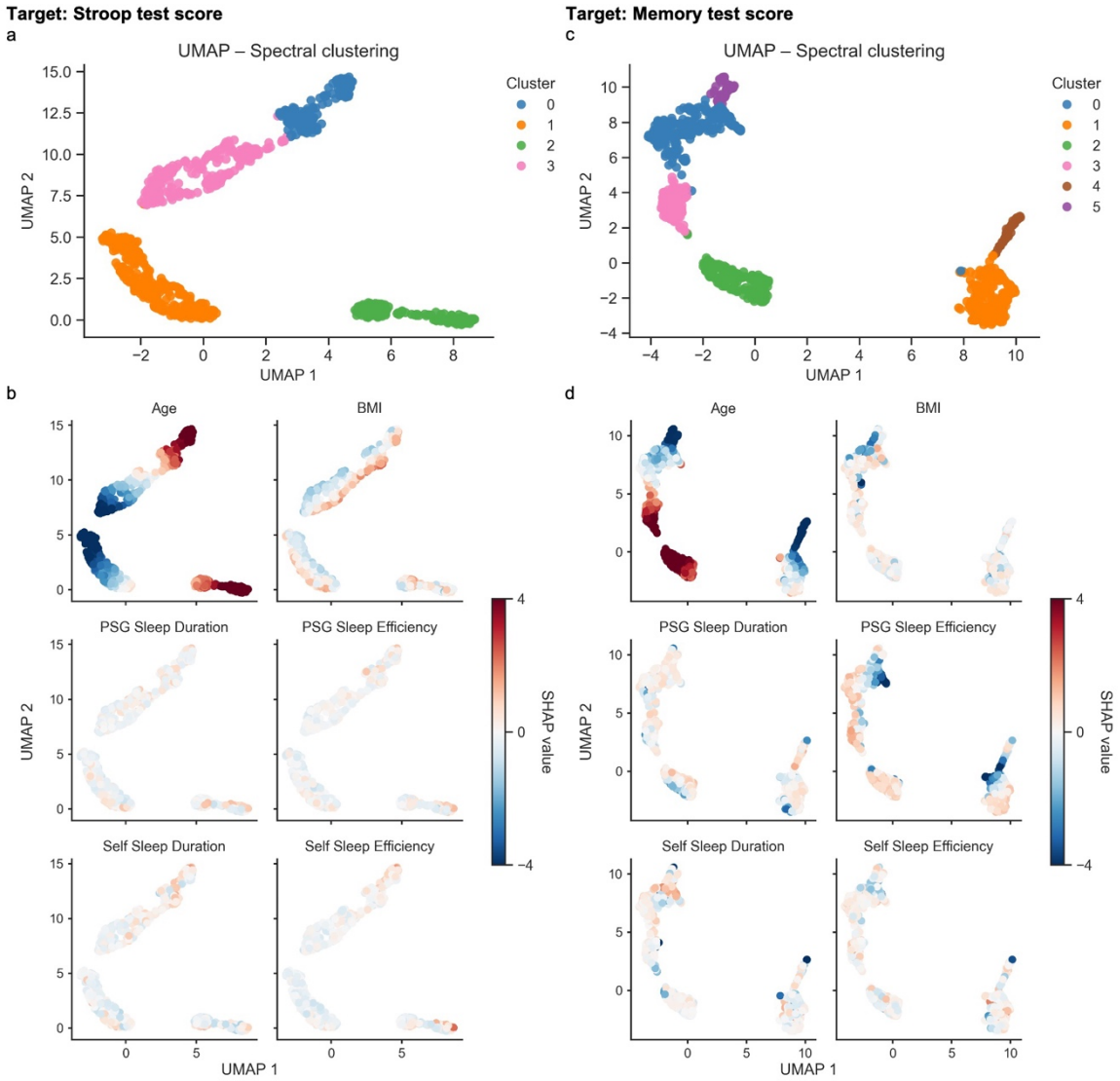
Supplementary Figure 5. Model explanation by SHAP for Memory test score prediction. **a–b**, SHAP based model explanations for memory test score prediction in Greifswald (a) and Stockholm (b) using AutoGluon models trained with sleep, demographics, and parcel-based brain morphometry. Each panel shows a beeswarm plot of SHAP values for individual features, where each point represents a participant and points color corresponds to the underlying feature value. Each panel also includes a horizontal bar plot that displays the mean absolute SHAP value of each feature across all participants. Blue numbers next to each bar report the mean absolute SHAP value of that feature and red numbers report the percentage contribution of that feature to the total mean absolute SHAP value across all features. For memory test scores, higher values indicate better performance, so positive SHAP values correspond to features that increase the predicted memory score, whereas negative SHAP values correspond to features that decrease the predicted score. **c–e**, Cortical and subcortical maps of parcel-wise mean absolute SHAP values in the Greifswald cohort for Schaefer cortical thickness (c), Schaefer cortical surface area (d), and FreeSurfer Aseg subcortical volumes (e). **f–h**, Corresponding maps for the Stockholm cohort for cortical thickness (f), cortical surface area (g), and subcortical volumes (h). Within each set of maps, warmer colors denote parcels with larger mean absolute SHAP values.



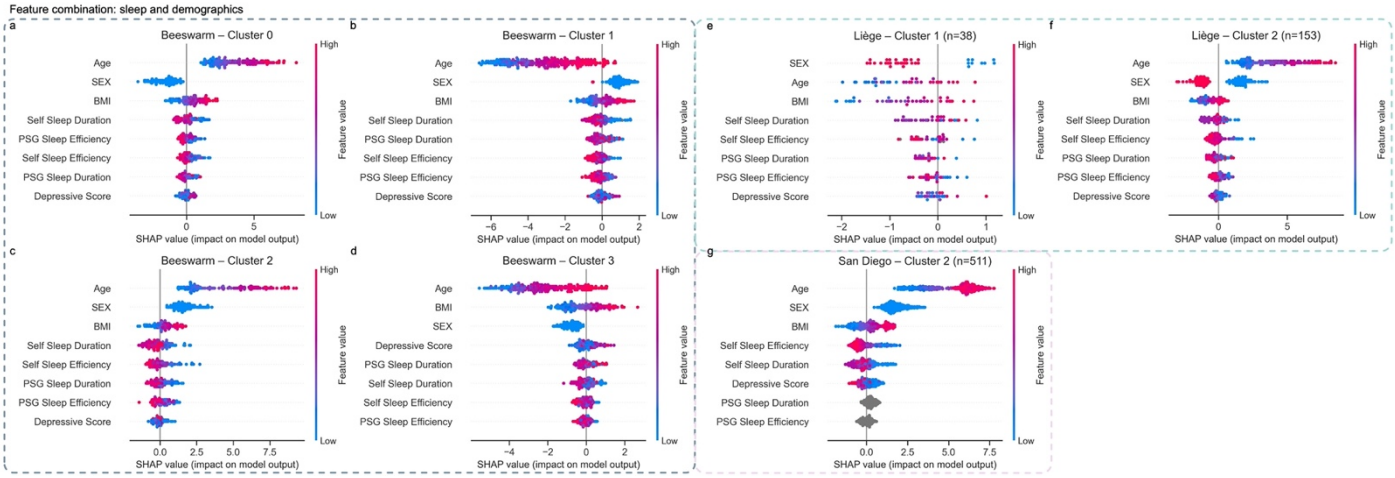
Supplementary Figure 6. Feature interactions based on SHAP dependence plots between age and sleep measurements. **a-d,** SHAP dependence plots from the AutoGluon model explaining predictions of Stroop test scores based on sleep measures and demographic variables in the Greifswald dataset. **e-h,** SHAP dependence plots from the AutoGluon model explaining predictions of memory test scores using the same feature combination in the Greifswald dataset. Each panel plots the raw value of a sleep metric on the x-axis against its SHAP value on the y-axis (feature effect on the model output for that participant). Points are colored by age (blue = younger, red = older) to visualize the interaction between age and the sleep feature.



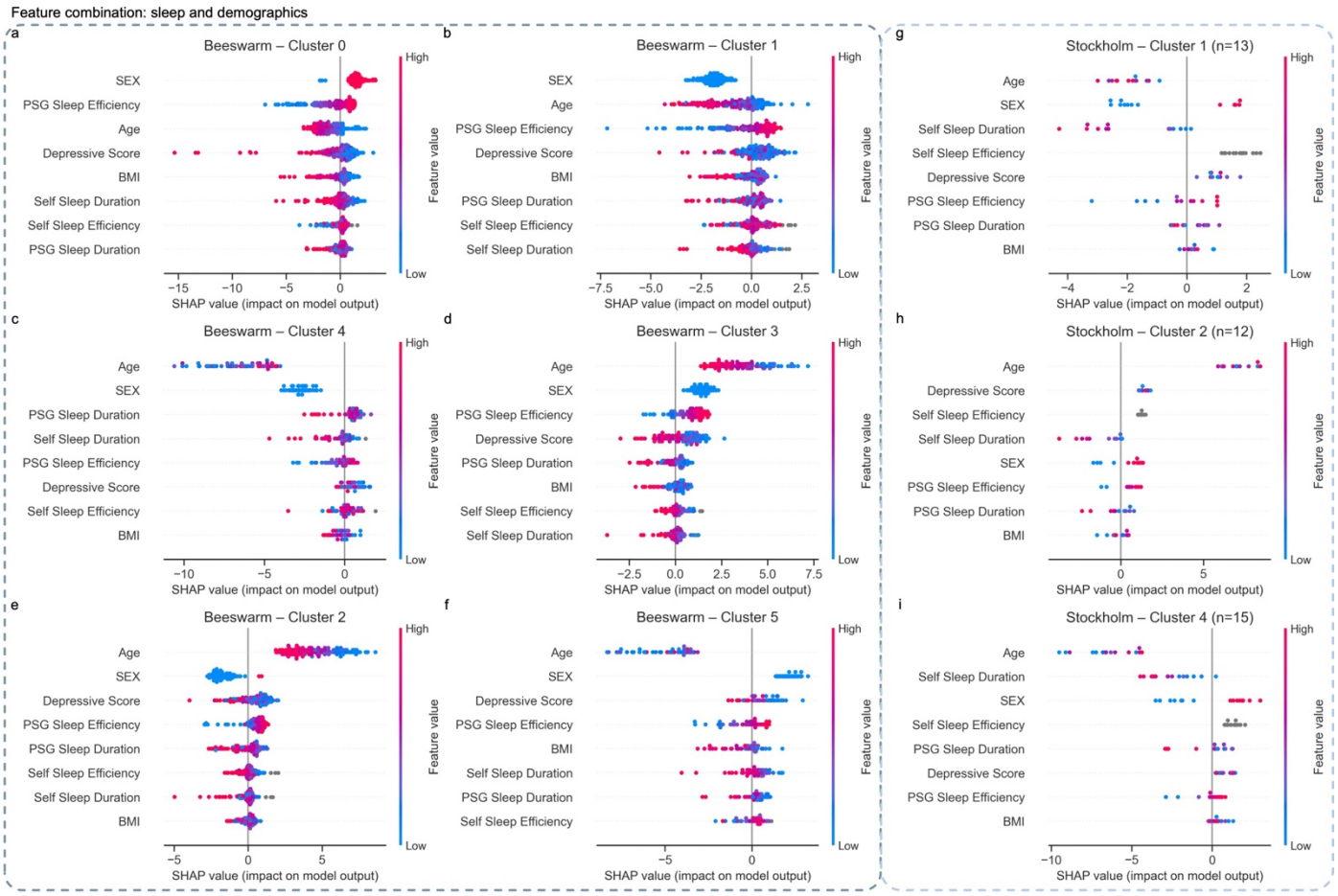
Supplementary Figure 7. Evaluation of clustering metrics and determination of cluster number based on SHAP values. Multiple clustering evaluation metrics were calculated, including gap statistics, silhouette scores, Bayesian information criterion (BIC), and within-cluster sum of squares (WCSS; Elbow method). These metrics were used collectively, along with UMAP dimensionality reduction visualization, to determine the optimal number of clusters. Four clusters were selected for Stroop test score prediction, and six clusters were identified for memory test score prediction based on SHAP values of AutoGluon with model input sleep measurements and demographic information. Additionally, clustering methods including Gaussian Mixture Model (GMM), K-means, and spectral clustering were compared. Spectral clustering was chosen as it provided clusters with the clearest interpretability and aligned best with the UMAP visualization.



