Influence of sleep quality and brain volume in Alzheimer's disease

Weiping Zhang
Data Science
FHNW School of Engineering
Windisch, Switzerland
weiping.zhang@students.fhnw.ch

Antonia Bonfà
Data Science
FHNW School of Engineering
Windisch, Switzerland
antonia.bonfa@students.fhnw.ch

***Abstract*— Introduction Despite its prevalence and many promising recent developments in the study of Alzheimer's Disease (AD), the factors that contribute to AD are still poorly understood. In this study, we examine the influence of sleep disturbances and relative brain volumes on the chances of developing AD and cognitive conversion using the ADNI dataset and state-of-the-art machine learning models.**

Methods To assess sleep quality, both answers from the NPIK questionnaire as well as medical history data mentioning insomnia and/or obstructive sleep apnea (OSA) were used. Using these features as well as brain size measurements, a series of machine learning classification algorithms (logistic regression, decision tree, multiple layer perceptron (MLP) and random forest) were applied to find the best performing model to predict a conversion in cognitive diagnosis. To examine the influence of sleep disturbances markers on a model predicting AD diagnosis, we applied feature selection on the logistic regression and random forest models. An unsupervised machine learning algorithm (HDBSCAN) was used to segment MRI images of AD, MCI and NL patients.

Results The Random forest models performed best overall. The most important features for predicting a deterioration of diagnosis were caregiver distress, sleep disturbances frequency and excessive sleep during daytime. Absent caregiver distress was the most important feature for patients that improved.

This suggests that lowering the caregiver distress and/or treating sleep disorders in elderly patients might help to slow down the progression of Alzheimer's.

Keywords—Alzheimer's Disease, sleep disturbances, OSA, insomnia, MRI, machine learning, HDBSCAN

I. INTRODUCTION

Alzheimer's disease (AD) is a neuro-degenerative disease that affects over 50 million people over the age of 65 worldwide [1]. As the population grows older, the number of affected individuals is expected to rise. Alzheimer's disease is the most common form of dementia. It is characterized by progressive memory problems and loss of cognitive functions. While the underlying cause is still not well understood, some potential risk factors have been identified.

Bad sleep quality might be such a risk factor. While some decline in sleep quality can be expected with growing age, it is more prevalent in patients with AD. Bad sleep is also associated with faster cognitive decline [2]. A meta study that analyzed 12,926 papers on the topic of sleep disturbances concluded that subjects who reported sleep disturbances had a higher risk of developing AD. [3]

Given the current state of knowledge, our goal was to see if we can develop a machine learning algorithm that can utilize sleep quality assessment data from the ADNI dataset to predict changes in diagnosis. Additionally, we wanted to see if an unsupervised learning algorithm would be able to distinguish different parts of the brain.

II. DATA

A. Source

All data were taken from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database [3], a longitudinal study conducted in the US over several years. ADNI has run over several phases (ADNI1, ADNIGO, ADNI2 and ADNI3) and the data collected varies for each phase. For this paper, we made our own master file, based on several tables from the ANDI datasets. (Appendix 1). The tables were connected using the visit codes (VISCODE), roster ID (RID) and phase keys.

B. Alzheimer's disease cognitive Diagnosis

In the ADNI dataset, patients were diagnosed based on multiple cognitive tests. There are three types of diagnosis: normal (NL), mild cognitive impairment (MCI) and Alzheimer's disease (AD). Each patient was diagnosed at each visit, however some of the diagnoses were missing. Since AD diagnosis is unlikely to change in a short time, we filled in the missing data for each patient by first forward and then backward filling it. As we were interested in the potential factors that will cause or influence AD conversion, we added a feature that labeled the previous and current diagnosis.

C. Brain volume

As cognitive decline is linked with brain atrophy, we wanted to know if and how brain volume reduction is correlated with AD conversion. The ADNI database has rich numeric brain volumes data (including ventricles, hippocampus, whole brain, entorhinal, fusiform, and ICV (Total Intracranial Volume)), which are converted from MRI (Magnetic resonance imaging) scans. To avoid the affect from personal brain volume differences, only the relative brain volume data was used: for each patient, the brain volume was divided by the patient's brain volume at the first visit.

D. Biomarkers

Aβ42 and tau have been shown to have a strong correlation with AD pathologies as well as sleep disturbances. In the ADNIMERGE file, three features called abeta, tau, and ptau were available. However, as the source and exact nature of these values was unclear, we chose to exclude those features despite their potential importance.

TABLE 1. DISTRIBUTION OF INSOMNIA AND OSA

	Number of observations		
	Total	***Insomnia***	***OSA***
AD	6419	507 (7.3%)	694 (10%)
MCI	6736	870 (11.4%)	1002 (13.1%)
NL	7116	561 (7.3%)	533 (6.9%)

E. Sleep quality assessment data

To investigate the effect of sleep in AD conversion, we identified three potential topics to include in our data: insomnia, obstructive sleep apnea (OSA) and the NPI sleep questionnaire (NPI - sleep).

1) Insomnia and obstructive sleep apnea

To identify patients suffering from insomnia, both data from the medical history and health assessments were used, as well as the AXINSOMN/BCINSOMN column where available. Patients whose records mentioned insomnia were flagged as insomnia positive. For patients with multiple visits, only the observations where insomnia was detected were flagged.

The OSA status of patients was extracted from health assessment and medical history records using the same method we used for insomnia: if the medical record mentioned it, the patient was flagged as OSA positive. Since the OSA status is unlikely to change over time, the OSA marker was applied to all observations for each patient.

Table 1 shows the distribution of insomnia and OSA by diagnosis group. As we can see, both insomnia and OSA are most prevalent in the MCI group, followed by the AD group. However, the difference is not statistically significant.

