Associations between sleep, brain, and cognitive performance using the ENIGMA-Sleep data

Hanwen Bi[1,2], Robin Bülow[3], Michele Deantoni[4], David Elmenhorst[1], Eva-Maria Elmenhorst[5], Ralf Ewert[3], Fabio Ferrarelli[6], Stefan Frenzel[3], Hans J. Grabe[3,7], Sanne J.W. Hoepel[8], Felix Hoffstaedter[1,2], Neda Jahanshad[9], Ahmadreza Keihani[6], Vincent Küppers[1,10], Annemarie I. Luik[8], Ahmad Mayeli[6], Nasrin Mortazavi[4], Gustav Nilsonne[11,12], Julia S Rupp[6], Amin Saberi[13,1,2], Christina Schmidt[4], Kai Spiegelhalder[14], Sandra Tamm[11,12], Sophia I. Thomopoulos[9], Paul M. Thompson[9], Sofie Valk[13,1,2], Gilles Vandewalle[4], Henry Völzke[3,15], Antoine Weihs[3], Joseph Wexler[11,16], Katharina Wittfeld[3,7], Simon B. Eickhoff[1,2], Kaustubh R. Patil[1,2], Federico Raimondo[1,2,*], Masoud Tahmasian[1,10,2,*], for the ENIGMA-Sleep Working Group

[1]Research Center Jülich, Jülich, Germany; [2]Heinrich-Heine University Düsseldorf, Düsseldorf, Germany; [3]University Medicine Greifswald, Greifswald, Germany; [4]University of Liège, Liège, Belgium; [5]German Aerospace Center, Cologne, Germany; [6]University of Pittsburgh, Pittsburgh, United States; [7]German Centre for Neurodegenerative Diseases (DZNE), Greifswald, Germany; [8]Erasmus MC University Medical Center Rotterdam, Rotterdam, Netherlands; [9]University of Southern California, Marina del Rey, United States; [10]University of Cologne, Cologne, Germany; [11]Karolinska Institute, Stockholm, Sweden; [12]Stockholm University, Stockholm, Sweden; [13]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany; [14]University of Freiburg, Freiburg, Germany; [15]German Center for Cardiovascular Research (DZHK), Greifswald, Germany; [16]Stanford University, Stanford, United States

Background

- Sleep disturbance is considered a potential risk factor for cognitive decline and dementia[1,2].
- Large-scale studies are needed to reveal the relationship.
- Previous large-scale studies (N ≈ 40,000 & 500,000) using the UK-Biobank data highlighted a significant, non-linear relationship between sleep duration and cognitive performance but with a **small effect size**[3,4].

Gap:

- These studies were primarily cross-sectional and limited to a UK-based sample, thereby constraining the generalizability of the findings.
- Results from conventional statistical approaches are missing inter-individual variability.
- Machine learning methods enable individual-level predictions and can validate models on unseen data, thus providing a more robust analytical framework.

Aim:

In this study, we performed machine learning (ML) analysis based on both sleep duration and sleep efficiency from self-reported questionnaires and polysomnography and brain structure data using the ENIGMA-Sleep[5] data to predict cognitive scores at the individual level.

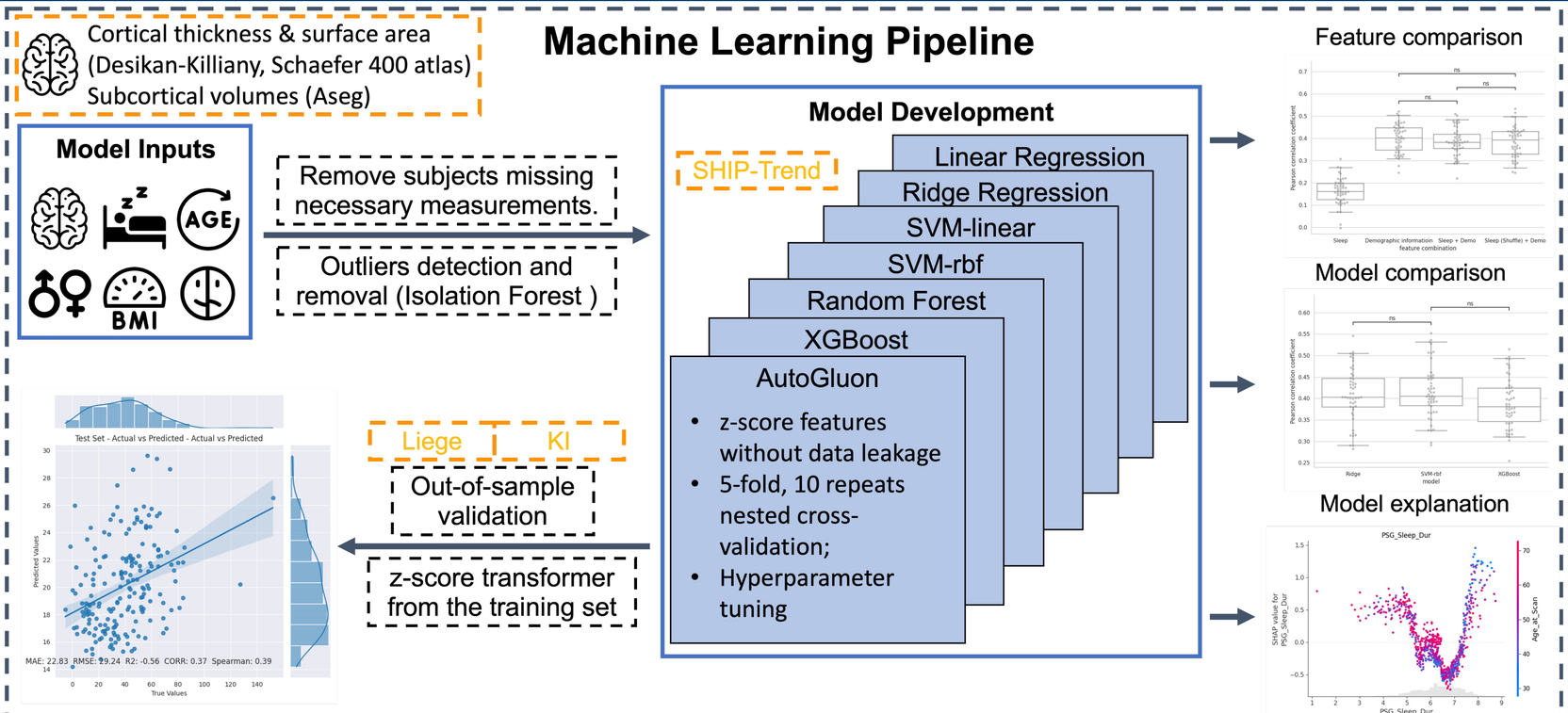
Sleep / Brain + Demographic → Cognitive Performance

