Multi-scale brain simulation with integrated positron emission tomography yields hidden local field potential activity that augments machine-learning classification of Alzheimer's disease

Paul **Triebkorn*** (1, 2, 3), Leon **Stefanovski*** (1, 2), Kiret **Dhindsa** (1, 2), Margarita-Arimatea **Diaz-Cortes** (1, 2), Patrik **Bey** (1, 2), Konstantin **Bülau** (1, 2), Roopa **Pai** (1, 2, 4), Andreas **Spiegler** (1, 2 ,5), Ana **Solodkin** (6), Viktor **Jirsa** (3), Anthony Randal **McIntosh** (7), Petra **Ritter**° (1, 2, 4); for the Alzheimer's Disease Neuroimaging Initiative†

1) Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Berlin, Germany
2) Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Department of Neurology with Experimental Neurology, Brain Simulation Section, Berlin, Germany
3) Institut de Neurosciences des Systèmes, Aix Marseille Université, Marseille, France
4) Bernstein Center for Computational Neuroscience Berlin, Germany
5) Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
6) Neuroscience, Behavioral and Brain Sciences, UT Dallas. Richardson, TX, USA
7) Baycrest Health Sciences, Rotman Research Institute, Toronto, Ontario, Canada

* The authors Paul Triebkorn and Leon Stefanovski had equal contributions to this article.

° Corresponding author:

Prof. Dr. med. Petra Ritter
Department of Neurology and Experimental Neurology
Brain Simulation Section
Charitéplatz 1
10117 Berlin
Germany
Email: petra.ritter@charite.de

†Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at:

http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

ABSTRACT.

Introduction. While the prevalence of neurodegenerative diseases and dementia increases, our knowledge of the underlying pathomechanisms and related diagnostic biomarkers, outcome predictors, or therapeutic targets remains limited. In this article, we

show how computational multi-scale brain network modeling using The Virtual Brain (TVB) simulation platform supports revealing potential disease mechanisms and can lead to improved diagnostics.

Methods. TVB allows standardized large-scale structural connectivity (SC)-based modeling and simulation of whole-brain dynamics. We combine TVB with a cause-and-effect model for amyloid-beta, and machine-learning classification with support vector machines and random forests. The amyloid-beta burden as quantified from individual AV-45 PET scans informs parameters of local excitation/inhibition balance. We use magnetic resonance imaging (MRI), positron emission tomography (PET, specifically Amyloid-beta (Abeta) binding tracer AV-45-PET, and Tau-protein (Tau) binding AV-1451-PET) from 33 participants of Alzheimer's Disease Neuroimaging Initiative study 3 (ADNI3). The frequency compositions of simulated local field potentials (LFP) are under investigation for their potential to classify individuals between Alzheimer's disease (AD), Mild Cognitive Impairment (MCI), and healthy controls (HC) using support vector machines and random forest classifiers.

Results. The combination of empirical features (subcortical volumetry, AV-45- and AV-1451- PET standard uptake value ratio, SUVR per region) and simulated features (mean LFP frequency per brain region) significantly outperformed the classification accuracy of empirical data alone by about 10% in the accuracy index of weighted F1-score (empirical 64.34% vs. combined 74.28%). There was no significant difference between empirical and simulated features alone. The features with the highest feature importance showed high biological plausibility with respect to the AD-typical spatial distribution of the features. This was demonstrated for all feature types, e.g., increased importance indices for the left entorhinal cortex as the most important Tau-feature, the left dorsal temporopolar cortex for Abeta, the right thalamus for LFP frequency, and the right putamen for volume.

Discussion. In summary, here we suggest a strategy and provide proof of concept for TVB-inferred mechanistic biomarkers that are direct indicators of pathogenic processes in neurodegenerative disease. We show how the cause-and-effect implementation of local hyperexcitation caused by Abeta can improve the machine-learning-driven classification of AD. This proves TVBs ability to decode information in empirical data by means of SC-based brain simulation.

Keywords: Alzheimer's Disease, The Virtual Brain, Machine Learning, Positron Emission Tomography

Contains colored figures.

1. INTRODUCTION

Alzheimer's disease (AD) is a major public health concern in industrialized societies, leading to direct and indirect estimated annual costs of $1 trillion in the US (Alzheimer's Association 2019). This figure is set to double by 2030 with a predicted burden of 152 million cases by 2050, considering the development of the disease in a person every three seconds (Alzheimer's Association 2019).

Currently, no cure or disease-modifying therapy is available for AD patients and actual treatment-regimens only provide temporary symptomatic improvement (Grassi, Perna et al. 2018). Additionally, due to the large gap between probable disease onset with beginning neuropathologic brain changes, and the first clinical symptoms of AD of up to 20 years (Nunomura, Perry et al. 2001), early diagnosis plays a crucial role in the development of potential future treatments. Moreover, many patients suffer from Mild Cognitive Impairment (MCI) years before the onset of manifest dementia (Petersen, Caracciolo et al. 2014), which, in contrast to MCI, also impairs the activity of daily living (McKhann, Knopman et al. 2011).

Although the spectrum of AD-related disease burden is broad and its early diagnosis a common modern health problem, the knowledge of underlying disease mechanisms remains incomplete. Besides the two hallmark proteins Amyloid-beta (Abeta) (Sadigh-Eteghad, Sabermarouf et al. 2015, Selkoe and Hardy 2016) and Tau (Jadhav, Avila et al. 2019, Tapia-Rojas, Cabezas-Opazo et al. 2019), other involved factors have been identified, such as e.g. impairment of the blood-brain-barrier (Storck Steffen and Pietrzik Claus 2018, Sweeney, Sagare et al. 2018, Zetterberg and Schott 2019), synaptic dysfunction (Tonnies and Trushina 2017, Jackson, Jambrina et al. 2019), network disruption, (Selkoe 2019), mitochondrial dysfunction (Swerdlow and Khan 2009), neuroinflammation (Heneka, Carson et al. 2015) as well as genetic risk factors (Van Cauwenberghe, Van Broeckhoven et al. 2016, Pimenova, Raj et al. 2018, Takatori, Wang et al. 2019). While Abeta and Tau are widely accepted as involved core features (Blennow, de Leon et al. 2006, Bloom 2014, Jack, Bennett et al. 2018), their mutual interaction (Bloom 2014) and interaction with other factors (Gauthier, Zhang et al. 2018, Zetterberg and Schott 2019) are incompletely understood. Comprehensive knowledge of this multifactorial interaction in the pathogenesis of AD is crucial for further therapeutic developments.