2) NPI sleep questionnaire

The Neuropsychiatric Inventory Sleep Questionnaire (NPI-sleep) is used to assess abnormal sleep based on questions addressed to an informed caregiver [4]. The NPI-sleep questionnaire contains eleven questions to assess the sleep disturbances of patient. (Table 2)

TABLE 2. NPI-SLEEP QUESTIONNAIRE.

ID	*Question*
NPIK1	Difficulties in falling asleep?
NPIK2	Getting up during the night?
NPIK3	Pacing and wandering at night?
NPIK4	Awakening the caregiver during night?
NPIK5	Awaken at night or too early in the morning and thinking that it is time to start the day?
NPIK6	Waking up too early in the morning?
NPIK7	Excessive daytime napping?
NPIK8	Other abnormal nighttime behaviors?
NPIK9A	Frequency ratings?
NPIK9B	Severity ratings?
NPIK9C	Caregiver distress?

Fig. 1 shows the distribution of the total score (NPIKTOT) for patients that had replied to the NPIK sleep questionnaire (Fig 1A). This score is calculated by multiplying NPIK9A and NPIK9B. Overall, most of the scores were low, indicating that, while the patients may suffer from sleep disturbances, the majority of these were either not very frequent or not severe. This pattern is concurrent for patients with insomnia and / or OSA, as seen in Fig. 1B.

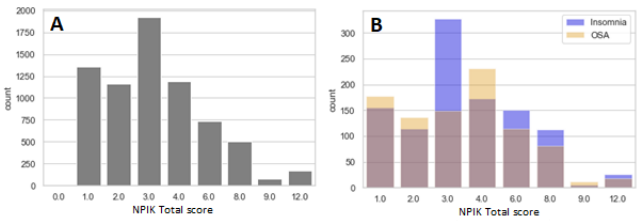


Fig. 1. Distribution of NPIK total scores

III. METHODS

A. Machine learning models for AD conversion

We designed seven binary classification models (three based on sleep assessment attributes, four based on relative brain volumes) to answer two critical questions: Will the patient's AD status get worse or not? Will the patient's AD status get better or not? (Table 2.)

For each model, we extracted the observations where all the relevant (sleep disturbances or relative brain volume) attributes are available. In the three models using sleep disturbances (Model 1 - 3), we got 7,089 observations, and 16,939 observations in the models using brain volume (Model 4 - 7). We applied under-sampling to balance the target diagnosis changes and used 10–fold cross-validation (cv) in all models instead of a one-time train-test splitting to increase the models' reliability. To select a suitable machine learning algorithm for classification and to avoid overfitting, we tried out four different machine learning algorithms with corresponding hyperparameters tuning: logistic regression (regularization strength, penalty, and solver), decision tree (tree depth), random forest (number of trees), and multiple layer perceptron (MLP, with different hidden layer settings). To improve the performance, we used three types of input data: raw data, standard scaled data, and PCA–transformed data (90% data variance is explained). Through 168 tests with different input data, algorithms and hyperparameter combinations, it turned out that random forest (between 30 and 90 trees) overall performed very stably and gave better, but not extremely outstanding, accuracy (f1-score) for all seven criteria. Using raw data, scaled data, or PCA-reduced data didn't affect the random forest performance much. Therefore, random forest model with 10-fold cross validation on raw data was applied for all seven AD conversion analyses.

B. Feature selection for predicting AD

To investigate if sleep assessment markers would improve a model which simply predicted whether a patient's cognitive diagnosis is AD or not, feature selection was used. For this approach, the same methods used for the AD conversion models were applied to the data: undersampling was used to balance the target labels and the data was then scaled and standardized.

We established a random forest classifier with 10-fold cross validation as baseline model, because this algorithm performed most stably in the model selection. In this model,

only brain volume was used. Three new models were built, using the baseline criteria and one of the sleep markers (insomnia, OSA and the NPI total score (NPIKTOT)). Lastly, a model using all features was generated.

C. *Exploring MRI scans by HDBSCAN clustering*

To better understand differences in brain structure, we applied HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise), an unsupervised machine learning algorithm (developed by Campello, Moulavi, and Sander), to detect clusters in the brain MRI scans. We reduced the 3D MRI scans to the resolution of 64 × 64 × 64 pixels, as the corresponding segmentation result is comparable to the images with higher resolution (128 × 128 × 128 pixels), but much faster. Afterwards, each 3D image had 262,144 points (3D pixels). Our goal was to optimize the clustered segmentations so that they could well represent different brain regions. This was achieved through adjusting the cluster and noise sizes by tuning the parameter min-samples (minimum number of neighbors to a core point).

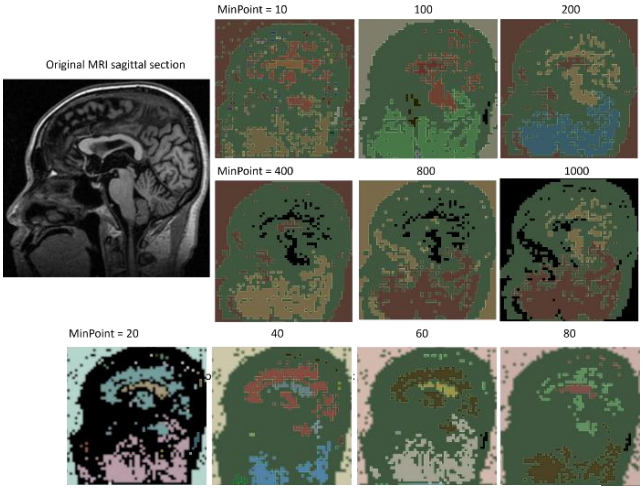


Fig. 2. Parameter tuning for HDBSCAN: Different clusters with different minimal points.