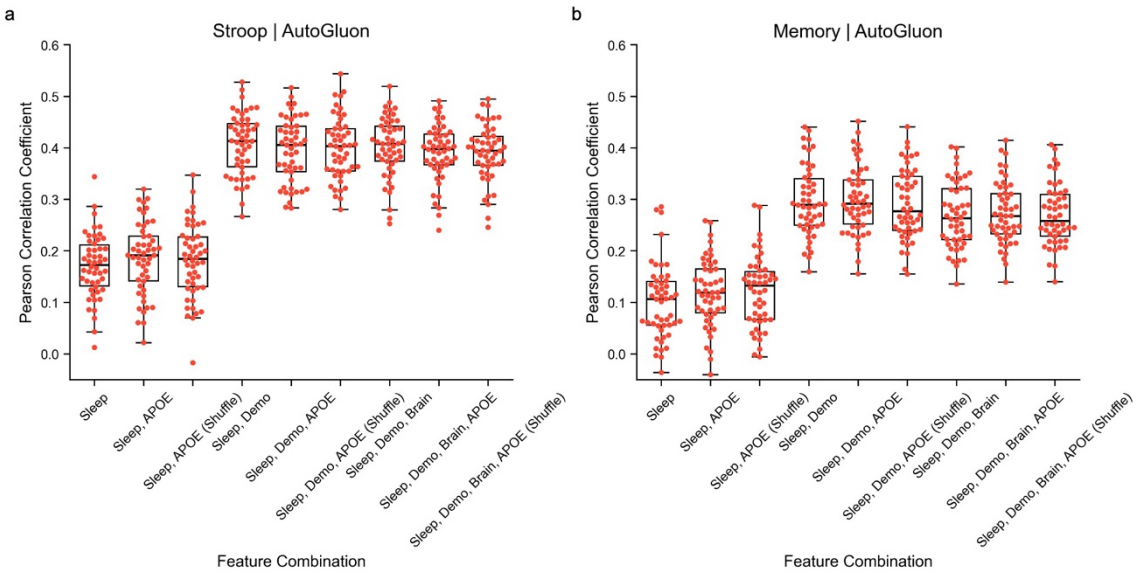
Supplementary Figure 8. SHAP-based clustering and feature-specific SHAP patterns in the Greifswald cohort. **a,** Two-dimensional UMAP embedding of participant-level SHAP profiles from the AutoGluon model trained with sleep measurements and demographics for Stroop test score prediction. Points are colored by spectral-clustering assignment, with four clusters identified. **b,** The same UMAP embedding for Stroop test score prediction, with points colored by feature-specific SHAP values for age, BMI, PSG-derived sleep duration, PSG-derived sleep efficiency, self-reported sleep duration, and self-reported sleep efficiency. **c,** Two-dimensional UMAP embedding of participant-level SHAP profiles from the AutoGluon model trained with sleep measurements and demographics for memory test score prediction. Points are colored by spectral-clustering assignment, with six clusters identified. **d,** The same UMAP embedding for memory test score prediction, with points colored by feature-specific SHAP values for the same variables. In **b** and **d**, color indicates the direction and magnitude of each feature contribution to the model prediction. Positive SHAP values contribute to increase the predicted Stroop or memory test score, whereas negative SHAP values contribute to decrease the predicted score. The cluster labels shown here are used consistently in Supplementary Table 3, Figures 6 and 7, and Supplementary Figures 9 and 10.



Supplementary Figure 9. SHAP results for each subgroup identified from clustering the SHAP explanations of the Stroop test score prediction. **a–d,** SHAP beeswarm plots of four clusters in Greifswald cohort. **e–f,** SHAP beeswarm plots of two clusters in Liège cohort. **g,** SHAP beeswarm plots of one cluster in San Diego cohort.



Supplementary Figure 10. SHAP results for each subgroup identified from clustering the SHAP explanations of the Stroop test score prediction. **a–f,** SHAP beeswarm plots of six clusters in Greifswald cohort. **g–i,** SHAP beeswarm plots of three clusters in Stockholm cohort.



Supplementary Figure 11. Evaluation of APOE-ε4 status in the prediction of Stroop and memory test scores. To assess the contribution of APOE-ε4 status to Stroop and memory test scores prediction, models were trained with feature sets including APOE status as well as with shuffled APOE labels, using AutoGluon. Model performance for both Stroop (**a**) and memory (**b**) test scores was compared across feature combinations by corrected t-test. The inclusion of APOE-ε4 information did not significantly enhance prediction performance for Stroop and memory test score.

Supplementary tables

Supplementary Table 1. Participants' demographics, depressive score, APOE-ε4 carriership, and sleep measurements by cohorts

Location / Cohort name			N	Age, (Years)		Female	BMI	Depression score		APOE-ε4 present	Sleep Duration, (Hours)		Sleep Efficiency, (%)	
				Mean ± SD	Range	N, (%)	Mean ± SD	Mean ± SD	Type	N, (%)	Mean ± SD	Type	Mean ± SD	Type
Greifswald	SHIP-Trend		845	52.7 ± 13.4	21.8 – 81.4	404 (47.8)	28.1 ± 4.4	8.3 ± 6.0	BDI-II	186 (22.0)	6.3 ± 1.0	PSG	82.1 ± 11.7	PSG
											6.5 ± 1.2	PSQI	83.0 ± 14.2	PSQI
Rotterdam	The Rotterdam Study		700	62.7 ± 6.4	51.5 – 95.1	372 (53.1)	27.3 ± 4.1	5.2 ± 6.5	CES-D	195 (27.8)	6.3 ± 1.1	PSG	81.1 ± 10.5	PSG
											6.9 ± 1.2	PSQI	92.2 ± 12.8	PSQI
San Diego	VETSA		511	67.5 ± 2.6	61.4 – 71.7	0 (0.0)	29.2 ± 4.3	6.8 ± 7.0	CES-D	101 (19.8)	6.7 ± 1.2	PSQI	84.2 ± 14.2	PSQI
Liège	COF		101	59.4 ± 5.3	50.0 – 69.0	68 (67.3)	24.6 ± 2.9	5.2 ± 4.4	BDI-II	20 (19.8)	6.6 ± 0.8	PSG	82.9 ± 9.5	PSG
											7.1 ± 1.1	PSQI	82.1 ± 18.5	PSQI
	COGNAP		90	68.9 ± 5.2	59.5 – 81.8	30 (33.3)	25.1 ± 2.8	4.0 ± 3.4	BDI-II	\	6.4 ± 0.7	PSG	80.1 ± 8.8	PSG
											7.3 ± 1.2	PSQI	85.9 ± 15.7	PSQI
Stockholm	Sleepy Brain		48	55.4 ± 21.1	21.0 – 75.0	28 (58.3)	24.1 ± 3.9	2.0 ± 1.8	HADS*	16 (33.3)	6.3 ± 1.2	PSQI	80.9 ± 10.5	PSQI
											8.3 ± 1.0	KSQ	\	\
Pittsburg	PyNEL		41	20.4 ± 4.2	18.0 – 33.0	22 (53.7)	24 ± 5.6	\	\	\	7.6 ± 1.0	PSG	88.8 ± 9.3	PSG
											8.0 ± 1.3	PSQI	94.6 ± 7.3	PSQI
Jülich	Somnosafe	Sess-1	36	27.1 ± 4.6	20.4 – 39.3	14 (38.9)	23.3 ± 2.4	\	\	\	6.9 ± 0.8	PSG	85.6 ± 9.8	PSG
		Sess-2									7.2 ± 0.4	PSG	89.5 ± 4.4	PSG
		Sess-SD									7.1 ± 0.4	PSG	88.6 ± 4.4	PSG
Total			2372	58.8 ± 13.0	18.0 – 95.1	938 (39.5)								

* Hospital Anxiety and Depression Scale (HADS) transformed to the Beck Depression Inventory (BDI-II) scale.

Abbreviations: N, number of participants; SD, standard deviation; BMI, body mass index; PSG, Polysomnography; PSQI, Pittsburgh Sleep Quality Index; CES-D, Center for Epidemiologic Studies Depression Scale; KSQ, Karolinska Sleep Questionnaire; Sess, session; SD, sleep deprivation.