The Virtual Brain (TVB) platform for modeling and simulating large-scale brain networks by using personalized structural connectivity models (Ritter, Schirner et al. 2013, Sanz Leon, Knock et al. 2013) enables the model-based inference of underlying neurophysiological mechanisms across different brain scales that are involved in the generation of macroscopic neuroimaging signals including functional Magnetic Resonance Imaging (fMRI), Electroencephalography (EEG) and Magnetoencephalography (MEG). Moreover, TVB facilitates the reproduction and evaluation of individual configurations of the brain through the use of subject-specific data. In this study, we make use of virtual local field potentials (LFPs) from simulated brain data from a recent experiment with TVB presented in (Stefanovski, Triebkorn et al. 2019). By integrating individual patterns from Abeta, obtained from Positron Emission Tomography (PET) with the Abeta-binding tracer ^{18}F-AV-45 into the brain model, distinct spectral patterns in simulated LFPs and EEG could be observed for patients with AD, MCI, and healthy control (HC) subjects (**Figure 1**). Such integration was done by transferring the local concentration of Abeta to a variation in the brain model's local excitation-inhibition balance. This resulted in a shift from alpha to theta rhythms in AD patients, which was located in a similar pattern as local hyperexcitation in core structures of the brain network. The frequency shift was reversible by applying "virtual memantine", i.e., virtual NMDA antagonistic drug therapy. An overview of the study results is provided in **Figure 1.**