We selected the optimal min-samples based on the segmentations in both 3D and 2D (coronal, sagittal and horizontal). Here we displayed only the sagittal 2D images as representations. We started with a series of number between 10 and 1000 to narrow the range of optimal min-samples. As we could see from the Fig. 2, when increasing the value of min-samples, there will be fewer segmentations, as more points will be discarded as noise. When min-samples = 10, more brain areas were recognized, but with many small segmentations. When min-samples was set to 100 or higher, the major part of cortex was labelled as noise. With another four tests between 10 and 100 (20, 40, 60, 80), we found out the optimal value of min-samples is 60.

With the optimal parameters, we generated segmentations of MRI scans from for NL, MCI and AD patients (one of each) to see if the brain structure is different.

IV. Results

A. *Predicting diagnosis changes using sleep markers*

The accuracies of the three models were between 60.0% and 69.7%. Model 1 (Fig 3. A - D) aimed to answer the question if we could predict the AD diagnosis conversion of a normal patient. Similarly, Model 2 (Fig 3. E - H) and Model 3 (Fig 2. I - L) predicted the worsening conversion and reverse conversion of MCI patients. The feature importance of each random forest model was plotted in a boxplot (Fig 3. A, E, I). The top three important features with diagnosis changes were plotted in histograms (Fig 3. B - D, F - H, J - L) to better understand the correlation with conversion. Caregiver distress (NPIK9C, Fig 3. B, F, J) and sleep disturbances frequency (NPIK9A, Fig 3. C, G, K) appeared as the most important features in all three analyses. The third most important feature for a worsening conversion (NL-MCI/AD or MCI-AD) was excessive day sleep (NPIK7, Fig 3. D, H), whereas for the

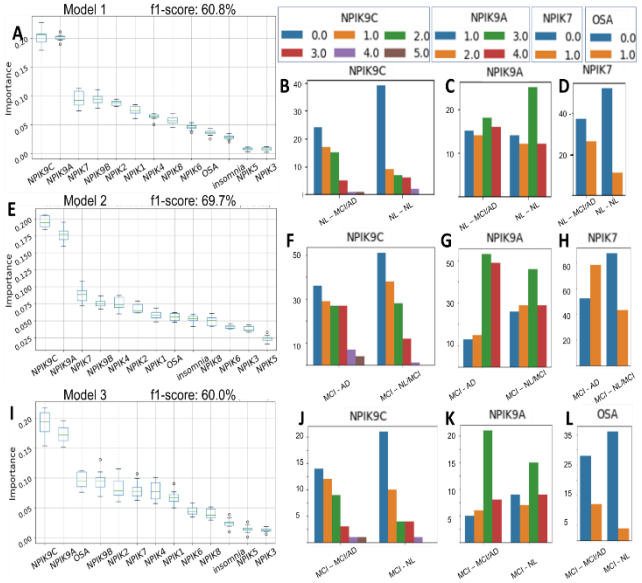


Fig. 3. Sleep-relevant feature importance rankings in random forest models and the distribution of three most important features in each model. 13 sleep-relevant features are used in random forest models to predict A – D: if a normal patient will get worse (NL – MCI/AD) or not (NL - NL); **E – H**: if an MCI patient will get a worse (MCI - AD) or not (MCI – NL/MCI); **I – L**: if an MCI patient will get better (MCI - NL) or not (MCI – MCI/AD).

reverse conversion it was OSA (Fig 3. L). In each bar plot, the right side was the group where the diagnosis did not get worse. Comparing to the bars at left side, there are less caregiver distress (NPIK9C = 0, Fig 3. B, F, J), smaller disturbances frequency, fewer excessive day sleep cases, and fewer OSA patients.

B. *Predicting diagnosis changes using brain volume*

Four random forest models based on brain volume yield accuracy between 60.9% and 79.8%. The most important feature of Model 4 (Fig 4. A - D) and Model 5 (Fig 4. E - H), which aimed to predict worsening conversion, was relative ventricles volume (Fig 4. A, E). From the boxplots (Fig. 4, C and G), we can see that the ventricles increase more over time for patients whose diagnosis changed to worse (NL-MCI/AD or MCI-AD), as the medians for that patient group are higher (approx. 1.2). The second and third most important features are relative entorhinal (Fig 4. C, H), hippocampus (Fig 4.G), and fusiform volumes (Fig 4. D). For these features, the boxplots show that the volume for the stable or improving

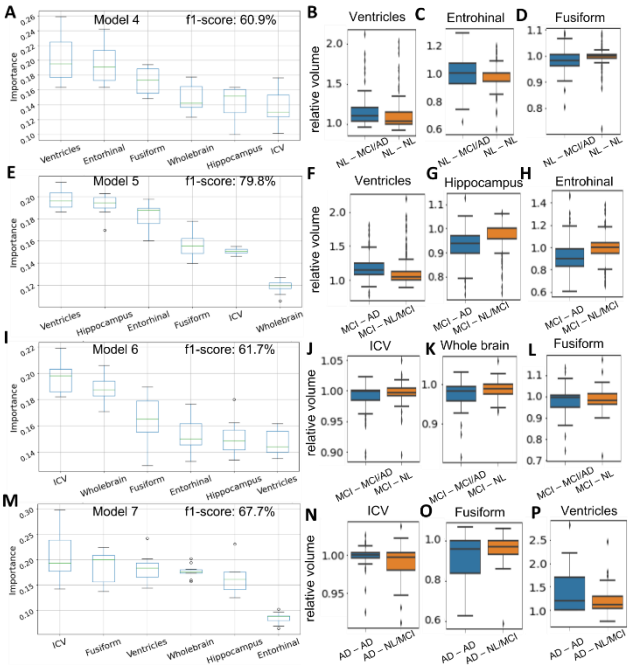


Fig. 4. Relative brain volume-relevant feature importance rankings in random forest models and the distribution of three most important features in each model. Six features are used in random forest models to predict (A – D): if a normal patient will get worse (NL – MCI/AD) or not (NL - NL); **(E – H)**: if an MCI patient will get worse (MCI - AD) or not (MCI – NL/MCI); **(I – L)**: if an MCI patient will get better (MCI - NL) or not (MCI – MCI/AD); and (M – P): if an AD patient will get better (AD – NL/MCI) or not (AD - AD).

patient groups stayed roughly the same, while it decreased or had larger variance for worsening conversions.