Supplementary Table 2. Cohort-specific information

Cohort	Dataset name	Exclusion/inclusion criteria	Scanner manufacturer and type
Greifswald	SHIP-Trend	Participants were included if they completed both whole-body MRI and overnight polysomnography with adequate recording duration and quality, and if MRI showed no major structural brain abnormalities. Individuals with contraindications to MRI, for example, a pacemaker or those with multiple sclerosis, epilepsy, Parkinson's disease, stroke, or other major neurologic or psychiatric disorders, were excluded.	1.5T Magnetom Avanto (Siemens, Erlangen, Germany)
Rotterdam	The Rotterdam Study	Participants from the population-based Rotterdam Study cohort were invited to a polysomnography substudy, with no exclusion criteria other than the ability to understand instructions. To be included in this analysis, participants needed to have 1) valid PSG data; 2) MRI data collected within 1 year of PSG; 3) Cognitive data collected within 2 years of PSG	1.5T Signa Excite II (GE Healthcare)
San Diego	VETSA	Raw images were visually inspected for quality and excluded for excessive motion or acquisition artifacts. Processed structural images were manually edited to correct errors in cortical surface reconstruction, and subcortical segmentation was visually reviewed. Subjects with major segmentation or surface reconstruction failures were excluded from analysis. Standard MRI exclusion criteria (e.g., no metal in the body) were applied.	GE 3T Discovery 750x scanners with 8-channel phased-array head coils
Liège	COF	\	\
	COGNAP	Participants were included if they had normal global cognition and no clinically relevant depression or anxiety based on standardized questionnaires and if screening polysomnography did not indicate a major sleep disorder. Exclusion criteria included underweight or obesity, a history of diagnosed psychiatric conditions or severe brain trauma, chronic use of medication affecting the central nervous system, for example, sleep medication, anxiolytics or beta blockers, diabetes, smoking, high caffeine intake, excessive alcohol or other drug use, and recent transmeridian travel.	\
Stockholm	Sleepy Brain	Participants were included if they could safely undergo MRI and had normal or corrected to normal color vision without major refractive error and without self-reported insomnia or extreme morningness or eveningness. Exclusion criteria included any MRI contraindications such as ferromagnetic items in the body, claustrophobia or pregnancy, left handedness, psychiatric including high degree of self-reported depressive or anxiety symptoms or neurological illness, addiction, hypertension, diabetes, use of psychoactive or immune modulatory medication, daily nicotine use, high caffeine intake, clinically significant insomnia and marked circadian preference, including habitual bedtime later than one in the morning.	3T Discovery 750 (General Electric)
Pittsburg	PyNEL	Participants were included if they had no DSM diagnosis and no first-degree family history of psychotic disorder or mood disorder with psychotic features. Exclusion criteria included night shift work, current pregnancy, traumatic brain injury within the past year, medical conditions that could affect brain structure or function, and drug dependence within the past year.	\
Jülich	Somnosafe	Only healthy, non-smoking participants with regular sleep times were included.	\

Supplementary Table 3. Stroop test score prediction in Greifswald across models and feature combinations

Sleep								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	7.55 (0.44)	7.41 (0.45)	7.41 (0.45)	**7.20 (0.45)**	7.43 (0.50)	7.48 (0.43)	7.75 (0.56)	7.75 (0.54)
Test RMSE	10.76 (1.16)	**10.55 (1.16)**	**10.55 (1.16)**	10.72 (1.18)	10.83 (1.15)	10.62 (1.15)	11.14 (1.57)	11.13 (1.57)
Test R2	-0.00 (0.01)	**0.03 (0.04)**	**0.03 (0.04)**	0.00 (0.03)	-0.02 (0.03)	0.02 (0.04)	0.02 (0.02)	0.02 (0.02)
Test Pearson r	-0.00 (0.00)	**0.21 (0.08)**	**0.21 (0.08)**	**0.21 (0.08)**	0.13 (0.07)	0.19 (0.07)	0.16 (0.06)	0.17 (0.06)
Test Spearman r	nan (nan)	**0.21 (0.07)**	**0.21 (0.07)**	**0.21 (0.07)**	0.16 (0.07)	0.18 (0.06)	0.16 (0.06)	0.16 (0.06)
Demo								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	7.55 (0.44)	6.84 (0.41)	6.84 (0.41)	**6.65 (0.43)**	6.67 (0.44)	6.94 (0.41)	7.18 (0.56)	7.18 (0.52)
Test RMSE	10.76 (1.16)	**9.85 (1.10)**	**9.85 (1.10)**	10.02 (1.17)	10.00 (1.17)	9.95 (1.10)	10.38 (1.51)	10.36 (1.48)
Test R2	-0.00 (0.01)	**0.16 (0.05)**	**0.16 (0.05)**	0.13 (0.04)	0.13 (0.04)	0.14 (0.06)	**0.15 (0.05)**	**0.15 (0.05)**
Test Pearson r	-0.00 (0.00)	**0.41 (0.06)**	**0.41 (0.06)**	0.41 (0.06)	0.41 (0.06)	0.39 (0.07)	0.40 (0.06)	**0.41 (0.06)**
Test Spearman r	nan (nan)	**0.45 (0.06)**	**0.45 (0.06)**	**0.45 (0.06)**	0.44 (0.06)	0.41 (0.06)	0.42 (0.05)	0.43 (0.05)
Demo + Sleep								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	7.55 (0.44)	6.86 (0.43)	6.86 (0.43)	6.66 (0.44)	**6.65 (0.44)**	6.89 (0.41)	7.21 (0.55)	7.16 (0.52)
Test RMSE	10.76 (1.16)	**9.86 (1.11)**	**9.86 (1.11)**	10.01 (1.16)	9.97 (1.18)	9.88 (1.15)	10.44 (1.54)	10.34 (1.49)
Test R2	-0.00 (0.01)	**0.16 (0.06)**	**0.16 (0.06)**	0.13 (0.04)	0.14 (0.05)	0.15 (0.05)	0.14 (0.05)	0.15 (0.04)
Test Pearson r	-0.00 (0.00)	**0.41 (0.06)**	**0.41 (0.06)**	**0.41 (0.06)**	**0.41 (0.06)**	0.40 (0.07)	0.39 (0.06)	**0.41 (0.06)**
Test Spearman r	nan (nan)	**0.44 (0.06)**	**0.44 (0.06)**	**0.44 (0.06)**	**0.44 (0.06)**	0.42 (0.06)	0.41 (0.06)	0.43 (0.05)
Demo + Sleep (shuffle)								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	7.55 (0.44)	6.90 (0.40)	6.90 (0.40)	6.68 (0.42)	**6.75 (0.44)**	6.94 (0.39)	7.18 (0.57)	7.18 (0.54)
Test RMSE	10.76 (1.16)	**9.87 (1.09)**	**9.87 (1.09)**	10.07 (1.16)	10.11 (1.16)	9.94 (1.09)	10.44 (1.49)	10.41 (1.47)
Test R2	-0.00 (0.01)	**0.15 (0.05)**	**0.15 (0.05)**	0.12 (0.04)	0.11 (0.04)	0.14 (0.06)	0.14 (0.06)	0.14 (0.05)
Test Pearson r	-0.00 (0.00)	**0.41 (0.06)**	**0.41 (0.06)**	0.40 (0.06)	0.39 (0.06)	0.39 (0.06)	0.39 (0.07)	0.39 (0.07)
Test Spearman r	nan (nan)	**0.44 (0.06)**	**0.44 (0.06)**	**0.44 (0.06)**	0.42 (0.07)	0.42 (0.06)	0.41 (0.06)	0.41 (0.05)
Brain								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	7.55 (0.44)	14.81 (0.76)	14.50 (0.73)	7.46 (0.43)	**6.99 (0.44)**	7.23 (0.40)	7.34 (0.56)	7.35 (0.56)
Test RMSE	10.76 (1.16)	18.99 (0.85)	18.62 (0.84)	10.75 (1.10)	10.38 (1.14)	**10.27 (1.07)**	10.66 (1.47)	10.60 (1.49)
Test R2	-0.00 (0.01)	-2.23 (0.69)	-2.10 (0.65)	-0.01 (0.06)	0.07 (0.03)	0.08 (0.04)	0.10 (0.04)	**0.11 (0.04)**
Test Pearson r	-0.00 (0.00)	0.09 (0.07)	0.10 (0.07)	0.25 (0.06)	0.30 (0.06)	0.31 (0.06)	0.33 (0.06)	**0.36 (0.07)**
Test Spearman r	nan (nan)	0.09 (0.06)	0.09 (0.07)	0.25 (0.05)	0.33 (0.06)	0.33 (0.06)	0.35 (0.06)	**0.37 (0.06)**
Demo + Sleep + Brain								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	7.55 (0.44)	14.64 (0.77)	14.35 (0.75)	7.39 (0.44)	**6.95 (0.44)**	7.04 (0.38)	7.23 (0.57)	7.20 (0.54)
Test RMSE	10.76 (1.16)	18.80 (0.87)	18.45 (0.87)	10.66 (1.12)	10.33 (1.14)	**10.08 (1.04)**	10.50 (1.51)	10.42 (1.52)
Test R2	-0.00 (0.01)	-2.16 (0.67)	-2.04 (0.64)	0.01 (0.06)	0.08 (0.03)	0.12 (0.05)	0.13 (0.05)	**0.14 (0.04)**
Test Pearson r	-0.00 (0.00)	0.10 (0.07)	0.10 (0.07)	0.27 (0.06)	0.32 (0.06)	0.36 (0.06)	0.37 (0.06)	**0.40 (0.06)**
Test Spearman r	nan (nan)	0.10 (0.07)	0.10 (0.07)	0.27 (0.05)	0.35 (0.06)	0.40 (0.06)	0.39 (0.06)	**0.42 (0.05)**
Demo + Sleep + Brain (shuffle)								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	\	\	\	\	\	\	7.31 (0.55)	**7.27 (0.53)**
Test RMSE	\	\	\	\	\	\	10.61 (1.53)	**10.53 (1.53)**
Test R2	\	\	\	\	\	\	0.11 (0.05)	**0.12 (0.03)**
Test Pearson r	\	\	\	\	\	\	0.35 (0.06)	**0.38 (0.06)**

Test Spearman r	\	\	\	\	\	\	0.36 (0.06)	**0.40 (0.05)**
Sleep + APOE								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	\	\	\	\	\	\	7.75 (0.55)	**7.63 (0.55)**
Test RMSE	\	\	\	\	\	\	11.14 (1.57)	**11.11 (1.55)**
Test R2	\	\	\	\	\	\	**0.02 (0.02)**	0.02 (0.03)
Test Pearson r	\	\	\	\	\	\	0.16 (0.06)	**0.19 (0.07)**
Test Spearman r	\	\	\	\	\	\	0.16 (0.06)	**0.19 (0.08)**
Sleep + APOE (shuffle)								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	\	\	\	\	\	\	7.74 (0.56)	**7.66 (0.54)**
Test RMSE	\	\	\	\	\	\	11.14 (1.57)	**11.12 (1.55)**
Test R2	\	\	\	\	\	\	**0.02 (0.02)**	0.02 (0.03)
Test Pearson r	\	\	\	\	\	\	0.16 (0.06)	**0.18 (0.07)**
Test Spearman r	\	\	\	\	\	\	0.16 (0.07)	**0.17 (0.08)**
Demo + Sleep + APOE								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	\	\	\	\	\	\	7.21 (0.57)	**7.08 (0.52)**
Test RMSE	\	\	\	\	\	\	10.44 (1.55)	**10.40 (1.52)**
Test R2	\	\	\	\	\	\	**0.14 (0.05)**	**0.14 (0.05)**
Test Pearson r	\	\	\	\	\	\	0.39 (0.07)	**0.40 (0.06)**
Test Spearman r	\	\	\	\	\	\	0.41 (0.06)	**0.42 (0.05)**
Demo + Sleep + APOE (shuffle)								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	\	\	\	\	\	\	7.21 (0.56)	**7.12 (0.54)**
Test RMSE	\	\	\	\	\	\	10.42 (1.53)	**10.37 (1.50)**
Test R2	\	\	\	\	\	\	0.14 (0.05)	**0.15 (0.05)**
Test Pearson r	\	\	\	\	\	\	0.39 (0.06)	**0.40 (0.06)**
Test Spearman r	\	\	\	\	\	\	0.41 (0.05)	**0.42 (0.05)**
Demo + Sleep + Brain + APOE								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	\	\	\	\	\	\	**7.21 (0.58)**	7.25 (0.54)
Test RMSE	\	\	\	\	\	\	10.46 (1.50)	**10.45 (1.53)**
Test R2	\	\	\	\	\	\	0.13 (0.05)	**0.14 (0.04)**
Test Pearson r	\	\	\	\	\	\	0.38 (0.06)	**0.39 (0.05)**
Test Spearman r	\	\	\	\	\	\	0.40 (0.06)	**0.42 (0.05)**
Demo + Sleep + Brain + APOE (shuffle)								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	\	\	\	\	\	\	**7.22 (0.60)**	7.24 (0.54)
Test RMSE	\	\	\	\	\	\	10.48 (1.51)	**10.44 (1.53)**
Test R2	\	\	\	\	\	\	0.13 (0.05)	**0.14 (0.04)**
Test Pearson r	\	\	\	\	\	\	0.38 (0.06)	**0.39 (0.05)**
Test Spearman r	\	\	\	\	\	\	0.40 (0.06)	**0.41 (0.05)**