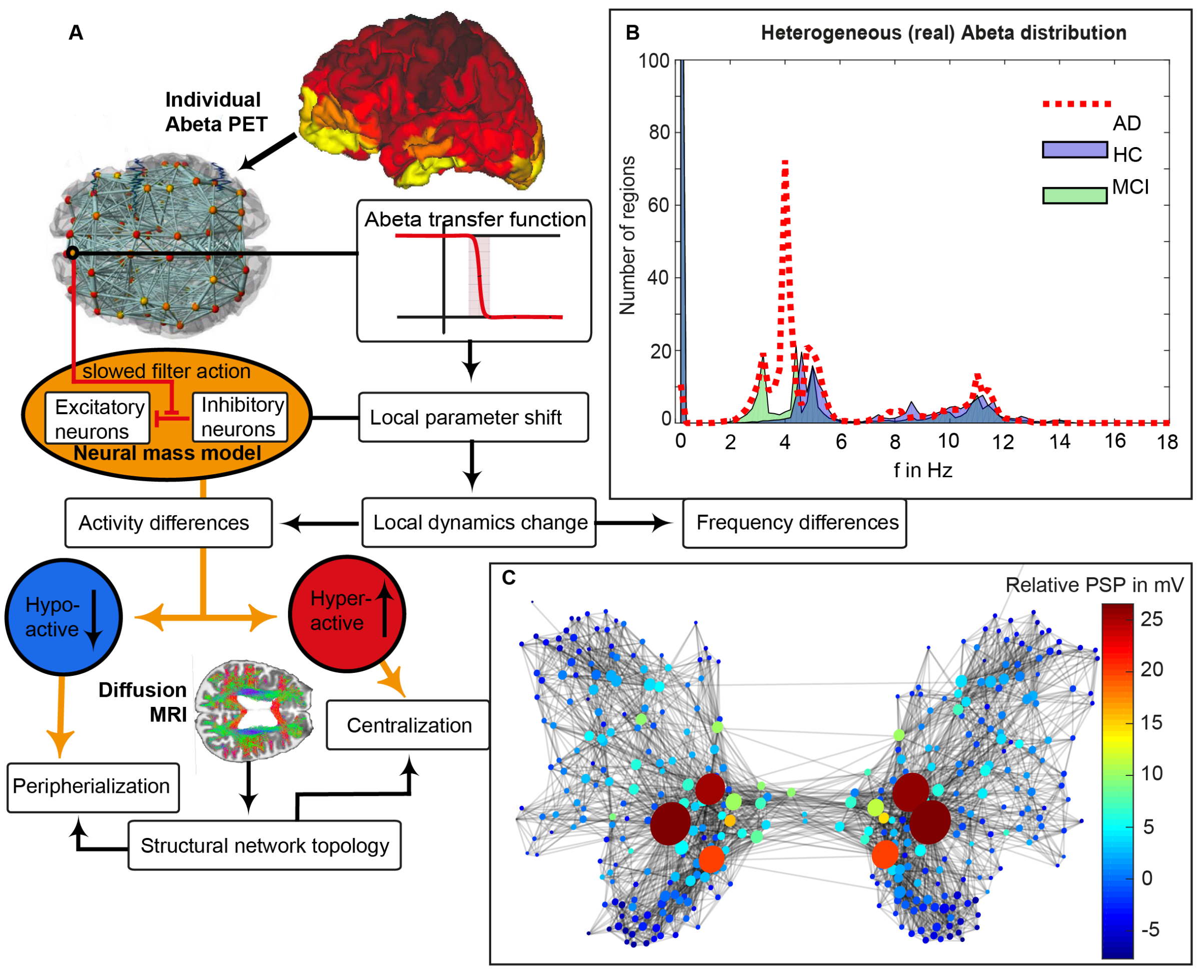

Figure 1. Modified from (Stefanovski, Triebkorn et al. 2019) Abeta-PET-driven brain simulation model of AD (A): Regional PET intensity constraints regional parameters. A sigmoidal transfer function translates the regional Abeta load to changes in the E/I-balance. (B) Virtual AD patient brains exhibited significantly slower simulated LFPs than MCI and HC virtual brains and showed a shift from alpha to theta frequency range. This slowing was spatially associated with local hyperexcitation (C). The graph in (C) represents the SC, wherein the nodes' size reflects the degree, while color corresponds to the relative PSP (relative to the mean PSP of the simulation). The graph indicates that local hyperexcitation occurs in central parts of the networks. (D – F): A three-dimensional histogram shows the distribution of regional LFP frequencies across the scale coupling parameter G. While the AD group is dominated by two clusters in the alpha and theta band (D), the groups of HC (E) and MCI (F) have an additional strong cluster exhibiting no oscillations ("frequency of zero"). This phenomenon is absent in the AD group. The stable focus in HC and MCI virtual brains, which disables oscillations for G's higher values, provides an additional – simulation inferred - distinctive criterion between groups.

Simulation of the individual brain is possible through the inclusion of personal empirical data. AD specific pathologies, such as deposition of Abeta in neuritic plaques, Tau deposition in neurofibrillary tangles, and atrophy of neural tissue, have been widely studied - including with machine learning (ML) approaches (van Rossum, Vos et al. 2010, Forlenza, Diniz et al. 2013). The major advantage of employing ML-based classification algorithms on neuroimaging data is the potential for recognizing complex high dimensional, previously unknown disease patterns in the data, potentially identifying AD before clinical manifestation or predicting a disease trajectory. over the last 15 years, ML has evolved as a fast-growing multidisciplinary field of research, with many applications in various scientific areas (Dhall, Kaur et al. 2020), including neuroimaging in AD (Martí-Juan, Sanroma-Guell et al. 2020). ML techniques range from relatively simple mathematical models to complex approaches. Their principal aim has been to "give computers the ability to learn without being explicitly programmed to do so" (Samuel 1959). In other words, ML can be seen as a collection of methods developed to enable computational systems to learn from the data with the primary purpose of making predictions and inferences. The advantage of ML algorithms over traditional statistical or model-based approaches is that they can discover subtle and even complex patterns in high-dimensional data that would be difficult to identify or encode otherwise (Sajda 2006). However, the ability of ML algorithms to discover patterns in data can also result in the reliance on spurious correlations that appear in data by chance, or which are otherwise not clinically generalizable beyond the data used to train the algorithm. Many algorithms have been developed with successful convincing classification capabilities (Kotsiantis, Zaharakis et al. 2006, Binkhonain and Zhao 2019).

Generally, the problem of classification is the prediction of categories for each point in a (multidimensional) data cloud based on their representations, that is, identifying the class to which an input belongs among a set of labeled categories (Rebala, Ravi et al. 2019). A classification model, also called a classifier, can be binary or multi-class depending on the number of groups or labels to predict. A classifier is then tailored from the learning process of categorizing a set of "training" data. Despite the numerous classification methods currently available for this inference, it is impossible to conclude which one is generally superior to the other. This particular fact depends on the application and nature of the available data set and is commonly referred to as the "no-free-lunch theorem" (Wolpert and Macready 1997).

Classification allows considering other factors, like lifestyle risks, genetic conditions, among others (Ithapu, Singh et al. 2015). Numerous ML methods have been used to classify and predict AD stages with promising results (Haller, Lövblad et al. 2011, Falahati, Westman et al. 2014, Rathore, Habes et al. 2017). While some studies have made use of a single screening modality, such as MRI (Fan, Batmanghelich et al. 2008, Kloppel, Stonnington et al. 2008, Cuingnet, Gerardin et al. 2011, Liu, Zhang et al. 2012, Tong, Wolz et al. 2014), or electroencephalography (EEG) (Blinowska, Rakowski et al. 2017, Farina, Emek-Savaş et al. 2020, Ferri, Babiloni et al. 2020, Oltu, Akşahin et al. 2021), others have used a combination of multiple imaging techniques including MRI, PET, and cerebrospinal fluid (CSF) biomarkers (Zhang, Wang et al. 2011, Gray, Aljabar et al. 2013, Jie, Zhang et al. 2013, Young, Modat et al. 2013, Teipel, Kurth et al. 2015, Yun, Kwak et al. 2015, Samper-González, Burgos et al. 2018). Although many of those studies presented interesting and promising results in AD classification, most focused on a so-called two-class problem. It has been pointed out, that standardization and quality checks for robustness, CV schema, etc. are essential but often lacking (Blinowska, Rakowski et al. 2017). Only a few studies have made use of more complex classification, as a three-class problem with AD, MCI, and

HC (Suk, Lee et al. 2014, Farina, Emek-Savaş et al. 2020) or differentiation in between MCI to converters and non-converters (Aksu, Miller et al. 2011, Cui, Liu et al. 2011, Moradi, Pepe et al. 2015, Beheshti, Maikusa et al. 2017).

Among the spectrum of ML classifiers, in the present study, we developed a dual methodology for solving the classification task, by making use of Support Vector Machine (SVM) and Random Forest (RF). Both employed ML-classification approaches were cross-validated within a sample of 33 human subjects with MCI, AD, or HC from the ADNI database. The use of two different ML-classifiers in a nested approach provides a robust, generalizable, and appropriate evaluation of the classification; on the other hand, it enables exploring the empirical and simulated features of highest importance for separation between the three groups under study.