Model 6 (Fig 4. I-L) and Model 7 (Fig 4. M-P) aimed to predict the reverse or no-reverse conversions. The most important feature was ICV (Fig 4. J, N). The second and third most important features are relative fusiform (Fig 4. L, O), whole brain (Fig 4. K), and ventricles volumes (Fig 4. P).

TABLE 4: RESULTS OF FEATURE SELECTION

Model	***Number of Trees***	***f1 - Score***
Baseline	32	73.8%
OSA	512	73.4%
Insomnia	1024	73.1%
NPIKTOT	1024	73.3%
all variables	128	74.5%

C. Feature selection for detecting AD patients

The baseline random forest model using only brain size achieved an accuracy of 73.8%. When including one of the sleep disturbances factors (OSA, insomnia, NPIKTOT), the model performance was not improved, even when using more trees. The model including all three sleep disturbance factors and baseline yielded the best performance, with an f1-score of 74.5% (Table 4). For the model using all variables we analyzed the importance of each feature (Fig. 5). The sleep disturbance markers scored lowest on importance, while the relative hippocampus volume scored the highest, which fits with our f1-scores from the models.

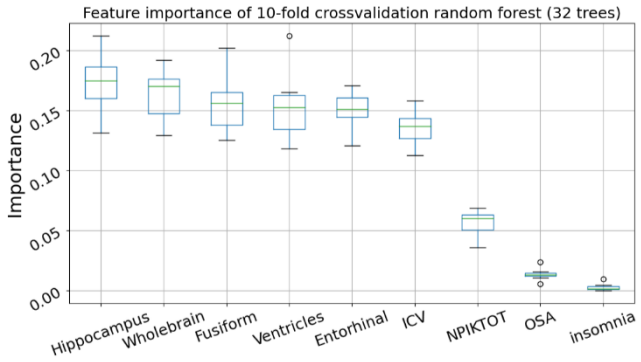


Fig. 5: Feature importance for brain volume and sleep markers

TABLE 3. TOP-3 IMPORTANT FACTORS IN EACH RANDOM FOREST MODEL WITH THE TARGET OF AD CONVERSION

Goal	**Model index**	**target labels**	**Nr. Sample per label**	**F1-score**	**1st impor. factor**	**2nd impor. factor**	**3rd impor. factor**
Will the AD status get ***worse***? (based on ***sleep*** disturbances features)	1	Worse: NL - MCI/AD Not worse: NL - NL	63 63	60.8%	**NPIK9C**	NPIK9A	NPIK7
	2	Worse: MCI - AD Not worse: MCI - NL/MCI	130 130	69.7%	**NPIK9C**	NPIK9A	NPIK7
Will the AD status get ***better***? (based on ***sleep*** disturbances features)	3	Better: MCI - NL Not better: MCI - MCI/AD	40 40	60.0%	**NPIK9C**	NPIK9A	OSA
Will the AD status get ***worse***? (based on relative ***brain volumes***)	4	Worse: NL - MCI Not worse: NL - NL	87 87	60.9%	Ventricles	Entorhinal	Fusiform
	5	Worse: MCI - AD Not worse: MCI - NL/MCI	215 215	79.8%	Ventricles	Hippocampus	Entorhinal
Will the AD status get ***better***? (based on relative ***brain volumes***)	6	Better: MCI - NL Not better: MCI - MCI/AD	74 74	61.7%	ICV	Wholebrain	Fusiform
	7	Better: AD - MCI/AD Not better: AD - AD	34 34	67.7%	ICV	Fusiform	Ventricles

D. HDBSCAN Segmentation of MRI brain scans

The HDBSCAN was able to pick up certain structures in all three patients, such as the hemispheres, the midbrain and the thalamus (Fig 6.). For the NL brain, almost all of the brain was considered to be one connected structure. In that brain, the HDBSCAN also made clusters that seem to be the eyes and maybe the ear canals. For the MCI and AD brain, there appears to be larger gap between the hemispheres than for the NL brain. Interestingly, in the MCI brain, the ventricles were identified as separate structures. These are however only three examples: if we used other patients in the same diagnosis groups, the results would likely be different.