Supplementary Table 4. Memory test score prediction in Greifswald across models and feature combinations

Sleep								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	12.69 (0.57)	12.57 (0.60)	12.57 (0.60)	12.56 (0.61)	**12.48 (0.58)**	12.64 (0.60)	12.64 (0.51)	12.65 (0.49)
Test RMSE	15.80 (0.71)	**15.62 (0.75)**	**15.62 (0.75)**	15.77 (0.76)	15.75 (0.75)	15.76 (0.76)	15.71 (0.63)	15.70 (0.62)
Test R2	-0.01 (0.01)	**0.01 (0.03)**	**0.01 (0.03)**	-0.00 (0.03)	-0.00 (0.03)	-0.00 (0.04)	0.00 (0.02)	0.00 (0.02)
Test Pearson r	-0.00 (0.00)	0.16 (0.06)	0.16 (0.06)	0.16 (0.06)	**0.17 (0.07)**	0.13 (0.08)	0.10 (0.07)	0.10 (0.07)
Test Spearman r	nan (nan)	0.14 (0.05)	0.14 (0.05)	**0.15 (0.06)**	**0.15 (0.06)**	0.12 (0.06)	0.10 (0.08)	0.10 (0.08)
Demo								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	12.69 (0.57)	12.05 (0.51)	12.05 (0.51)	**12.00 (0.55)**	12.02 (0.56)	12.01 (0.49)	12.03 (0.57)	12.02 (0.56)
Test RMSE	15.80 (0.71)	**15.05 (0.66)**	**15.05 (0.66)**	15.12 (0.71)	15.17 (0.72)	15.14 (0.59)	15.10 (0.68)	15.08 (0.67)
Test R2	-0.01 (0.01)	**0.08 (0.03)**	**0.08 (0.03)**	0.08 (0.04)	0.07 (0.04)	0.07 (0.05)	0.08 (0.04)	0.08 (0.04)
Test Pearson r	-0.00 (0.00)	**0.31 (0.05)**	**0.31 (0.05)**	**0.31 (0.05)**	0.30 (0.05)	0.29 (0.06)	0.29 (0.06)	0.30 (0.06)
Test Spearman r	nan (nan)	**0.30 (0.06)**	**0.30 (0.06)**	**0.30 (0.06)**	0.29 (0.06)	0.28 (0.06)	0.29 (0.06)	0.29 (0.06)
Demo + Sleep								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	12.69 (0.57)	11.91 (0.56)	11.91 (0.56)	11.85 (0.57)	**11.83 (0.56)**	12.04 (0.55)	12.01 (0.59)	11.98 (0.58)
Test RMSE	15.80 (0.71)	**14.95 (0.68)**	**14.95 (0.68)**	14.97 (0.72)	14.95 (0.71)	15.18 (0.65)	15.16 (0.70)	15.06 (0.69)
Test R2	-0.01 (0.01)	0.10 (0.04)	0.10 (0.04)	0.09 (0.03)	**0.10 (0.03)**	0.07 (0.04)	0.07 (0.04)	0.08 (0.04)
Test Pearson r	-0.00 (0.00)	0.33 (0.05)	0.33 (0.05)	0.33 (0.05)	**0.34 (0.05)**	0.29 (0.06)	0.28 (0.07)	0.30 (0.07)
Test Spearman r	nan (nan)	0.32 (0.05)	0.32 (0.05)	0.32 (0.05)	**0.33 (0.05)**	0.29 (0.06)	0.28 (0.07)	0.29 (0.07)
Demo + Sleep (shuffle)								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	12.69 (0.57)	12.13 (0.53)	12.12 (0.53)	**12.09 (0.56)**	12.09 (0.57)	12.17 (0.53)	12.17 (0.57)	12.11 (0.53)
Test RMSE	15.80 (0.71)	**15.16 (0.68)**	**15.16 (0.68)**	15.23 (0.74)	15.24 (0.75)	15.25 (0.66)	15.29 (0.68)	15.19 (0.65)
Test R2	-0.01 (0.01)	**0.07 (0.03)**	**0.07 (0.03)**	0.06 (0.03)	0.06 (0.03)	0.06 (0.03)	0.05 (0.04)	0.07 (0.04)
Test Pearson r	-0.00 (0.00)	**0.29 (0.05)**	**0.29 (0.05)**	**0.29 (0.05)**	**0.29 (0.05)**	0.27 (0.06)	0.25 (0.07)	0.27 (0.07)
Test Spearman r	nan (nan)	**0.28 (0.06)**	**0.28 (0.06)**	**0.28 (0.06)**	**0.28 (0.06)**	0.26 (0.06)	0.25 (0.07)	0.27 (0.07)
Brain								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	12.69 (0.57)	24.52 (1.02)	24.00 (1.00)	12.83 (0.60)	**12.29 (0.58)**	12.46 (0.56)	12.34 (0.51)	12.35 (0.48)
Test RMSE	15.80 (0.71)	30.94 (1.21)	30.30 (1.17)	16.12 (0.80)	15.51 (0.78)	15.52 (0.71)	15.45 (0.68)	**15.42 (0.63)**
Test R2	-0.01 (0.01)	-2.90 (0.53)	-2.74 (0.50)	-0.05 (0.06)	0.03 (0.04)	0.03 (0.02)	0.03 (0.03)	**0.04 (0.02)**
Test Pearson r	-0.00 (0.00)	0.04 (0.09)	0.04 (0.09)	0.16 (0.07)	0.21 (0.07)	0.19 (0.05)	0.21 (0.06)	**0.22 (0.06)**
Test Spearman r	nan (nan)	0.05 (0.09)	0.05 (0.09)	0.15 (0.07)	0.22 (0.07)	0.20 (0.05)	0.21 (0.06)	**0.22 (0.06)**
Demo + Sleep + Brain								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	12.69 (0.57)	23.62 (0.95)	23.15 (0.93)	12.77 (0.60)	12.26 (0.59)	12.21 (0.57)	12.23 (0.53)	**12.18 (0.49)**
Test RMSE	15.80 (0.71)	29.84 (1.11)	29.25 (1.09)	16.02 (0.80)	15.46 (0.77)	15.26 (0.68)	15.31 (0.64)	**15.23 (0.63)**
Test R2	-0.01 (0.01)	-2.63 (0.47)	-2.48 (0.45)	-0.04 (0.06)	0.03 (0.04)	**0.06 (0.03)**	0.05 (0.04)	**0.06 (0.03)**
Test Pearson r	-0.00 (0.00)	0.06 (0.09)	0.06 (0.09)	0.18 (0.07)	0.23 (0.07)	0.26 (0.05)	0.25 (0.06)	**0.27 (0.06)**
Test Spearman r	nan (nan)	0.07 (0.09)	0.07 (0.09)	0.17 (0.07)	0.23 (0.07)	**0.27 (0.06)**	0.25 (0.07)	**0.27 (0.06)**
Demo + Sleep + Brain (shuffle)								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	\	\	\	\	\	\	12.51 (0.48)	**12.43 (0.48)**
Test RMSE	\	\	\	\	\	\	15.56 (0.60)	**15.45 (0.61)**
Test R2	\	\	\	\	\	\	0.02 (0.03)	**0.03 (0.02)**
Test Pearson r	\	\	\	\	\	\	0.17 (0.05)	**0.21 (0.06)**

Test Spearman r	\	\	\	\	\	\	0.18 (0.05)	**0.21 (0.06)**
Sleep + APOE								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	\	\	\	\	\	\	12.65 (0.52)	**12.60 (0.54)**
Test RMSE	\	\	\	\	\	\	15.73 (0.64)	**15.69 (0.69)**
Test R2	\	\	\	\	\	\	-0.00 (0.02)	**0.00 (0.02)**
Test Pearson r	\	\	\	\	\	\	0.10 (0.07)	**0.12 (0.07)**
Test Spearman r	\	\	\	\	\	\	0.09 (0.07)	**0.11 (0.07)**
Sleep + APOE (shuffle)								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	\	\	\	\	\	\	12.65 (0.52)	**12.58 (0.53)**
Test MSE	\	\	\	\	\	\	15.73 (0.65)	**15.69 (0.68)**
Test R2	\	\	\	\	\	\	-0.00 (0.02)	**0.00 (0.03)**
Test Pearson r	\	\	\	\	\	\	0.10 (0.08)	**0.12 (0.07)**
Test Spearman r	\	\	\	\	\	\	0.09 (0.08)	**0.11 (0.07)**
Demo + Sleep + APOE								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	\	\	\	\	\	\	12.01 (0.57)	**12.00 (0.58)**
Test MSE	\	\	\	\	\	\	15.14 (0.67)	**15.06 (0.69)**
Test R2	\	\	\	\	\	\	0.07 (0.04)	**0.08 (0.04)**
Test Pearson r	\	\	\	\	\	\	0.28 (0.07)	**0.30 (0.07)**
Test Spearman r	\	\	\	\	\	\	0.28 (0.07)	**0.29 (0.07)**
Demo + Sleep + APOE (shuffle)								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	\	\	\	\	\	\	12.02 (0.59)	**12.02 (0.57)**
Test MSE	\	\	\	\	\	\	15.15 (0.70)	**15.10 (0.69)**
Test R2	\	\	\	\	\	\	0.07 (0.04)	**0.08 (0.04)**
Test Pearson r	\	\	\	\	\	\	0.28 (0.07)	**0.29 (0.07)**
Test Spearman r	\	\	\	\	\	\	0.28 (0.08)	**0.29 (0.07)**
Demo + Sleep + Brain + APOE								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	\	\	\	\	\	\	12.21 (0.52)	**12.17 (0.49)**
Test MSE	\	\	\	\	\	\	15.29 (0.63)	**15.22 (0.63)**
Test R2	\	\	\	\	\	\	0.05 (0.03)	**0.06 (0.02)**
Test Pearson r	\	\	\	\	\	\	0.25 (0.06)	**0.27 (0.06)**
Test Spearman r	\	\	\	\	\	\	0.25 (0.06)	**0.28 (0.06)**
Demo + Sleep + Brain + APOE (shuffle)								
Metric	**Dummy**	**Linear**	**Ridge**	**SVM-linear**	**SVM-rbf**	**Random forest**	**XGBoost**	**AutoGluon**
Test MAE	\	\	\	\	\	\	12.22 (0.53)	**12.17 (0.50)**
Test MSE	\	\	\	\	\	\	15.29 (0.64)	**15.22 (0.62)**
Test R2	\	\	\	\	\	\	0.05 (0.03)	**0.06 (0.02)**
Test Pearson r	\	\	\	\	\	\	0.25 (0.05)	**0.27 (0.06)**
Test Spearman r	\	\	\	\	\	\	0.25 (0.06)	**0.27 (0.06)**