SVM is a widely used method for supervised ML classification problems and has been well established in recent Neuroscience literature (Rathore, Habes et al. 2017, Rondina, Ferreira et al. 2018, Shaikh and Ali 2019, Zhao, Ding et al. 2019, Soumaya, Taoufiq et al.). SVMs have been extensively employed due to their robustness, simplicity to implement, and because they can also be employed as non-linear classifiers by making simple variations. SVM is a specific type of so-called maximum margin classifiers that tries to find an optimal separating hyperplane (with the largest possible margin) between two (or more) groups within a higher dimensional representation of the original data. Due to its algorithm particularities, SVM has some advantages in small sample analysis (Jiang, Ching et al. 2017), which in this approach represents an advantage since we employ a sample size of 33 subjects.

Similar to SVMs, Random Forest (RF) has been widely used for classification within the ML community and has performed well in a range of applications for classification in neuroimaging studies (for a review see (Sarica, Cerasa et al. 2017)). It is based on a large number of decision trees performing binary splits on randomly selected subsets of features, and therefore uses a fundamentally different classification technique than SVMs. The main advantages are the resulting direct interpretability of the feature importance and high robustness towards overfitting of the algorithm (Breiman 2001). One of the main differences between RF and SVM is that RF directly provides a probability for each point of belonging to a defined class. In contrast, SVM provides the distance to the hyperplanes or boundaries.

To summarize, the design of our methodology posits:

1. a three-class task for AD, MCI, and HC with ML-classification
2. a nested dual classifier approach with SVM and RF
3. various sources of biological information in a “hybrid” methodology: multimodal empirical imaging data as well as simulated brain dynamics

One of our primary objectives is to determine if the use of extracted features from TVB adds to the classifiers’ predictive power. For achieving this, we repeated the ML procedure with three different feature sets: a) using empirical features alone, b) using simulated features alone, and c) combining both types of features into a hybrid model.

We aim to provide evidence (or falsify) that TVB inferred features improve the classification performance compared to imaging data features.

We show that TVB simulations provide additional unique diagnostic information that is not readily available without brain simulations. This lends support to the idea that TVB provides value and real-world applicability above and beyond merely reorganizing empirical data, and suggests that its simulations are biologically plausible.

2. MATERIALS AND METHODS

2.1. Alzheimer's disease Neuroimaging Initiative (ADNI) database

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see www.adni-info.org. We used data from ADNI in our recent work to simulate electrophysiological neuronal activity (EEG) and explore their spectral characteristics in AD, MCI, and HC (Stefanovski, Triebkorn et al. 2019). In the present study, we assess whether the results of the previous study can be used to improve ML classification of AD, MCI, and HC. Basic epidemiological properties can be found in (Stefanovski, Triebkorn et al. 2019).

2.2. Data acquisition, processing, and simulation

The detailed methodology of data acquisition, selection, processing, and simulation are described in (Stefanovski, Triebkorn et al. 2019). In the following, we will provide a summary. All study participants were from ADNI 3 and were scanned using Siemens scanners with a magnetic field strength of 3T. Compare Supplementary Table S1-S5 in (Stefanovski, Triebkorn et al. 2019) for the metadata. We included T1 MPRAGE (TE = 2.95 - 2.98 ms, TR = 2.3s), FLAIR (TE differs slightly, TR = 4.8s, matrix size = 160 x 256 x 256), DWI (TE = 56 -71 ms, TR = 3.4 - 7.2s, matrix size = 116 x 116 x 80, voxel size = 2 x 2 x 2, bvals = [0, 1000] or [0, 500, 1000, 2000], bvecs = 49 or 115), fieldmaps and PET Data (AV-45 for Abeta and AV-1451 for Tau). We used the human connectome project (HCP) minimal preprocessing pipeline (Glasser, Sotiropoulos et al. 2013) for the processing of structural data. This included Freesurfer (Reuter, Schmansky et al. 2012)

(https://surfer.nmr.mgh.harvard.edu/fswiki/FreeSurferMethodsCitation),

FSL (Smith, Jenkinson et al. 2004, Woolrich, Jbabdi et al. 2009, Jenkinson, Beckmann et al. 2012) and connectome workbench

(https://www.humanconnectome.org/software/connectome-workbench).

The adjustments on the HCP guidelines to fit our data are described in the methods of (Stefanovski, Triebkorn et al. 2019). For the subcortical volumetrics used in this study, we obtained the volumetry statistics provided by the *-autorecon2* command. The segmentation is performed with the modified Fischl parcellation (Fischl, Salat et al. 2002) of subcortical regions in Freesurfer

(http://freesurfer.net/fswiki/SubcorticalSegmentation).

For AV-45 and AV-1451 PET images, we used the already preprocessed images available in ADNI. We aligned the PET images to HCP-processed T1 images and performed linear registration with FLIRT (FSL). The resulting PET images were masked with subject-specific HCP brain masks. We calculated SUVRs by dividing image intensities by the mean intensity in the white matter of the cerebellum. Partial volume correction was applied using grey and white matter from Freesurfer segmentation and the Müller-Gärtner method from the PETPVC toolbox (Thomas, Cuplov et al. 2016). To get the average SUVR per region, subcortical SUVRs were defined as the average SUVR in subcortical GM. Simultaneously, cortical GM PET intensities were mapped onto the cortical surfaces using the connectome workbench tool. DWI preprocessing was performed by the MRtrix3 software package (http://www.mrtrix.org). We used the following functions: *Dwidenoise* (Veraart, Novikov et al. 2016), *Dwipreproc,*