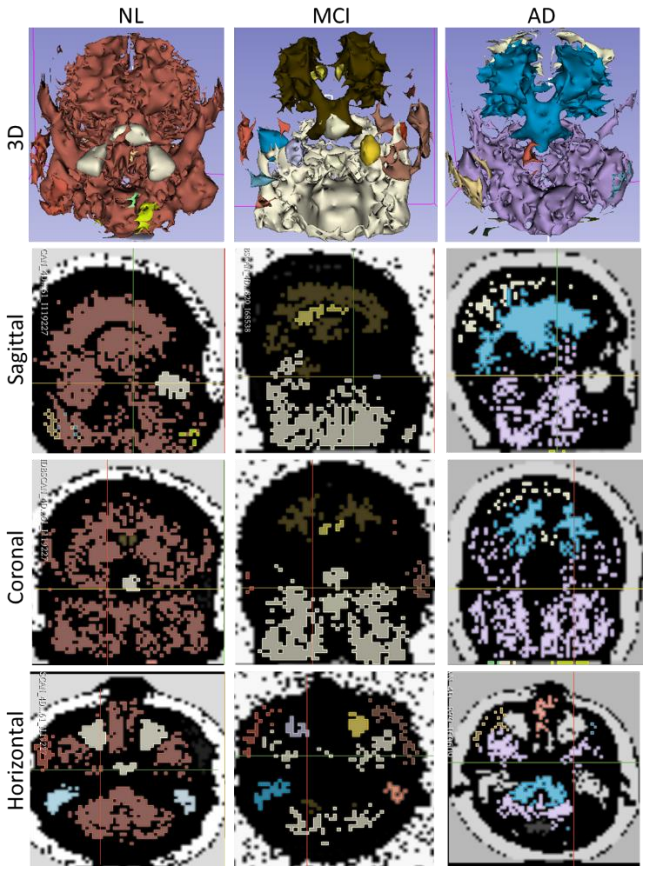


Fig. 6 HDBSCAN segmentations of MRI brain scans from NL, MCI and AD patients.

V. Discussion

It has been suggested by three researchers independently that sleep disturbances might increase the risk of dementia [10]. The goal of this study was to understand the relationship between sleep disturbances and AD cognitive diagnosis conversion. Shokouhi (2019) has found that abnormal night-time behavior was significantly associated with AD pathologies. NPIK3, 5, and 8 questions were indicated to have a high association with AD pathologies. Compared to Shokouhi's work, our study investigated the correlation between AD cognitive diagnosis conversion with more sleep disturbances factors. Our results showed that caregiver distress, sleep disturbances frequency, excessive sleep during daytime, and OSA are the major factors that influence patients' AD cognitive progression. We also found that more patients without caregiver distress (NPIK9C = 0, Fig2. B, F, J) were in the group where the diagnosis did not get worse. This may suggest that patients without caregiver distress are more likely to improve in their cognitive performance. Similarly, low sleep disturbances frequency (NPIK9A, Fig2. C, G, K), no excessive sleep over daytime (Fig2. D, H) and absence of OSA (Fig2. L) may help patients to slow down their AD cognitive progression.

Though it is yet unclear if sleep disturbances are a cause or consequence of AD, Lloret MA [4] suggested that sleep disturbances could be used to support AD cognitive diagnosis and evaluate the AD progression, because they are well correlated with the AD cognitive diagnosis.

One of the limitations of this study is that some of the sleep disturbances factors, such as sleep-related movement disorder (SRMD), circadian rhythm sleep disorder (CRSD), and nonspecific sleep problems were excluded, as they are not yet available in ADNI database. Furthermore, the sleep information that we extracted was based on keyword search. This may have resulted in false positive flags for insomnia and OSA.

Abnormal brain atrophy starts often at the beginning of MCI, and mostly in the medial temporal lobe (MTL). Therefore, brain volumes are considered as a good indicator for assessing the AD progression in the early phase. Several groups [5-9] have independently developed or are developing advanced diagnostic tools using brain MRI scans based on machine learning algorithms. Our random forest models have identified important brain reduction regions which could be helpful to diagnose AD cognitive progression. Using HDBSCAN 3D segmentation to identify brain regions using MRI scans has given us impressive results. Our long-term goal is to optimize the algorithms to yield highly accurate segmentations able to differentiate between NL and AD brain structures. Potentially this method may also be an improvement for calculating volume of brain structures.

One limitation could be the difficulty of identifying the brain regions which are very important but have small volumes, for example the hippocampus. To achieve high accuracy, we should generate high resolution segmentations which could be very time- and memory-consuming, although it is already improved compared to DBSCAN.

VI. Conclusion

Our results suggest that patients that suffer from AD or MCI may have a slower progression of cognitive decline if their support system is less distressed. To verify this, further research would be needed that focuses on caregivers. Treating sleep disorders and OSA could have a positive effect on the progression of AD. Further research focusing on disordered sleep and its effect on cognitive decline would be needed, especially with better data than self-declarations. While HDBSCAN is an exciting new way of visualizing the brain that could potentially help to identify abnormal brain structures, the algorithm would have to be improved vastly for this.

Acknowledgment

We'd like to thank our Alberto Rafael Ramos and Azizi Seixas for the support in medical aspects. Parts of the data preparation were done using code written by Kajenthini Kobivasan, Chantal Strasser and Joseph Weiber. The HDBCSCAN notebook was supplied by Leticia Fernández Moguel.

REFERENCES

[1] Shokouhi S. Associations of informant-based sleep reports with Alzheimer's disease pathologies. Clin Interv Aging. 2019 Sep 9;14:1631-1642. doi: 10.2147/CIA.S215208. PMID: 31571843; PMCID: PMC6750709.

[2] Li P, Gao L, Yu L, Zheng X, Ulsa MC, Yang HW, Gaba A, Yaffe K, Bennett DA, Buchman AS, Hu K, Leng Y. Daytime napping and Alzheimer's dementia: A potential bidirectional relationship. Alzheimers Dement. 2023 Jan;19(1):158-168. doi: 10.1002/alz.12636. Epub 2022 Mar 17. PMID: 35297533; PMCID: PMC9481741.

[3] 2019 Alzheimer's Statistics. (2019). Retrieved from https://www.alzheimers.net

[4] Lloret MA, Cervera-Ferri A, Nepomuceno M, Monllor P, Esteve D, Lloret A. Is Sleep Disruption a Cause or Consequence of Alzheimer's Disease? Reviewing Its Possible Role as a Biomarker. Int J Mol Sci. 2020 Feb 10;21(3):1168. doi: 10.3390/ijms21031168. PMID: 32050587; PMCID: PMC7037733.