Methods

ENIGMA-Sleep Datasets

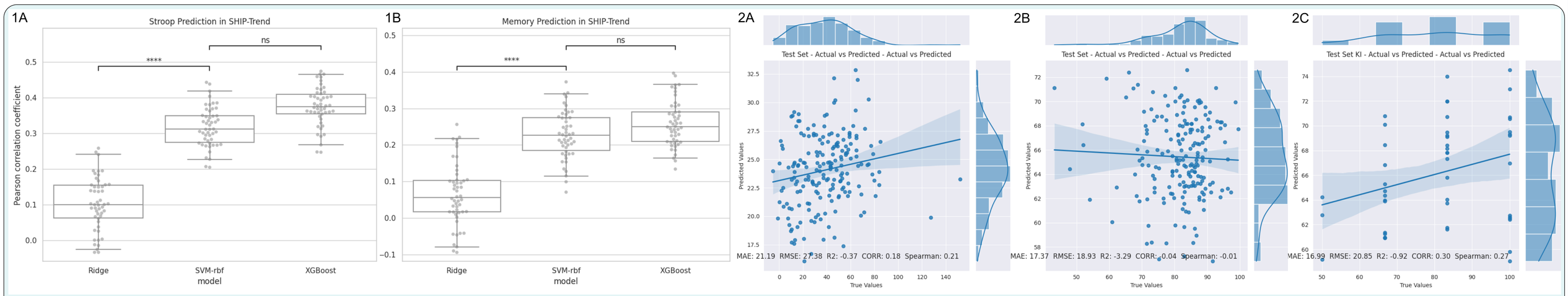
	SHIP-Trend	Liege	Karolinska
Sample Size (N)	831	192	48
Gender (F/M)	396/435	99/93	28/20
Age Range (Mean ± SD)	21-81 (52.7 ± 13.5)	50-82 (64.0 ± 7.1)	21-75 (55.4 ± 21.1)
Stroop Interference Score (second)	1-113 (20.2 ± 11.4)	-5.3-152 (38.9 ± 23.4)	\
Memory Test Accuracy (%)	69.6 ± 15.9	81.6 ± 9.2	81.2 ± 15.2

Machine Learning Pipeline



Results

Non-linear machine learning models showed Sleep + Demographic + Brain Structure can predict Stroop and memory test scores



ns: p <= 1.00e+00; *: 1.00e-02 < p <= 5.00e-02 ; **: 1.00e-03 < p <= 1.00e-02; ***: 1.00e-04 < p <= 1.00e-03; ****: p <= 1.00e-04 (same below)

Figure 1: Non-linear models showed much better prediction compared to ridge regression on SHIP-Trend. Results of prediction of Stroop (1A) and memory (1B).

Figure 2: Out-of-sample validation by XGBoost on Liege (2A: Stroop, 2B: memory) and Karolinska (2C: memory).