(https://mrtrix.readthedocs.io/en/latest/dwi_preprocessing/dwipreproc.html),

Dwibiascorrect, *Diw2mask, Dwiintensitynorm, Dwi2response,* (Tournier, Calamante et al. 2013), *Average_response, Dwi2fod*, (Tournier, Calamante et al. 2007), *Tckgen* (Tournier, Calamante et al. 2010, Smith, Tournier et al. 2012, Smith, Tournier et al. 2015). The detailed steps are described in the original paper (Stefanovski, Triebkorn et al. 2019). The preprocessed cortical surfaces and T1 images were used to compute the Boundary Element Model in Brainstorm (Tadel, Baillet et al. 2011), wherein inner and outer Scalp, as well as outer skull, were modeled with 1922 vertices per layer and the default 'BrainProducts EasyCap 65' EEG cap. We estimated the leadfield matrix with the adjoint method in OpenMEEG (default conductivities 1 (scalp), 0.0125 (skull), and 1 (brain)). After processing the empirical imaging data, we used the SC of the HC population to generate an averaged standard SC for all participants. Model parameters were chosen due to considerations in former studies (Spiegler, Kiebel et al. 2010) and are described in (Stefanovski, Triebkorn et al. 2019). We explored a range of global scaling factor G to capture different dynamic states of the simulation. The novelty in our recent simulation study was the introduction of a mechanistic model for Abeta-driven effects. We linked local Abeta concentrations, measured by Abeta PET, to the E/I balance in the model by defining the inhibitory time constant τ_i as a function of local Abeta burden. We showed that this implementation led to a significantly lower LFP frequency in AD participants. Therefore, we calculate the EEG signal as a projection of the LFP from within the brain to the surface of the head, taking into the concept of a lead-field matrix simplification to three compartment borders brain-skull, skull-scalp, and scalp-air (Jirsa, Jantzen et al. 2002, Bojak, Oostendorp et al. 2010, Litvak, Mattout et al. 2011, Ritter, Schirner et al. 2013). However, the mean frequency is only one of a plethora of simulated features produced by the mechanistic disease model. In this study, we aim to further examine its properties for diagnostic classification.

In addition to the data used in our former study (Stefanovski, Triebkorn et al. 2019), we also used the distribution of Tau in the AV-14-51 PET for our analyses to obtain the best available empirical data basis. The information on the local Tau burden is taken from the ^{18}F-AV-14-51 PET imaging data. Again, the nuclear signal intensity is related to a reference volume in the cerebellum. High significance in Tau PET using ROI can be achieved from the usage of the Braak-and-Braak stages: regions of interest that are intended to represent the transentorhinal, limbic, or neocortical Braak stages (Schöll, Lockhart et al. 2016). The burden of tau measured by AV-1451 showed the highest group differences between AD patients and HC ($p < 0.001$) in the regions of the earlier stages I-IV. The cut-off SUVR was here >2.79 mean SUVR (Schöll, Lockhart et al. 2016). The composites of the Braak and Braak stages can be found in our description above in the introductory part. The regions of interest defined in that study can be found in the supplementary material of (Schöll, Lockhart et al. 2016).

2.3. General Machine Learning Approach

Our ML approach is designed to fulfill two goals:

1. Providing a robust, reproducible, and accurate evaluation of classification performance with the data.
2. Facilitating exploration of the empirical and simulated features that are most important for achieving optimal separation between the AD, MCI, and HC groups.

To satisfy the first goal, we implement a strict nested cross-validation scheme that allows us to obtain statistically reliable classification performance metrics while minimizing overfitting in a P>>N setting (i.e., we have a small sample size N, but a very large number of features P). Our cross-validation method is adapted from earlier work in machine learning for clinical neuroscience (Boshra, Dhindsa et al. 2019), and is described in greater detail below and in **Figure 2**. Furthermore, two classifiers based on different principles (SVM and RF) were used to rule out the possibility that our results could be explained as an artifact of a particular classification algorithm.

We satisfy the second goal in two ways. First, our cross-validation scheme provides a natural metric for feature relevance, i.e., feature selection frequency across cross-validation runs. Besides, we use feature importance metrics inherent to each feature selection method explored. In our case, the F-statistic and the entropy criterion were two metrics used for feature selection for the SVM and the RF respectively.
The entropy criterion uses the notion of entropy from Shannon's Information Theory (Shannon 1948) as a measure of feature importance, i.e., by measuring how well it separates classes.

$$\text{Entropy} = \sum_{i=1}^{C} -f_i \log(f_i)$$

Where f_i is the proportion of class label i that meets a splitting criterion learned for a given feature (e.g., the proportion of class label i for which the feature is less than a particular value), and C is the number of classes.

The F-statistic ranks each feature based on their ANOVA F-value to measure class separability and select the top k features (Conover and Iman 1982).

Since our overall goal is to assess whether the inclusion of features extracted from TVB simulations contributes diagnostic information independent from just the empirical features, we repeat the entire machine learning process with three feature sets: 1) with the empirical features only; 2) with the simulated features only; 3) with both the simulated and empirical features combined.

We used a part of the population with subjects of all groups as a training population while providing specific data aspects to the machine-learning engine. For an overview of machine learning approaches on AD MRI and PET, see (Samper-González, Burgos et al.

2018). Our approach uses a more complex cross-validation with an inner and an outer cross-validation loop.