[5] Blinkouskaya Y, Weickenmeier J. Brain Shape Changes Associated With Cerebral Atrophy in Healthy Aging and Alzheimer's Disease. Front Mech Eng. 2021 Jul;7:705653. doi: 10.3389/fmech.2021.705653. Epub 2021 Jul 19. PMID: 35465618; PMCID: PMC9032518.

[6] Beheshti I, Demirel H, Matsuda H; Alzheimer's Disease Neuroimaging Initiative. Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm. Comput Biol Med. 2017 Apr 1;83:109-119. doi: 10.1016/j.compbiomed.2017.02.011. Epub 2017 Feb 27. PMID: 28260614.

[7] Shaikh TA, Ali R. Automated atrophy assessment for Alzheimer's disease diagnosis from brain MRI images. Magn Reson Imaging. 2019 Oct;62:167-173. doi: 10.1016/j.mri.2019.06.019. Epub 2019 Jul 4. PMID: 31279772.

[8] Zhang Y, Liu S. Analysis of structural brain MRI and multi-parameter classification for Alzheimer's disease. Biomed Tech (Berl). 2018 Jul 26;63(4):427-437. doi: 10.1515/bmt-2016-0239. PMID: 28622141.

[9] M Arabi E, S Ahmed K, S Mohra A. Advanced Diagnostic Technique for Alzheimer's Disease using MRI Top-Ranked Volume and Surface-based Features. J Biomed Phys Eng. 2022 Dec 1;12(6):569-582. doi: 10.31661/jbpe.v0i0.2112-1440. PMID: 36569569; PMCID: PMC9759646.

[10] Shi L, Chen SJ, Ma MY, Bao YP, Han Y, Wang YM, Shi J, Vitiello MV, Lu L. Sleep disturbances increase the risk of dementia: A systematic review and meta-analysis. Sleep Med Rev. 2018 Aug;40:4-16. doi: 10.1016/j.smrv.2017.06.010. Epub 2017 Jul 6. PMID: 28890168.

CONTRIBUTION

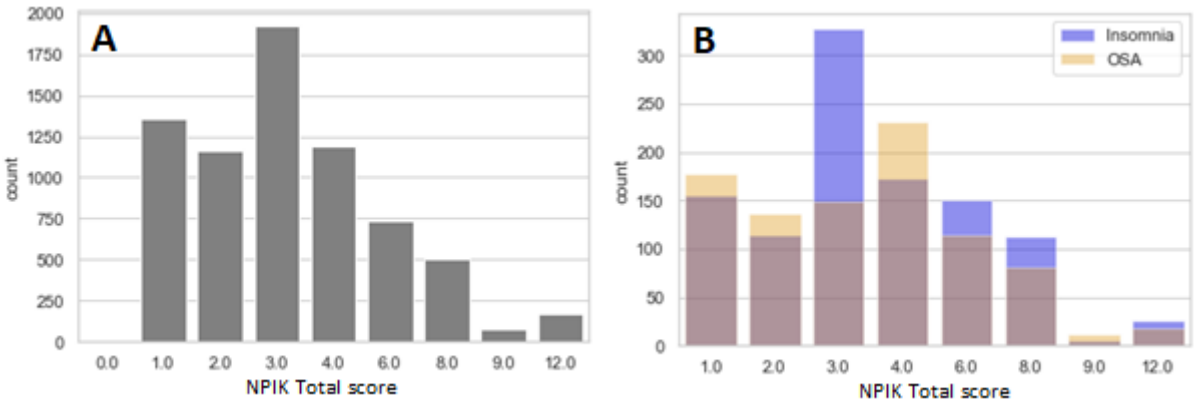
Both authors contributed equally to the work. The authors declare that there is no conflict of interest.

APPENDIX

A.1: OVERVIEW OF TABLES USED

Source tables	**Usage**
ADSXLIST	Creation / selection of OSA and insomnia marker
BLCHANGE	Creation / selection of OSA and insomnia marker
INITHEALTH	Creation / selection of OSA and insomnia marker
RECMHIST	Creation / selection of OSA and insomnia marker
RECBLLOG	Creation / selection of OSA and insomnia marker
BLSCHECK	Creation / selection of OSA and insomnia marker
ADNIMERGE	Brain volumes, imaging information, age, demographic information
DXSUM_PDXCONV_ADNIALL	Diagnosis information
NPI	NPIK answers
NPIQ	NPIK answers