Supplementary Table 5. Clustering rules of SHAP-derived subgroups for Stroop test score prediction in Greifswald

Cluster	Rules	Precision	Recall (coverage)	Participants number
0	Age > 58.1, Sex = Female	1.00	0.98	167
1	Age <= 57.6, Sex = Male	1.00	0.99	289
2	Age > 57.6, Sex = Male	1.00	1.00	154
3	25.8 < Age <= 57.7, Sex = Female	0.99	0.96	235

Supplementary Table 6. Clustering rules of SHAP-derived subgroups for memory test score prediction in Greifswald

Cluster	Rules	Precision	Recall (coverage)	Participants number
0	47.9< Age <= 67.5, Sex = Female	1.00	0.89	240
1	47.9 < Age <= 67.4, Sex = Male	0.99	0.99	213
2	PSG Sleep Efficiency > 0.608, Depressive score <= 39.5, Age <= 47.9, Sex = Male	1.00	0.98	168
3	PSG Sleep Efficiency > 0.677, Depressive score <= 20.0, Age <= 47.8, Sex = Male	0.98	1.00	117
4	PSG Sleep Duration > 3.27, Age > 67.4, SEX = Male	1.00	1.00	58
5	PSG Sleep Duration > 5.15, Depressive score <= 11.9, Age > 67,4, SEX = Female	1.00	0.95	49

Supplement Table 7. Out-of-cohort validation of Stroop test score prediction in Rotterdam

Model	Feature Combination	Performance Test Set MAE	Performance Test Set RMSE	Performance Test Set R2	Performance Test Set Pearson r	Performance Test Set Spearman r
AutoGluon	Sleep	8.82	11.05	-0.14	0.15****	0.07
	Demographic	9.80	11.97	-0.34	0.27****	0.24****
	Sleep, Demo	9.82	11.92	-0.33	0.29****	0.26****
	Sleep (Shuffled), Demo	10.00	12.21	-0.40	0.28****	0.23****
	Brain	10.04	12.12	-0.37	0.26****	0.18****
	Sleep, Demo, Brain	10.78	12.72	-0.51	0.34****	0.26****
	Sleep, Demo, Brain (Shuffled)	9.71	11.82	-0.31	0.23****	0.21****
	Sleep, APOE	8.86	11.07	-0.15	0.16****	0.14***
	Sleep, APOE (Shuffled)	8.94	11.17	-0.17	0.14***	0.15****
	Sleep, Demo, APOE	9.83	11.90	-0.33	0.30****	0.26****
	Sleep, Demo, APOE (Shuffled)	9.92	11.99	-0.34	0.31****	0.27****
	Sleep, Demo, Brain, APOE	10.62	12.59	-0.48	0.33****	0.26****
	Sleep, Demo, Brain, APOE (Shuffled)	10.69	12.66	-0.50	0.33****	0.26****
XGBoost	Sleep	8.95	11.16	-0.17	0.17****	0.16****
	Demographic	9.49	11.66	-0.27	0.24****	0.23****
	Sleep, Demo	9.31	11.45	-0.23	0.24****	0.21****
	Sleep (Shuffled), Demo	9.61	11.82	-0.31	0.15****	0.14***
	Brain	10.00	12.16	-0.38	0.20****	0.18****
	Sleep, Demo, Brain	10.17	12.22	-0.40	0.28****	0.25****
	Sleep, Demo, Brain (Shuffled)	10.00	12.10	-0.37	0.21****	0.19****
	Sleep, APOE	8.81	11.07	-0.15	0.17****	0.15****

	Sleep, APOE (Shuffled)	8.89	11.09	-0.15	0.22****	0.19****
	Sleep, Demo, APOE	9.56	11.70	-0.28	0.18****	0.16****
	Sleep, Demo, APOE (Shuffled)	9.80	11.89	-0.32	0.24****	0.22****
	Sleep, Demo, Brain, APOE	10.67	12.74	-0.52	0.24****	0.23****
	Sleep, Demo, Brain, APOE (Shuffled)	9.90	12.02	-0.35	0.19****	0.19****

Supplement Table 8. Out-of-cohort validation of Stroop test score prediction in San Diego

Model	Feature Combination	Performance Test Set MAE	Performance Test Set RMSE	Performance Test Set R2	Performance Test Set Pearson r	Performance Test Set Spearman r
AutoGluon	Sleep	20.19	21.36	-8.96	-0.05	-0.05
	Demographic	12.46	14.21	-3.41	0.11*	0.11*
	Sleep, Demo	12.04	13.80	-3.16	0.08	0.10*
	Sleep (Shuffled), Demo	13.31	14.90	-3.85	0.13**	0.14**
	Brain	14.76	16.22	-4.74	0.19****	0.18****
	Sleep, Demo, Brain	13.34	14.85	-3.82	0.15***	0.16***
	Sleep, Demo, Brain (Shuffled)	14.88	16.32	-4.82	0.06	0.08
	Sleep, APOE	19.26	20.53	-8.21	-0.03	-0.02
	Sleep, APOE (Shuffled)	18.90	20.21	-7.92	-0.05	-0.03
	Sleep, Demo, APOE	13.14	14.76	-3.76	0.08	0.10*
	Sleep, Demo, APOE (Shuffled)	12.70	14.27	-3.45	0.10*	0.11*
	Sleep, Demo, Brain, APOE	13.30	14.82	-3.80	0.15***	0.16***
	Sleep, Demo, Brain, APOE (Shuffled)	13.01	14.56	-3.63	0.14**	0.15***
XGBoost	Sleep	20.53	21.62	-9.20	0.03	0.03
	Demographic	15.45	16.83	-5.19	0.15***	0.14**

	Sleep, Demo	17.75	18.97	-6.86	0.15***	0.15***
	Sleep (Shuffled), Demo	17.68	18.94	-6.83	0.13**	0.13**
	Brain	14.77	16.27	-4.78	0.14**	0.15***
	Sleep, Demo, Brain	14.63	16.08	-4.64	0.14**	0.14**
	Sleep, Demo, Brain (Shuffled)	15.01	16.44	-4.90	0.04	0.02
	Sleep, APOE	20.13	21.25	-8.86	-0.05	-0.06
	Sleep, APOE (Shuffled)	20.43	21.54	-9.13	-0.03	-0.02
	Sleep, Demo, APOE	17.32	18.61	-6.56	0.05	0.04
	Sleep, Demo, APOE (Shuffled)	15.89	17.29	-5.53	0.06	0.08
	Sleep, Demo, Brain, APOE	13.80	15.33	-4.13	0.12**	0.13**
	Sleep, Demo, Brain, APOE (Shuffled)	15.81	17.24	-5.49	0.02	0.02

Supplement Table 9. Out-of-cohort validation of digit-based memory test score prediction in San Diego

Model	Feature Combination	Performance Test Set MAE	Performance Test Set RMSE	Performance Test Set R2	Performance Test Set Pearson r	Performance Test Set Spearman r
AutoGluon	Sleep	16.86	19.50	-1.84	0.00	-0.06
	Demographic	12.35	14.73	-0.62	0.03	0.02
	Sleep, Demo	11.50	13.85	-0.44	0.03	0.03
	Sleep (Shuffled), Demo	12.70	15.30	-0.75	-0.07	-0.06
	Brain	15.25	17.83	-1.38	0.05	0.05
	Sleep, Demo, Brain	13.93	16.44	-1.02	0.01	0.01
	Sleep, Demo, Brain (Shuffled)	15.16	17.76	-1.36	-0.03	-0.03
	Sleep, APOE	16.75	19.41	-1.82	-0.02	-0.07
	Sleep, APOE (Shuffled)	16.55	19.20	-1.76	-0.03	-0.06

	Sleep, Demo, APOE	10.84	13.12	-0.29	0.02	0.03
	Sleep, Demo, APOE (Shuffled)	11.42	13.72	-0.41	0.01	0.01
	Sleep, Demo, Brain, APOE	13.23	15.74	-0.85	0.03	0.02
	Sleep, Demo, Brain, APOE (Shuffled)	13.71	16.22	-0.97	0.01	0.02
XGBoost	Sleep	17.72	20.37	-2.11	-0.03	-0.01
	Demographic	12.10	14.45	-0.56	0.06	0.05
	Sleep, Demo	11.63	14.04	-0.47	0.05	0.05
	Sleep (Shuffled), Demo	19.51	22.10	-2.66	0.01	0.02
	Brain	13.49	15.94	-0.90	0.03	0.03
	Sleep, Demo, Brain	12.92	15.44	-0.78	0.02	0.01
	Sleep, Demo, Brain (Shuffled)	16.61	19.26	-1.78	-0.06	-0.05
	Sleep, APOE	18.64	21.26	-2.38	-0.04	-0.02
	Sleep, APOE (Shuffled)	18.95	21.56	-2.48	0.02	0.01
	Sleep, Demo, APOE	14.31	16.88	-1.13	0.02	0.02
	Sleep, Demo, APOE (Shuffled)	11.89	14.34	-0.54	0.04	0.04
	Sleep, Demo, Brain, APOE	13.24	15.74	-0.85	-0.03	-0.03
	Sleep, Demo, Brain, APOE (Shuffled)	13.86	16.41	-1.01	0.05	0.05

Supplement Table 10. Out-of-cohort validation of letter-based memory test score prediction in San Diego

Model	Feature Combination	Performance Test Set MAE	Performance Test Set RMSE	Performance Test Set R2	Performance Test Set Pearson r	Performance Test Set Spearman r
AutoGluon	Sleep	26.38	28.54	-5.73	0.08	0.04
	Demographic	20.01	22.64	-3.23	0.03	0.02
	Sleep, Demo	18.45	21.17	-2.71	0.03	0.03