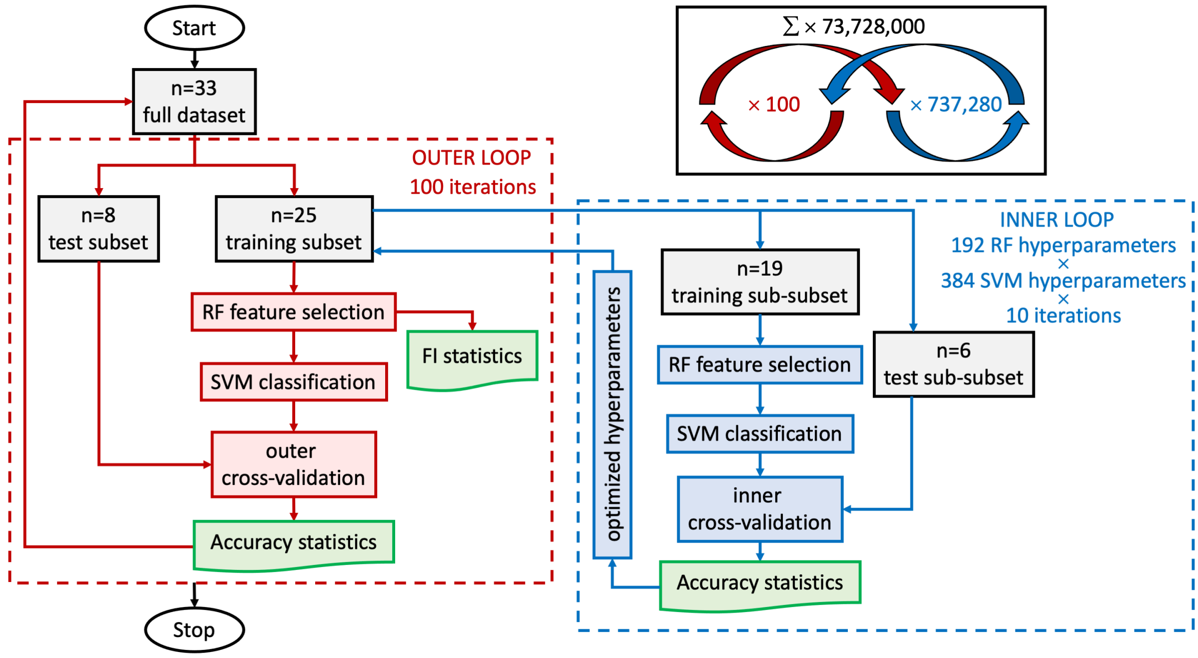

Figure 2. Double cross-validation loop design. Starting in the outer loop: stochastic cross-validation starts with 100 iterations using 25% of data (randomly selected per iteration) for testing. The training subset goes to the inner loop after the train-test split. In the inner loop: split data again just like in the outer loop to obtain training set and validation set for an inner 10 cross-validation iterations with each hyperparameter setting (in total 192 combinations for RF and 384 for SVM, leading to 73,728 combinations with every 10 iterations). Next, we scale training features by subtracting the median and dividing by the inter-quartile range (makes them robust to outliers we identified above). We apply these scaling statistics calculated from the training set also to the test set. Then, we iterate through hyperparameters (**Supplementary Tables 1 and 2**). RF is used for feature selection. Afterward, the remaining features are used for training the SVM classifier with specific hyperparameter settings. We track the selected features for each run and compute the frequency with which they are selected across iterations for the outer loop. The SVM classifications are validated with the test sub-subset (inner cross-validation). This provides optimized hyperparameter settings from the inner cross-validation loop. Back to the outer loop, we recombine training and validation data (which were separated in the inner loop) - still keeping test data separate. We set hyperparameters to the best settings obtained in the inner loop. Then, we train the model and record results: RF is again used for feature selection, which leads to feature importance (FI) statistics used for the results. Afterward, SVM classifies the reaming features, which are then validated with the test set (outer cross-validation). After this, the next iteration of the outer loop begins.

The inner cross-validation loop is the model selection loop wherein hyperparameters can be optimized without influence from the test set to ensure an unbiased estimate of model generalization performance. The outer cross-validation loop is the standard cross-validation loop used to estimate model generalization performance using training and test set partitioning. Using a nested cross-validation loop to separate hyperparameter optimization and a model performance optimization is a well-documented approach to ensuring the validity of machine learning results and preventing overfitting (Stone 1974, Cawley and Talbot 2010). This is particularly necessary for the problem presented in this work, where the number of subjects is low, and the number of features is relatively high. Under such conditions, both overfitting and underfitting due to a poor choice in hyperparameters can significantly affect model performance. Therefore more robust methods like nested cross-validation are required.

Since N<<P and the complexity of the problem is such that we expect the possibility of considerable individual differences among the patient population, the influence of a single data point can be significant. This means learning a classification rule that is nearly optimal for one problem is quite difficult unless hyperparameter tuning is done. Technically, hyperparameter tuning should always be done for a "final model". Still, for more straightforward problems with good data representation, the influence of hyperparameters should not be as large as they are in our situation. So hyperparameter tuning is more a necessity due to the complexity of our problem. The validity of our results comes from the nested cross-validation loop we used, the way we search hyperparameters to prefer less complex and therefore more general solutions (e.g., using fewer features), and the empirical plausibility of the features that were selected. In other words, we performed the necessary step of hyperparameter optimization, but in a cautious and conservative way to ensure the validity of our results.

As mentioned before, we focus here on two ML classifiers: SVM and RF. In the following, we describe their general functionality as well as their application in this work.

2.4. Feature Space

Prior to using the features in the machine learning procedure, we checked them manually for general plausibility. This led us to remove five Volume features from the freesurfer volumetrics outcome: ventricle and white matter hyperintensity measurements (5th Ventricle, left WM hyperintensities, right WM hyperintensities, left non-WM hyperintensities, and right non-WM hyperintensities). For technical reasons, these volumes were not obtained in the image preprocessing.

In the end, we used the following feature spaces:

1. Empirical features (800 dimensions)
    a. 379 values for regional Abeta burden in SUVR, measured from AV-45 PET
    b. One value for the averaged global Abeta burden
    c. 379 values for regional Tau burden in SUVR, measured from AV-1451 PET
    d. One value for the averaged global Tau burden
    e. 40 Volume measures from HCP standards image processing
2. Simulated features (379 dimensions)
    a. 379 regional LFP peak frequencies, averaged over 201 simulations with different scaling factor G
3. Combined features (1179 dimensions)
    a. All features from above

2.5. Nested Cross-Validation Scheme

Our nested cross-validation scheme is illustrated in **Figure 2**. By nesting two cross-validation loops, we are able to simultaneously perform feature selection and model selection robustly. Subjects are portioned into training and test sets in the main outer loop. The training set is then treated as if it were the full dataset in the inner cross-validation loop, where it is split again into a smaller training set and a validation set. In this inner loop, each parameter setting of the classifier (**Supplementary Tables 1 and 2**) undergoes 10-fold cross-validation with feature selection performed independently on each run. These variations on the model can then be compared by their performance against the validation set. The best performing model parameters are then used to train a new classifier on the larger training set defined in the outer loop, with performance measured on the actual test set. The features selected in the outer loop are stored so that the frequency with which each feature is selected can be measured. This entire process is repeated 100 times in order to obtain statistically reliable estimates of our chosen performance metrics and feature importance metrics. This is also why we use random sampling with replacement to partition training and test data since it allows a greater number of cross-validation iterations for statistical evaluation of our models.