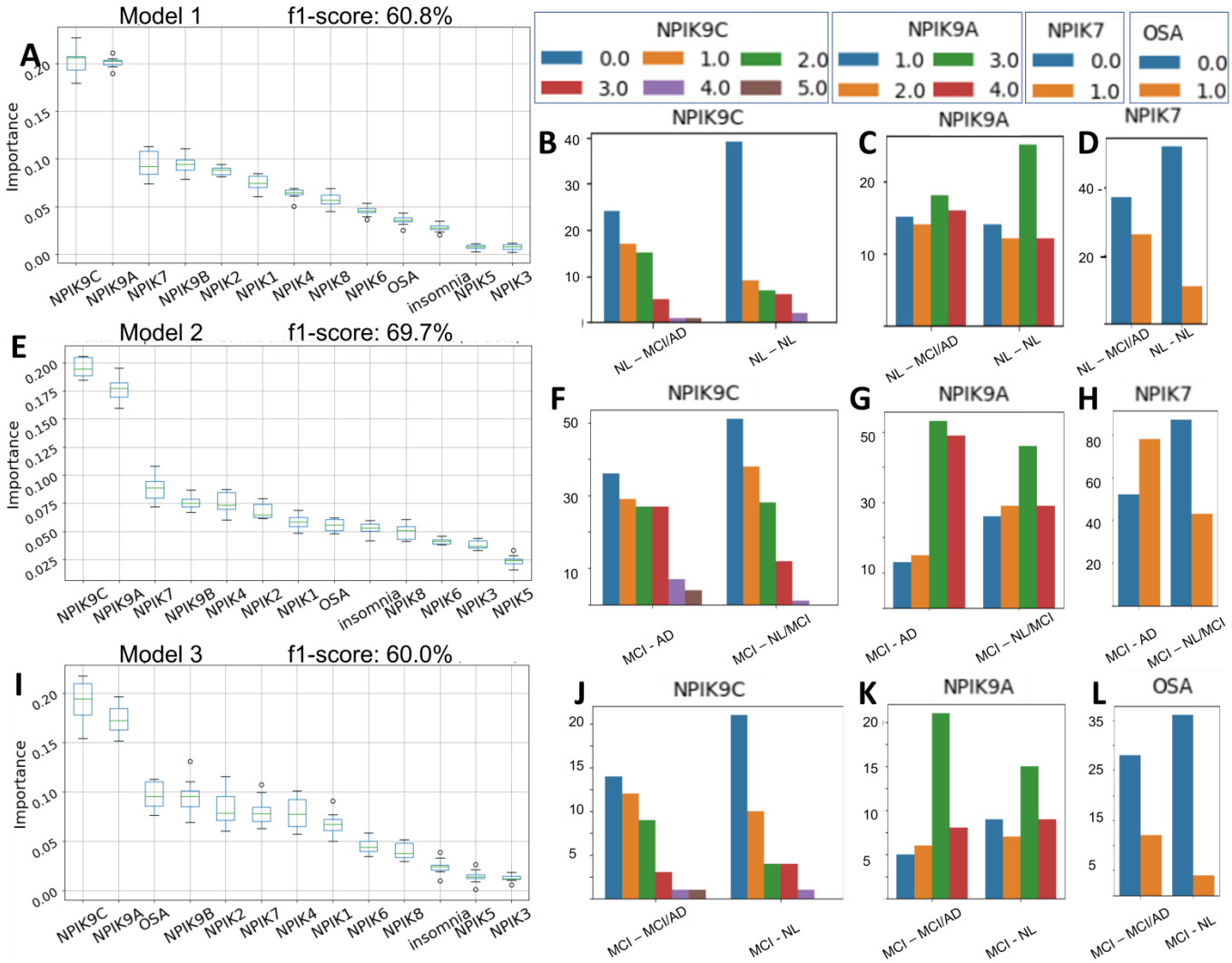
A.2: FIGURE 1



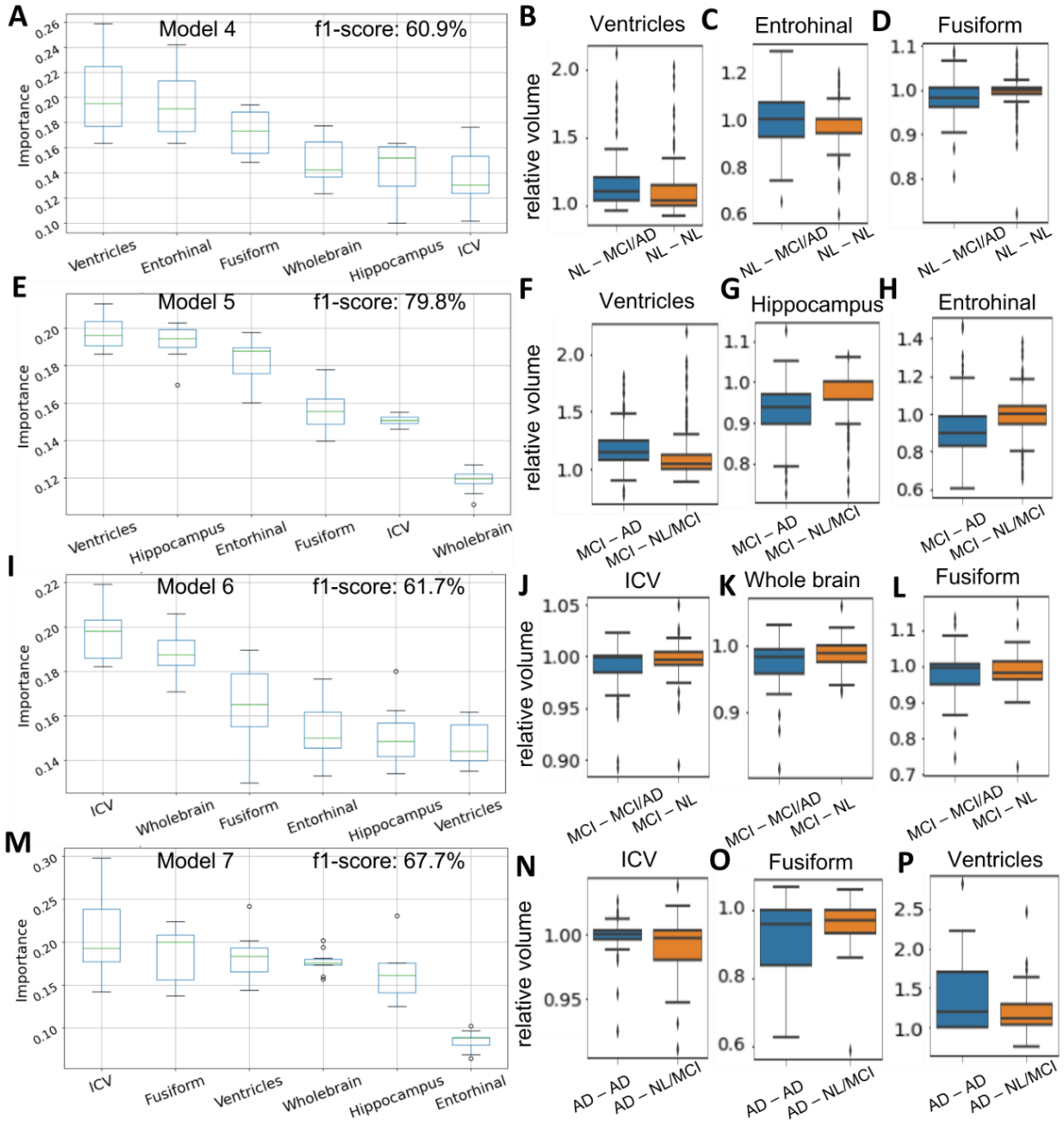
A.3: NPIK QUESTIONNAIRE WITH RESULTS

ID	*Question*	*Ratings*
NPIK1	Difficulties in falling asleep?	0: no; 1: yes
NPIK2	Getting up during the night?	0: no; 1: yes
NPIK3	pacing and wandering at night?	0: no; 1: yes
NPIK4	Awakening the caregiver during night?	0: no; 1: yes
NPIK5	Awaken at night or too early in the morning and thinking that it is time to start the day?	0: no; 1: yes
NPIK6	Waking up too early in the morning?	0: no; 1: yes
NPIK7	Excessive daytime napping?	0: no; 1: yes
NPIK8	Other abnormal nighttime behaviors?	0: no; 1: yes
NPIK9A	Frequency ratings?	1: occasionally; 2: often; 3: frequently; 4: very frequently
NPIK9B	Severity ratings?	1: mild; 2: moderate; 3: marked
NPIK9C	Caregiver distress?	0: not at all; 1: minimally; 2: mildly; 3: moderately; 4: severely; 5: very severely

A.4: FIGURE 3


Model 1
f1-score: 60.8%
A
Importance
NPIK9C NPIK9A NPIK7 NPIK9B NPIK2 NPIK1 NPIK4 NPIK8 NPIK6 OSA insomnia NPIK5 NPIK3
Model 2
f1-score: 69.7%
E
Importance
NPIK9C NPIK9A NPIK7 NPIK9B NPIK4 NPIK2 NPIK1 OSA insomnia NPIK8 NPIK6 NPIK3 NPIK5
Model 3
f1-score: 60.0%
I
Importance
NPIK9C NPIK9A OSA NPIK9B NPIK2 NPIK7 NPIK4 NPIK1 NPIK6 NPIK8 insomnia NPIK5 NPIK3
NPIK9C
0.0 1.0 2.0 3.0 4.0 5.0
NPIK9A
1.0 3.0 2.0 4.0
NPIK7
0.0 1.0
OSA
0.0 1.0
B
NPIK9C
NL – MCI/AD
NL – NL
C
NPIK9A
NL – MCI/AD
NL – NL
D
NPIK7
NL – MCI/AD
NL - NL
F
NPIK9C
MCI - AD
MCI – NL/MCI
G
NPIK9A
MCI - AD
MCI – NL/MCI
H
NPIK7
MCI - AD
MCI – NL/MCI
J
NPIK9C
MCI – MCI/AD
MCI - NL
K
NPIK9A
MCI – MCI/AD
MCI - NL
L
OSA
MCI – MCI/AD
MCI - NL