	Sleep (Shuffled), Demo	20.41	23.19	-3.45	-0.09*	-0.05
	Brain	24.26	26.61	-4.85	0.08	0.06
	Sleep, Demo, Brain	22.44	24.89	-4.12	0.04	0.04
	Sleep, Demo, Brain (Shuffled)	24.22	26.55	-4.83	-0.03	-0.03
	Sleep, APOE	26.26	28.43	-5.68	0.06	0.03
	Sleep, APOE (Shuffled)	25.97	28.17	-5.56	0.06	0.01
	Sleep, Demo, APOE	17.00	19.83	-2.25	0.04	0.04
	Sleep, Demo, APOE (Shuffled)	18.15	20.92	-2.62	0.02	0.02
	Sleep, Demo, Brain, APOE	21.47	23.98	-3.75	0.07	0.05
	Sleep, Demo, Brain, APOE (Shuffled)	22.13	24.61	-4.01	0.05	0.04
XGBoost	Sleep	27.47	29.58	-6.23	0.01	0.06
	Demographic	19.50	22.21	-3.08	0.03	0.02
	Sleep, Demo	18.51	21.33	-2.76	0.05	0.06
	Sleep (Shuffled), Demo	29.59	31.55	-7.23	-0.04	-0.03
	Brain	21.82	24.34	-3.90	-0.01	-0.01
	Sleep, Demo, Brain	20.70	23.41	-3.53	0.03	0.02
	Sleep, Demo, Brain (Shuffled)	26.11	28.31	-5.62	-0.02	-0.01
	Sleep, APOE	28.58	30.58	-6.73	0.09*	0.08
	Sleep, APOE (Shuffled)	28.91	30.91	-6.89	0.07	0.06
	Sleep, Demo, APOE	23.09	25.50	-4.37	0.01	0.01
	Sleep, Demo, APOE (Shuffled)	18.76	21.69	-2.89	0.03	0.03
	Sleep, Demo, Brain, APOE	21.17	23.79	-3.68	0.02	0.03
	Sleep, Demo, Brain, APOE (Shuffled)	22.27	24.82	-4.09	0.06	0.07

Supplement Table 11. Out-of-cohort validation of Stroop test score prediction in Liège

Model	Feature Combination	Performance Test Set MAE	Performance Test Set RMSE	Performance Test Set R2	Performance Test Set Pearson r	Performance Test Set Spearman r
AutoGluon	Sleep	23.28	29.94	-0.64	-0.04	0.04
	Demographic	21.23	27.55	-0.39	0.33****	0.36****
	Sleep, Demo	21.40	27.65	-0.40	0.35****	0.39****
	Sleep (Shuffled), Demo	21.18	27.45	-0.38	0.33****	0.38****
	Brain	22.30	28.50	-0.49	0.14*	0.21**
	Sleep, Demo, Brain	21.62	27.62	-0.40	0.18*	0.29****
	Sleep, Demo, Brain (Shuffled)	21.95	28.16	-0.45	0.22**	0.25***
	Sleep, APOE	23.30	29.98	-0.64	-0.03	0.04
	Sleep, APOE (Shuffled)	23.08	29.76	-0.62	-0.02	0.09
	Sleep, Demo, APOE	21.88	28.12	-0.45	0.31****	0.35****
	Sleep, Demo, APOE (Shuffled)	21.42	27.58	-0.39	0.39****	0.41****
	Sleep, Demo, Brain, APOE	21.49	27.54	-0.39	0.21**	0.28****
	Sleep, Demo, Brain, APOE (Shuffled)	21.44	27.47	-0.38	0.19**	0.27***
XGBoost	Sleep	23.30	29.86	-0.63	0.07	0.09
	Demographic	22.35	28.62	-0.50	0.28***	0.31****
	Sleep, Demo	22.84	29.16	-0.56	0.27***	0.29****
	Sleep (Shuffled), Demo	22.78	29.07	-0.55	0.20**	0.23**
	Brain	21.93	28.18	-0.45	0.17*	0.20**
	Sleep, Demo, Brain	21.71	27.92	-0.43	0.21**	0.26***
	Sleep, Demo, Brain (Shuffled)	21.86	28.07	-0.44	0.19**	0.19*

	Sleep, APOE	23.38	29.96	-0.64	0.08	0.05
	Sleep, APOE (Shuffled)	23.38	29.91	-0.64	0.04	0.01
	Sleep, Demo, APOE	22.35	28.48	-0.48	0.33****	0.32****
	Sleep, Demo, APOE (Shuffled)	21.97	28.09	-0.44	0.32****	0.34****
	Sleep, Demo, Brain, APOE	21.46	27.61	-0.39	0.15*	0.16*
	Sleep, Demo, Brain, APOE (Shuffled)	22.20	28.58	-0.49	0.17*	0.20**

Supplement Table 12. Out-of-cohort validation of memory test score prediction in Liège

Model	Feature Combination	Performance Test Set MAE	Performance Test Set RMSE	Performance Test Set R2	Performance Test Set Pearson r	Performance Test Set Spearman r
AutoGluon	Sleep	13.96	15.68	-1.94	-0.06	-0.01
	Demographic	15.83	17.60	-2.71	-0.05	-0.05
	Sleep, Demo	16.52	18.50	-3.10	0.01	-0.00
	Sleep (Shuffled), Demo	15.47	17.14	-2.51	0.03	0.02
	Brain	15.74	17.29	-2.58	-0.08	-0.02
	Sleep, Demo, Brain	16.78	18.29	-3.00	-0.08	-0.03
	Sleep, Demo, Brain (Shuffled)	14.78	16.40	-2.22	-0.03	-0.04
	Sleep, APOE	13.76	15.57	-1.90	-0.06	-0.02
	Sleep, APOE (Shuffled)	14.16	15.92	-2.03	-0.05	-0.03
	Sleep, Demo, APOE	16.42	18.51	-3.10	0.01	-0.01
	Sleep, Demo, APOE (Shuffled)	16.04	18.12	-2.93	-0.00	-0.01
	Sleep, Demo, Brain, APOE	17.82	19.32	-3.46	-0.08	-0.03
	Sleep, Demo, Brain, APOE (Shuffled)	17.06	18.57	-3.12	-0.07	-0.01
XGBoost	Sleep	15.04	16.64	-2.31	-0.09	-0.08

	Demographic	16.07	17.84	-2.81	-0.05	-0.04
	Sleep, Demo	15.82	17.57	-2.69	0.06	0.02
	Sleep (Shuffled), Demo	12.43	14.08	-1.37	0.09	0.05
	Brain	18.47	19.91	-3.74	-0.05	-0.04
	Sleep, Demo, Brain	19.73	21.42	-4.49	-0.03	-0.03
	Sleep, Demo, Brain (Shuffled)	14.08	15.68	-1.94	0.02	0.01
	Sleep, APOE	13.36	15.02	-1.70	-0.02	0.02
	Sleep, APOE (Shuffled)	13.34	15.04	-1.71	-0.03	-0.00
	Sleep, Demo, APOE	15.15	16.79	-2.37	-0.06	-0.04
	Sleep, Demo, APOE (Shuffled)	16.09	18.03	-2.89	0.01	-0.00
	Sleep, Demo, Brain, APOE	19.12	20.60	-4.08	-0.00	0.02
	Sleep, Demo, Brain, APOE (Shuffled)	17.58	19.19	-3.41	-0.09	-0.06

Supplement Table 13. Out-of-cohort validation of memory test score prediction in Stockholm

Model	Feature Combination	Performance Test Set MAE	Performance Test Set RMSE	Performance Test Set R2	Performance Test Set Pearson r	Performance Test Set Spearman r
AutoGluon	Sleep	16.37	19.89	-0.74	-0.03	-0.05
	Demographic	14.60	17.91	-0.41	0.41**	0.34*
	Sleep, Demo	15.40	18.86	-0.57	0.42**	0.36*
	Sleep (Shuffled), Demo	15.12	18.61	-0.53	0.28	0.24
	Brain	15.97	19.54	-0.68	0.50***	0.45**
	Sleep, Demo, Brain	16.13	19.90	-0.75	0.50***	0.47***
	Sleep, Demo, Brain (Shuffled)	14.89	18.76	-0.55	0.41**	0.32*
	Sleep, APOE	17.07	20.85	-0.92	-0.08	-0.14

	Sleep, APOE (Shuffled)	16.83	20.60	-0.87	-0.04	-0.07
	Sleep, Demo, APOE	15.26	18.66	-0.53	0.41**	0.37**
	Sleep, Demo, APOE (Shuffled)	15.47	18.99	-0.59	0.37**	0.31*
	Sleep, Demo, Brain, APOE	17.02	20.68	-0.89	0.47***	0.46***
	Sleep, Demo, Brain, APOE (Shuffled)	16.41	20.19	-0.80	0.47***	0.44**
XGBoost	Sleep	15.43	18.29	-0.47	0.06	0.10
	Demographic	15.07	18.81	-0.56	0.40**	0.34*
	Sleep, Demo	15.07	18.91	-0.58	0.40**	0.35*
	Sleep (Shuffled), Demo	15.46	18.30	-0.48	-0.20	-0.19
	Brain	18.71	22.57	-1.25	0.26	0.29*
	Sleep, Demo, Brain	19.95	23.61	-1.46	0.24	0.26
	Sleep, Demo, Brain (Shuffled)	14.75	18.28	-0.47	0.42**	0.38**
	Sleep, APOE	15.59	18.47	-0.50	0.05	0.02
	Sleep, APOE (Shuffled)	15.74	18.75	-0.55	-0.00	-0.03
	Sleep, Demo, APOE	14.92	18.44	-0.50	0.41**	0.36*
	Sleep, Demo, APOE (Shuffled)	14.17	18.06	-0.44	0.41**	0.37**
	Sleep, Demo, Brain, APOE	18.53	22.25	-1.18	0.42**	0.40**
	Sleep, Demo, Brain, APOE (Shuffled)	16.74	20.72	-0.89	0.32*	0.30*

Supplement Table 14. Out-of-cohort validation of executive function score prediction in Pittsburgh

Model	Feature Combination	Performance Test Set MAE	Performance Test Set RMSE	Performance Test Set R2	Performance Test Set Pearson r	Performance Test Set Spearman r
AutoGluon	Sleep	4.49	5.35	-0.70	0.05	0.14
	Demographic	5.05	6.43	-1.45	0.04	0.17