2.6. Classifiers

Two types of classifiers that are suitable for small-sample classification problems were used: the kernel-based SVM (Cortes and Vapnik 1995) and the decision-tree based Random Forest (RF) (Breiman 2001). SVMs and RFs operate on fundamentally different principles.

An SVM aims to define a decision-boundary, or a set of boundaries in the case of multi-class classification, which partitions the feature space into class-defining regions. In the very simple case of only two features and two classes, this would mean separating two clusters of points on a two-dimensional plot with a line. It does so in such a way as to maximize the margins between the boundaries and the points nearest to those boundaries, which are referred to as the support vectors. When this is not possible in the original feature space, nonlinearity is introduced via the SVM's kernel function. This function is used to project the data into an arbitrarily high dimensional space wherein the best decision boundary is linear (i.e., a hyperplane). The projection is then reversed so that the decision boundary can be projected back into the original feature space, resulting in a nonlinear boundary.

On the other hand, an RF builds many decision trees based on finding the best partitioning of random subsets of features. Each of those decision trees is constructed by a set of rules, learned from the data, organized in a hierarchy that determines the decision process for classification. So-called hyperparameters of the RF define how the trees should be constructed, e.g., how many layers they have or how many features are involved in each layer. Each data point is classified by each tree based on the path it travels, given the values of each feature. Each tree is then given a 'vote' on how to classify the datapoint. Finally, a classification decision of the RF is made by pooling all of the trees' decisions.

By training two classifiers based on different underlying machine learning mechanisms, we provide more robust evidence that the pattern in classification performance, when combining simulated and empirical features, is reliable and clinically relevant, and that this pattern is driven by a reliably reoccurring subset of the features themselves, rather than by particular mechanisms underlying a classification algorithm.

2.7. Feature Selection Methods

For each classifier, we used an appropriate feature selection method:

2.7.1. SVM feature selection

To select features for the SVM classifier, we ranked features in the training set according to their F-statistic, as in an ANOVA analysis, which estimates the linear dependence of each feature on the class labels. We chose the top k features for varying values of k as part of the model selection phase of our nested cross-validation loop (**Supplementary Table 1** for values).

2.7.2. RF feature selection

RFs incorporate their own embedded feature selection process. Here we used entropy to allow the classifier to rank features by information gain while building trees. The maximum number of allowable features was the square root of the total number of input features, 34.

2.8. Classification Experiments

In order to establish the contribution of our feature sets separately from the contribution of our classification approach, we completed nine experiments organized in a 3x3 grid (three feature sets by three classification approaches). As described above, the three feature sets used are the empirical features, the simulated features, and the combination of empirical and simulated features via concatenation. For our classification approaches, we use the SVM and RF approaches already described, as well as a combined approach. In the combined method, we use the RF for feature selection. Its embedded feature selection approach is well-designed for taking into account the interactions among large numbers of features instead of the univariate method used with the SVM. The SVM is used for classification due to its power in low-sample settings once a robust feature set is preselected, owed in part to its maximum-margin objective.

2.9. Classification Performance Evaluation

We used the weighted F1-score to evaluate classifier performance. This metric is particularly beneficial for classification problems with imbalanced classes, where classification accuracy alone can be misleading (consider a classifier that only predicts Healthy Control with probability 1: it would have a 45.5% classification accuracy, which can be misinterpreted as better than chance for a 3-class problem). The weighted F1-score is computed by taking each class-wise F1-score, which itself is an average of precision and recall for the given class, and averaging those F1-scores together weighted by the number of samples in each class:

$$weighted\ F1 = \frac{1}{N}\sum_{k\in K} fr(C_k)F1_{C_k}$$

where N is the total number of data samples, K is the number of classes represented by

C_k, $fr(C_k)$ is the frequency of class k, and

$$F1 = 2\frac{precision \cdot recall}{precision + recall}$$

For comparison of weighted F1-scores between the groups (empirical data, simulated data, and combined data in the feature space) we used the Wilcoxon Signed-Rank test, as the Shapiro-Wilk test revealed $p < 0.01$ for the empirical and combined approach (non-normal distributed) and $p = 0.07$ for the simulated approach (normal distributed), leading to the usage of a non-parametric test. We assessed the significance by using data from 100 cross-validation runs, leading to 100 data points per group.

2.10. Feature Importance Metrics

We tracked two metrics for feature importance. The most direct feature importance metric is the feature importance statistic used for feature selection itself, i.e., the F-statistic for the SVM and the entropy measure for the RF. In conjunction, we also tracked the selection frequency, defined as the proportion of outer cross-validation iterations in which each feature was selected. We compare and contrast these two metrics and discuss their agreement and differences below.

3. RESULTS

3.1. Data properties and feature selection

In the beginning, we performed basic statistical analysis to characterize the feature space with which we run the classification analysis to assess the quality of the data.

The distribution of simulated LFP frequencies, Abeta SUVR, Tau SUVR, and regional Volumes and their interdependency are shown in **Figure 3**. Abeta ($p = 0.0018$) and Tau SUVR ($p = 0.0005$) are significantly different between AD and HC after Bonferroni-correction. LFP frequency differs significantly between AD and MCI ($p = 0.0316$) but is not significant after Bonferroni-correction. Interestingly, in the existing data, there is no significant difference in the mean volumes between groups.