A.5: FIGURE 4


A
Model 4
f1-score: 60.9%
Importance
Ventricles
Entorhinal
Fusiform
Wholebrain
Hippocampus
ICV
B
Ventricles
relative volume
NL – MCI/AD
NL – NL
C
Entrohinal
D
Fusiform
E
Model 5
f1-score: 79.8%
Ventricles
Hippocampus
Entorhinal
Fusiform
ICV
Wholebrain
F
Ventricles
MCI – AD
MCI – NL/MCI
G
Hippocampus
H
Entrohinal
I
Model 6
f1-score: 61.7%
ICV
Wholebrain
Fusiform
Entorhinal
Hippocampus
Ventricles
J
ICV
MCI – MCI/AD
MCI – NL
K
Whole brain
L
Fusiform
M
Model 7
f1-score: 67.7%
ICV
Fusiform
Ventricles
Wholebrain
Hippocampus
Entorhinal
N
ICV
AD – AD
AD – NL/MCI
O
Fusiform
P
Ventricles

Goal	Model index	target labels	Nr. Sample per label	F1-score	1st impor. factor	2nd impor. factor	3rd impor. factor
Will the AD status get ***worse***? (based on ***sleep*** disturbances features)	1	Worse: NL - MCI/AD Not worse: NL - NL	63 63	60.8%	**NPIK9C**	NPIK9A	NPIK7
	2	Worse: MCI - AD Not worse: MCI - NL/MCI	130 130	69.7%	**NPIK9C**	NPIK9A	NPIK7
Will the AD status get ***better***? (based on ***sleep*** disturbances features)	3	Better: MCI - NL Not better: MCI - MCI/AD	40 40	60.0%	**NPIK9C**	NPIK9A	OSA
Will the AD status get ***worse***? (based on relative ***brain volumes***)	4	Worse: NL - MCI Not worse: NL - NL	87 87	60.9%	Ventricles	Entorhinal	Fusiform
	5	Worse: MCI - AD Not worse: MCI - NL/MCI	215 215	79.8%	Ventricles	Hippocampus	Entorhinal
Will the AD status get ***better***? (based on relative ***brain volumes***)	6	Better: MCI - NL Not better: MCI - MCI/AD	74 74	61.7%	ICV	Wholebrain	Fusiform
	7	Better: AD - MCI/AD Not better: AD - AD	34 34	67.7%	ICV	Fusiform	Ventricles

A.7: FIGURE 6