	Sleep, Demo	3.87	5.35	-0.69	-0.32*	-0.30
	Sleep (Shuffled), Demo	3.61	5.14	-0.56	-0.06	-0.09
	Brain	6.00	6.94	-1.85	-0.11	-0.06
	Sleep, Demo, Brain	4.43	5.04	-0.50	-0.03	-0.05
	Sleep, Demo, Brain (Shuffled)	3.47	4.20	-0.05	0.01	0.07
	Sleep, APOE	3.89	5.02	-0.49	0.03	0.16
	Sleep, APOE (Shuffled)	4.38	5.50	-0.79	-0.05	0.05
	Sleep, Demo, APOE	3.47	4.70	-0.31	-0.10	-0.13
	Sleep, Demo, APOE (Shuffled)	3.52	4.64	-0.27	-0.20	-0.24
	Sleep, Demo, Brain, APOE	4.39	4.96	-0.46	-0.03	-0.00
	Sleep, Demo, Brain, APOE (Shuffled)	4.56	5.20	-0.60	-0.03	-0.06
XGBoost	Sleep	5.06	5.82	-1.00	0.07	0.08
	Demographic	3.62	5.07	-0.52	0.03	-0.02
	Sleep, Demo	3.42	4.27	-0.08	-0.01	0.06
	Sleep (Shuffled), Demo	3.46	4.72	-0.32	0.00	0.01
	Brain	5.71	6.65	-1.62	-0.01	-0.00
	Sleep, Demo, Brain	4.47	5.10	-0.54	-0.05	-0.03
	Sleep, Demo, Brain (Shuffled)	3.49	4.17	-0.03	0.02	0.08
	Sleep, APOE	4.30	4.85	-0.39	0.08	0.10
	Sleep, APOE (Shuffled)	4.76	5.45	-0.76	0.12	0.10
	Sleep, Demo, APOE	3.51	4.22	-0.05	0.03	0.13
	Sleep, Demo, APOE (Shuffled)	3.51	4.41	-0.15	-0.11	-0.05
	Sleep, Demo, Brain, APOE	4.51	5.18	-0.59	-0.02	-0.09

	Sleep, Demo, Brain, APOE (Shuffled)	4.26	4.99	-0.48	-0.20	-0.29

Supplement Table 15. Out-of-cohort validation of spatial memory test score prediction in Pittsburgh

Model	Feature Combination	Performance Test Set MAE	Performance Test Set RMSE	Performance Test Set R2	Performance Test Set Pearson r	Performance Test Set Spearman r
AutoGluon	Sleep	11.91	13.93	-0.00	0.17	0.20
	Demographic	11.96	14.38	-0.07	0.25	0.25
	Sleep, Demo	12.01	14.48	-0.08	0.08	0.08
	Sleep (Shuffled), Demo	11.86	14.25	-0.05	0.28	0.19
	Brain	12.51	15.30	-0.21	0.09	0.07
	Sleep, Demo, Brain	12.00	14.30	-0.06	0.09	0.04
	Sleep, Demo, Brain (Shuffled)	11.66	13.87	0.01	0.09	0.13
	Sleep, APOE	12.00	14.02	-0.01	0.11	0.17
	Sleep, APOE (Shuffled)	12.14	14.10	-0.03	0.22	0.25
	Sleep, Demo, APOE	12.56	15.32	-0.21	0.04	0.08
	Sleep, Demo, APOE (Shuffled)	12.02	14.61	-0.10	0.07	0.12
	Sleep, Demo, Brain, APOE	12.13	14.50	-0.09	0.14	0.11
	Sleep, Demo, Brain, APOE (Shuffled)	12.15	14.43	-0.08	0.07	0.02
XGBoost	Sleep	12.21	14.30	-0.06	0.13	0.06
	Demographic	11.68	13.96	-0.01	0.17	0.19
	Sleep, Demo	11.82	14.24	-0.05	0.10	0.09
	Sleep (Shuffled), Demo	11.64	13.96	-0.01	0.17	0.24
	Brain	13.70	16.87	-0.47	0.04	0.05
	Sleep, Demo, Brain	13.91	17.16	-0.52	0.06	0.08

	Sleep, Demo, Brain (Shuffled)	11.60	13.91	0.00	0.05	0.06
	Sleep, APOE	12.06	14.10	-0.03	0.09	0.12
	Sleep, APOE (Shuffled)	12.44	14.18	-0.04	0.05	0.03
	Sleep, Demo, APOE	11.70	13.82	0.01	0.12	0.05
	Sleep, Demo, APOE (Shuffled)	11.59	14.20	-0.04	0.26	0.25
	Sleep, Demo, Brain, APOE	12.72	15.15	-0.19	-0.06	-0.10
	Sleep, Demo, Brain, APOE (Shuffled)	12.04	14.28	-0.05	0.22	0.19

Supplement Table 16. Out-of-cohort validation of letter-based memory test score prediction in Pittsburgh

Model	Feature Combination	Performance Test Set MAE	Performance Test Set RMSE	Performance Test Set R2	Performance Test Set Pearson r	Performance Test Set Spearman r
AutoGluon	Sleep	20.44	22.52	-4.79	0.07	0.06
	Demographic	27.46	29.10	-8.67	-0.14	-0.14
	Sleep, Demo	26.61	28.48	-8.26	-0.14	-0.09
	Sleep (Shuffled), Demo	27.35	28.88	-8.52	0.15	0.06
	Brain	16.75	18.98	-3.11	0.09	0.10
	Sleep, Demo, Brain	19.37	21.42	-4.24	0.17	0.20
	Sleep, Demo, Brain (Shuffled)	22.95	24.81	-6.03	0.00	0.03
	Sleep, APOE	21.67	23.80	-5.47	-0.03	0.07
	Sleep, APOE (Shuffled)	19.12	21.30	-4.18	0.16	0.14
	Sleep, Demo, APOE	28.76	30.54	-9.65	-0.15	-0.08
	Sleep, Demo, APOE (Shuffled)	26.75	28.73	-8.43	-0.18	-0.12
	Sleep, Demo, Brain, APOE	18.48	20.56	-3.83	0.18	0.18
	Sleep, Demo, Brain, APOE (Shuffled)	19.01	21.12	-4.09	0.15	0.13

XGBoost	Sleep	19.17	21.25	-4.16	0.15	0.16
	Demographic	25.42	27.19	-7.45	-0.10	-0.07
	Sleep, Demo	25.85	27.71	-7.77	-0.09	-0.08
	Sleep (Shuffled), Demo	21.00	22.97	-5.02	0.14	0.16
	Brain	14.10	16.36	-2.06	0.01	0.06
	Sleep, Demo, Brain	13.88	16.19	-1.99	0.16	0.17
	Sleep, Demo, Brain (Shuffled)	22.25	24.22	-5.70	-0.22	-0.19
	Sleep, APOE	20.27	22.34	-4.70	0.04	0.09
	Sleep, APOE (Shuffled)	21.71	23.83	-5.49	-0.01	0.01
	Sleep, Demo, APOE	23.20	25.13	-6.21	-0.09	-0.13
	Sleep, Demo, APOE (Shuffled)	27.44	29.19	-8.73	-0.00	0.02
	Sleep, Demo, Brain, APOE	17.90	20.24	-3.68	0.01	0.01
	Sleep, Demo, Brain, APOE (Shuffled)	18.47	20.66	-3.88	0.16	0.14

Supplement Table 17. Out-of-cohort validation of spatial memory test score prediction in Jülich session 1

Model	Feature Combination	Performance Test Set MAE	Performance Test Set RMSE	Performance Test Set R2	Performance Test Set Pearson r	Performance Test Set Spearman r
AutoGluon	Sleep	20.70	21.56	-13.94	-0.11	-0.21
	Demographic	16.37	17.25	-8.57	0.24	0.20
	Sleep, Demo	14.61	15.68	-6.91	0.11	0.11
	Sleep (Shuffled), Demo	16.19	17.14	-8.44	0.13	0.09
	Brain	24.02	24.70	-18.62	0.18	0.19
	Sleep, Demo, Brain	22.89	23.53	-16.81	0.26	0.26
	Sleep, Demo, Brain (Shuffled)	19.65	20.41	-12.40	0.14	0.16

	Sleep, APOE	19.39	20.39	-12.36	-0.12	-0.20
	Sleep, APOE (Shuffled)	20.57	21.38	-13.70	-0.12	-0.20
	Sleep, Demo, APOE	12.47	13.67	-5.01	0.18	0.24
	Sleep, Demo, APOE (Shuffled)	14.96	15.88	-7.11	0.33	0.29
	Sleep, Demo, Brain, APOE	25.26	26.02	-20.76	0.29	0.27
	Sleep, Demo, Brain, APOE (Shuffled)	24.00	24.71	-18.63	0.27	0.29
XGBoost	Sleep	21.59	22.30	-14.99	-0.00	-0.13
	Demographic	17.30	18.18	-9.63	0.14	0.07
	Sleep, Demo	14.53	15.55	-6.77	0.18	0.14
	Sleep (Shuffled), Demo	18.27	19.17	-10.81	-0.21	-0.26
	Brain	23.20	23.91	-17.38	0.09	0.20
	Sleep, Demo, Brain	25.57	26.52	-21.62	0.21	0.08
	Sleep, Demo, Brain (Shuffled)	20.49	21.26	-13.54	0.10	0.14
	Sleep, APOE	20.64	21.41	-13.73	-0.10	-0.12
	Sleep, APOE (Shuffled)	21.17	21.96	-14.50	-0.20	-0.20
	Sleep, Demo, APOE	17.15	18.01	-9.43	0.18	0.22
	Sleep, Demo, APOE (Shuffled)	13.38	14.45	-5.72	0.23	0.23
	Sleep, Demo, Brain, APOE	21.84	22.61	-15.44	0.15	0.24
	Sleep, Demo, Brain, APOE (Shuffled)	21.98	22.61	-15.43	0.32	0.28

Supplement Table 18. Out-of-cohort validation of letter-based memory test score prediction in Jülich session 1

Model	Feature Combination	Performance Test Set MAE	Performance Test Set RMSE	Performance Test Set R2	Performance Test Set Pearson r	Performance Test Set Spearman r
AutoGluon	Sleep	22.49	23.30	-17.49	-0.26	-0.32

	Demographic	18.19	18.93	-11.20	0.22	0.13
	Sleep, Demo	16.58	17.36	-9.26	0.04	0.02
	Sleep (Shuffled), Demo	18.06	18.73	-10.95	0.25	0.14
	Brain	25.81	26.41	-22.75	0.19	0.18
	Sleep, Demo, Brain	24.67	25.28	-20.75	0.21	0.15
	Sleep, Demo, Brain (Shuffled)	21.45	22.12	-15.66	0.05	0.07
	Sleep, APOE	21.27	22.14	-15.69	-0.26	-0.31
	Sleep, APOE (Shuffled)	22.36	23.11	-17.19	-0.28	-0.32
	Sleep, Demo, APOE	14.52	15.32	-6.99	0.11	0.14
	Sleep, Demo, APOE (Shuffled)	16.83	17.58	-9.52	0.26	0.21
	Sleep, Demo, Brain, APOE	27.05	27.78	-25.28	0.24	0.16
	Sleep, Demo, Brain, APOE (Shuffled)	25.79	26.46	-22.84	0.22	0.16
XGBoost	Sleep	23.37	24.04	-18.68	-0.29	-0.38*
	Demographic	19.09	19.86	-12.44	0.10	-0.02
	Sleep, Demo	16.52	17.22	-9.10	0.11	0.12
	Sleep (Shuffled), Demo	20.11	20.81	-13.74	-0.09	-0.11
	Brain	24.98	25.60	-21.31	0.14	0.17
	Sleep, Demo, Brain	27.36	28.24	-26.16	0.20	-0.01
	Sleep, Demo, Brain (Shuffled)	22.28	22.98	-16.99	0.04	0.02
	Sleep, APOE	22.43	23.14	-17.23	-0.34*	-0.33*
	Sleep, APOE (Shuffled)	22.96	23.67	-18.08	-0.32	-0.31
	Sleep, Demo, APOE	19.00	19.71	-12.23	0.09	0.15
	Sleep, Demo, APOE (Shuffled)	15.36	16.08	-7.80	0.23	0.12