Model comparison by features in SHIP-Trend by XGBoost

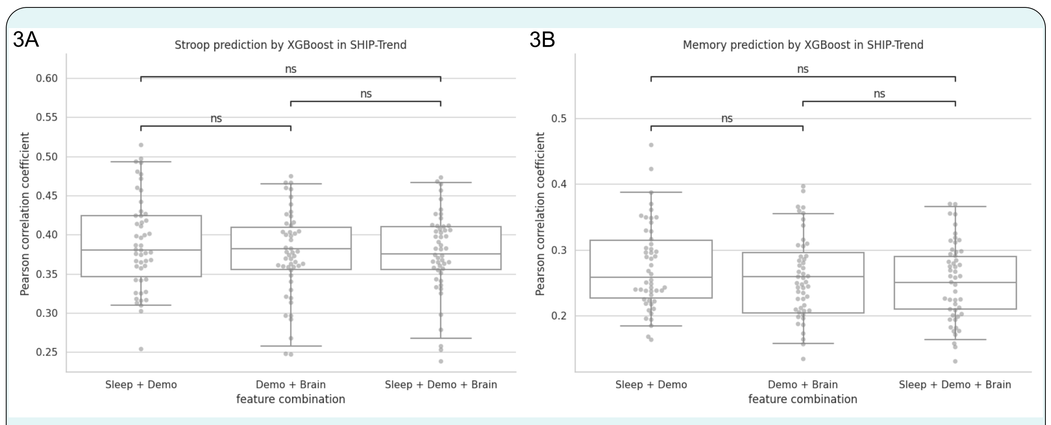


Figure 3: Both Stroop test and memory test scores can be predicted with sleep, brain structure, and demographic data by the complex non-linear model. (3A: Stroop, 3B: Memory)

Model explanation by SHAP in SHIP-Trend by XGBoost

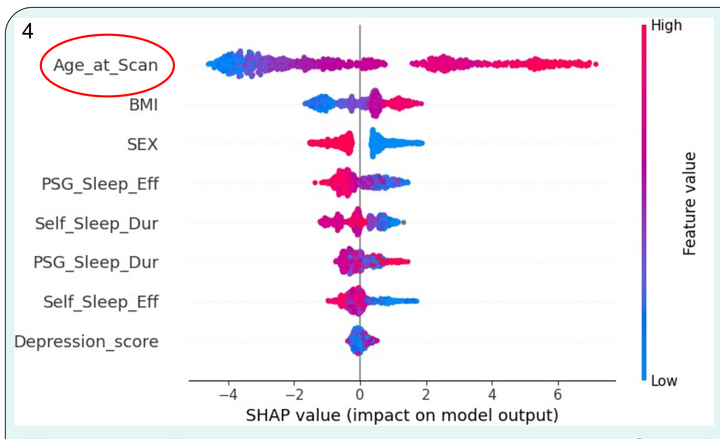


Figure 4: Feature importance based on SHAP value. (Model input: Sleep + Demographic information)

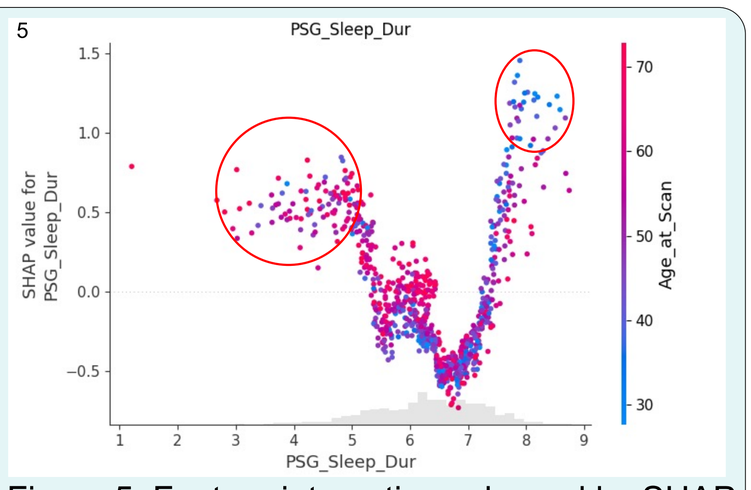


Figure 5: Feature interactions showed by SHAP value.

Conclusion

- Stroop test score can be weakly (XGBoost R2=0.13) but stably (out-of-sample validation R=0.18) predicted by sleep measurements, brain structure, and demographic data based on non-linear ML models.
- The nonlinear relationship between sleep measurements and cognitive measurements can be revealed by model explanation.
- Model explanation showed:
  - Predicted variance is driven by the COMPLEX interaction between sleep, brain, and demographic data (mainly age).
  - Older participants with shorter sleep and younger participants with longer sleep contribute mainly to the prediction of Stroop and memory test score.

References: 1. F. Emamian, et al. Front. Aging Neurosci. 8, 78 (2016); 2. R. S. Osorio, et al. Neurology 84.19, 1964-1971, 2015; 3. Y. Li, et al. Nat. Aging. 2.5 (2022); 4. S. D. Kyle, et al. Sleep Med. 38, 85-91 (2017); 5. M. Tahmasian, et al. J. Sleep Res. 30.6, e13347 (2021).

Contact: h.bi@fz-juelich.de; m.tahmasian@fz-juelich.de