	Sleep, Demo, Brain, APOE	23.63	24.31	-19.12	0.16	0.17
	Sleep, Demo, Brain, APOE (Shuffled)	23.77	24.35	-19.19	0.23	0.20

Supplement Table 19. Out-of-cohort validation of spatial memory test score prediction in Jülich session 2

Model	Feature Combination	Performance Test Set MAE	Performance Test Set RMSE	Performance Test Set R2	Performance Test Set Pearson r	Performance Test Set Spearman r
AutoGluon	Sleep	21.19	21.90	-15.95	-0.19	-0.33*
	Demographic	17.63	18.35	-10.90	0.29	0.22
	Sleep, Demo	15.65	16.53	-8.65	0.13	0.09
	Sleep (Shuffled), Demo	16.92	17.65	-10.01	0.33*	0.28
	Brain	25.27	25.97	-22.85	-0.00	0.08
	Sleep, Demo, Brain	24.06	24.72	-20.60	0.09	0.16
	Sleep, Demo, Brain (Shuffled)	20.73	21.37	-15.14	0.24	0.24
	Sleep, APOE	19.53	20.34	-13.63	-0.16	-0.33*
	Sleep, APOE (Shuffled)	21.46	22.16	-16.35	-0.26	-0.33*
	Sleep, Demo, APOE	13.82	14.87	-6.82	0.09	0.11
	Sleep, Demo, APOE (Shuffled)	16.90	17.82	-10.22	0.09	0.12
	Sleep, Demo, Brain, APOE	26.44	27.31	-25.35	0.10	0.19
	Sleep, Demo, Brain, APOE (Shuffled)	25.20	25.97	-22.85	0.08	0.17
XGBoost	Sleep	22.66	23.34	-18.25	-0.38*	-0.31
	Demographic	18.56	19.30	-12.17	0.16	0.05
	Sleep, Demo	15.78	16.69	-8.85	0.00	0.06
	Sleep (Shuffled), Demo	19.32	20.02	-13.17	0.14	0.08
	Brain	24.45	25.13	-21.32	-0.03	0.14

	Sleep, Demo, Brain	26.65	27.68	-26.08	0.03	-0.01
	Sleep, Demo, Brain (Shuffled)	21.26	21.90	-15.95	0.16	0.04
	Sleep, APOE	21.66	22.33	-16.62	-0.18	-0.19
	Sleep, APOE (Shuffled)	22.23	22.92	-17.57	-0.21	-0.28
	Sleep, Demo, APOE	18.25	19.03	-11.80	0.03	0.11
	Sleep, Demo, APOE (Shuffled)	14.78	15.66	-7.67	0.24	0.16
	Sleep, Demo, Brain, APOE	23.05	23.84	-19.08	-0.02	0.10
	Sleep, Demo, Brain, APOE (Shuffled)	23.18	23.80	-19.01	0.13	0.31

Supplement Table 20. Out-of-cohort validation of letter-based memory test score prediction in Jülich session 2

Model	Feature Combination	Performance Test Set MAE	Performance Test Set RMSE	Performance Test Set R2	Performance Test Set Pearson r	Performance Test Set Spearman r
AutoGluon	Sleep	21.72	22.43	-16.96	-0.32	-0.36*
	Demographic	18.16	18.88	-11.72	0.24	0.22
	Sleep, Demo	16.32	17.00	-9.31	0.18	0.14
	Sleep (Shuffled), Demo	17.58	18.19	-10.80	0.27	0.30
	Brain	25.81	26.43	-23.93	0.10	0.13
	Sleep, Demo, Brain	24.59	25.21	-21.69	0.13	0.12
	Sleep, Demo, Brain (Shuffled)	21.26	21.90	-16.12	0.09	0.18
	Sleep, APOE	20.16	20.89	-14.58	-0.30	-0.37*
	Sleep, APOE (Shuffled)	22.00	22.68	-17.36	-0.39*	-0.38*
	Sleep, Demo, APOE	14.57	15.31	-7.36	0.17	0.17
	Sleep, Demo, APOE (Shuffled)	17.51	18.29	-10.93	0.13	0.15
	Sleep, Demo, Brain, APOE	26.97	27.78	-26.53	0.14	0.14

	Sleep, Demo, Brain, APOE (Shuffled)	25.73	26.45	-23.97	0.12	0.11
XGBoost	Sleep	23.20	23.85	-19.30	-0.41*	-0.28
	Demographic	19.09	19.83	-13.03	0.10	0.03
	Sleep, Demo	16.45	17.13	-9.47	0.18	0.17
	Sleep (Shuffled), Demo	19.90	20.57	-14.09	-0.05	-0.02
	Brain	24.98	25.59	-22.37	0.09	0.20
	Sleep, Demo, Brain	27.18	28.11	-27.19	0.11	0.02
	Sleep, Demo, Brain (Shuffled)	21.79	22.38	-16.86	0.28	0.27
	Sleep, APOE	22.19	22.84	-17.62	-0.20	-0.20
	Sleep, APOE (Shuffled)	22.77	23.45	-18.63	-0.30	-0.27
	Sleep, Demo, APOE	18.79	19.49	-12.55	0.21	0.22
	Sleep, Demo, APOE (Shuffled)	15.45	16.15	-8.30	0.26	0.20
	Sleep, Demo, Brain, APOE	23.59	24.31	-20.09	0.04	0.12
	Sleep, Demo, Brain, APOE (Shuffled)	23.72	24.31	-20.09	0.13	0.27

Supplement Table 21. Out-of-cohort validation of spatial memory test score prediction in Jülich sleep deprivation session

Model	Feature Combination	Performance Test Set MAE	Performance Test Set RMSE	Performance Test Set R2	Performance Test Set Pearson r	Performance Test Set Spearman r
AutoGluon	Sleep	13.57	15.40	-1.88	-0.08	-0.00
	Demographic	10.62	12.23	-0.82	0.30	0.27
	Sleep, Demo	9.69	11.10	-0.50	0.22	0.13
	Sleep (Shuffled), Demo	10.42	12.03	-0.76	0.16	0.12
	Brain	16.46	18.66	-3.23	0.12	0.17
	Sleep, Demo, Brain	15.34	17.46	-2.71	0.22	0.27

	Sleep, Demo, Brain (Shuffled)	12.96	14.71	-1.63	0.30	0.35*
	Sleep, APOE	12.48	14.25	-1.47	-0.09	-0.00
	Sleep, APOE (Shuffled)	13.71	15.53	-1.93	0.01	0.02
	Sleep, Demo, APOE	8.79	10.10	-0.24	0.20	0.17
	Sleep, Demo, APOE (Shuffled)	10.10	11.61	-0.64	0.37*	0.38*
	Sleep, Demo, Brain, APOE	17.59	19.74	-3.74	0.23	0.28
	Sleep, Demo, Brain, APOE (Shuffled)	16.37	18.55	-3.18	0.21	0.26
XGBoost	Sleep	14.65	16.57	-2.34	0.26	0.22
	Demographic	11.43	13.03	-1.06	0.21	0.17
	Sleep, Demo	9.80	11.08	-0.49	0.24	0.26
	Sleep (Shuffled), Demo	12.01	13.67	-1.27	-0.02	-0.08
	Brain	15.76	17.96	-2.92	0.06	0.18
	Sleep, Demo, Brain	18.04	20.19	-3.96	0.19	0.17
	Sleep, Demo, Brain (Shuffled)	13.28	15.26	-1.83	0.07	0.01
	Sleep, APOE	13.85	15.71	-2.00	0.25	0.19
	Sleep, APOE (Shuffled)	14.45	16.35	-2.25	0.10	0.18
	Sleep, Demo, APOE	11.30	12.81	-1.00	0.11	0.10
	Sleep, Demo, APOE (Shuffled)	8.94	10.46	-0.33	0.26	0.23
	Sleep, Demo, Brain, APOE	14.72	16.84	-2.45	0.07	0.21
	Sleep, Demo, Brain, APOE (Shuffled)	14.63	16.70	-2.39	0.23	0.38*

Supplement Table 22. Out-of-cohort validation of letter-based memory test score prediction in Jülich sleep deprivation session

Model	Feature Combination	Performance Test Set MAE	Performance Test Set RMSE	Performance Test Set R2	Performance Test Set Pearson r	Performance Test Set Spearman r

AutoGluon	Sleep	15.93	17.44	-5.15	-0.20	-0.11
	Demographic	12.32	13.91	-2.91	0.23	0.21
	Sleep, Demo	10.65	12.28	-2.05	0.19	0.12
	Sleep (Shuffled), Demo	11.83	13.49	-2.68	0.14	0.13
	Brain	19.75	20.98	-7.90	0.18	0.19
	Sleep, Demo, Brain	18.55	19.80	-6.93	0.22	0.16
	Sleep, Demo, Brain (Shuffled)	15.27	16.70	-4.64	0.22	0.25
	Sleep, APOE	14.49	16.12	-4.26	-0.21	-0.11
	Sleep, APOE (Shuffled)	16.14	17.60	-5.26	-0.12	-0.09
	Sleep, Demo, APOE	9.10	10.82	-1.37	0.17	0.18
	Sleep, Demo, APOE (Shuffled)	11.65	13.18	-2.51	0.34*	0.36*
	Sleep, Demo, Brain, APOE	20.92	22.25	-9.01	0.23	0.19
	Sleep, Demo, Brain, APOE (Shuffled)	19.68	20.97	-7.89	0.22	0.17
XGBoost	Sleep	17.45	18.79	-6.14	0.13	0.12
	Demographic	13.21	14.80	-3.43	0.13	0.09
	Sleep, Demo	10.63	12.21	-2.01	0.29	0.30
	Sleep (Shuffled), Demo	13.85	15.45	-3.83	-0.16	-0.17
	Brain	18.93	20.18	-7.24	0.17	0.21
	Sleep, Demo, Brain	21.21	22.68	-9.40	0.20	0.13
	Sleep, Demo, Brain (Shuffled)	15.79	17.26	-5.02	0.12	0.05
	Sleep, APOE	16.41	17.83	-5.43	0.10	0.09
	Sleep, APOE (Shuffled)	17.15	18.55	-5.96	-0.01	0.08
	Sleep, Demo, APOE	12.75	14.38	-3.18	0.17	0.14

	Sleep, Demo, APOE (Shuffled)	9.74	11.40	-1.63	0.28	0.23
	Sleep, Demo, Brain, APOE	17.54	18.96	-6.27	0.14	0.16
	Sleep, Demo, Brain, APOE (Shuffled)	17.67	18.97	-6.27	0.20	0.34*