3.2. Classification results for all experiments

As described before, we were performing nine experiments with different classification schemes and feature sets. The combined classification scheme with SVM and RF performed best. However, in all schemes, it was shown that the combined feature space outperformed both the empirical and the simulated feature space (**Table 1**). A detailed description of the schemes with RF only and SVM only can be found in **Supplementary Tables 1 and 2**. For a more detailed visualization of the results of SVM only and RF only classification, consider **Supplementary Figures 1 and 2**.

Table 1. Classification performance for different experimental designs.

F1-score	SVM	RF	SVM + RF
Empirical features	0.6756	0.6304	0.6434
Simulated features	0.6338	0.6501	0.6607
Combined features	0.7182	0.6699	0.7428

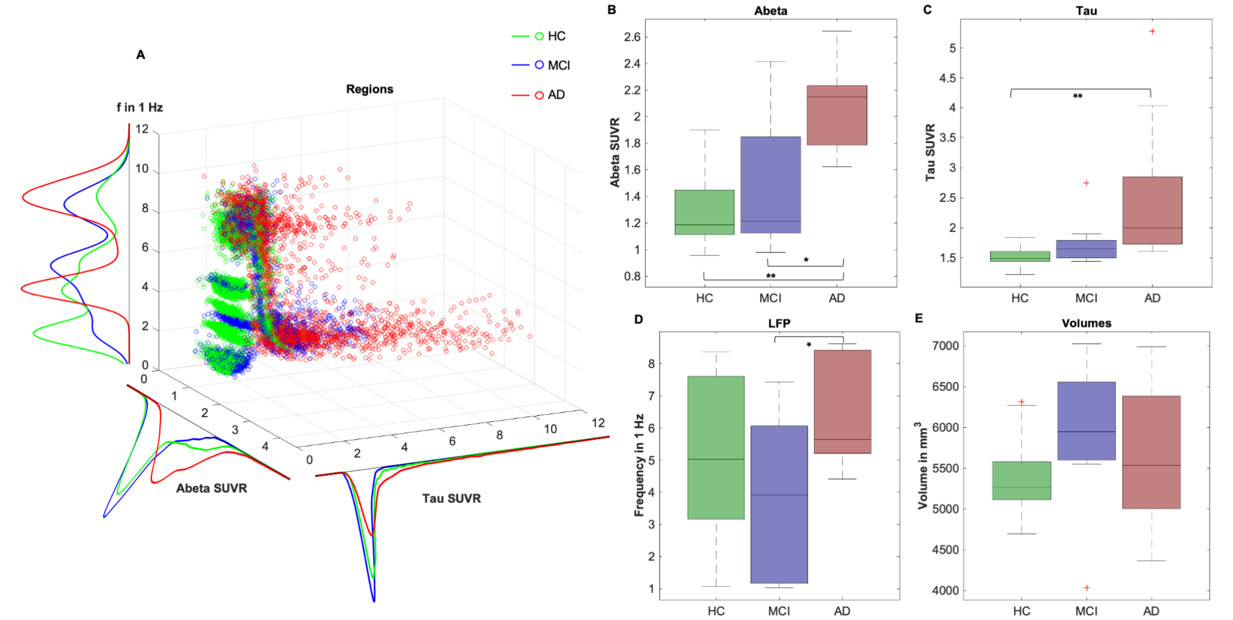

Figure 3. Characteristics of empirical feature space. In (**A**), regional distributions of Abeta, Tau, and LFP frequency are shown for all groups in a 3D scatterplot. Red datapoints symbolize regions of AD patients, green points MCI patients, and blue points HC. Each scatters point stands for one region of one subject. Color density is normalized between groups. A kernel density estimate of the corresponding histograms is shown (projection of the 3D-plot to one axis). In particular, it can be seen a string of outliers with very high Tau values in the AD group and in parts in the MCI group, which does not appear for HC. Moreover, AD participant's regions show higher Abeta values, in particular for lower frequencies. Besides, boxplots are presented for groupwise comparisons for the features mean Abeta per subject, mean Tau per subject, mean simulated LFP frequency per subject, and mean Volume per subject. A Kruskal-Wallis-test was performed to assess significance: * marks significance with $p>0.05$; ** marks significance after Bonferroni-correction with $p < 0.0042$. (**B**) Abeta SUVR is significantly different between AD and HC ($p = 0.0018$) and MCI ($p = 0.0454$), but not between HC and MCI ($p = 0.8113$). (**C**) Tau SUVR is only significantly different between AD and HC ($p = 0.0005$), but not between AD and MCI ($p = 0.1736$) or HC and MCI ($p = 0.2671$). (**D**) LFP frequency is only significantly different between AD and MCI ($p = 0.0316$), but not between AD and HC ($p = 0.2160$) or HC and MCI ($p = 0.4716$). (**E**) The mean Volume of all regions does not show significant differences: AD and MCI ($p = 0.7056$, AD and HC ($p = 0.5102$) or HC and MCI ($p = 0.1405$).

3.3. Classification Using Random Forest and SVM Together

Random Forest was used in the inner cross-validation loop (10 iterations) to perform multivariate feature selection. In the outer cross-validation loop (100 iterations), the features used by the best random forest model were then used to train an SVM.

Classification F1-score and normalized confusion matrices are given in **Figure 4**. The combined approach (0.7428) outperformed the empirical one (0.6434) by about 0.1

(**Figure 4D**), mainly because of an improvement in the classification of the MCI group (**Figure 4A-C**). We used the Wilcoxon Signed-Rank test from 100 cross-validation runs to assess significance (Shapiro-Wilk test of normality for the F1 score distributions revealed $p < 0.01$ for empirical and combined approach and $p=0.07$ for the simulated approach, leading to the usage of a non-parametric test). The differences between the combined approach and both other approaches were highly significant (combined and empirical: $p = 2.52 \times 10^{-11}$; combined and simulated: $p = 1.6 \times 10^{-7}$), meanwhile there was no significant difference between the empirical and simulated approach ($p = 0.34$).

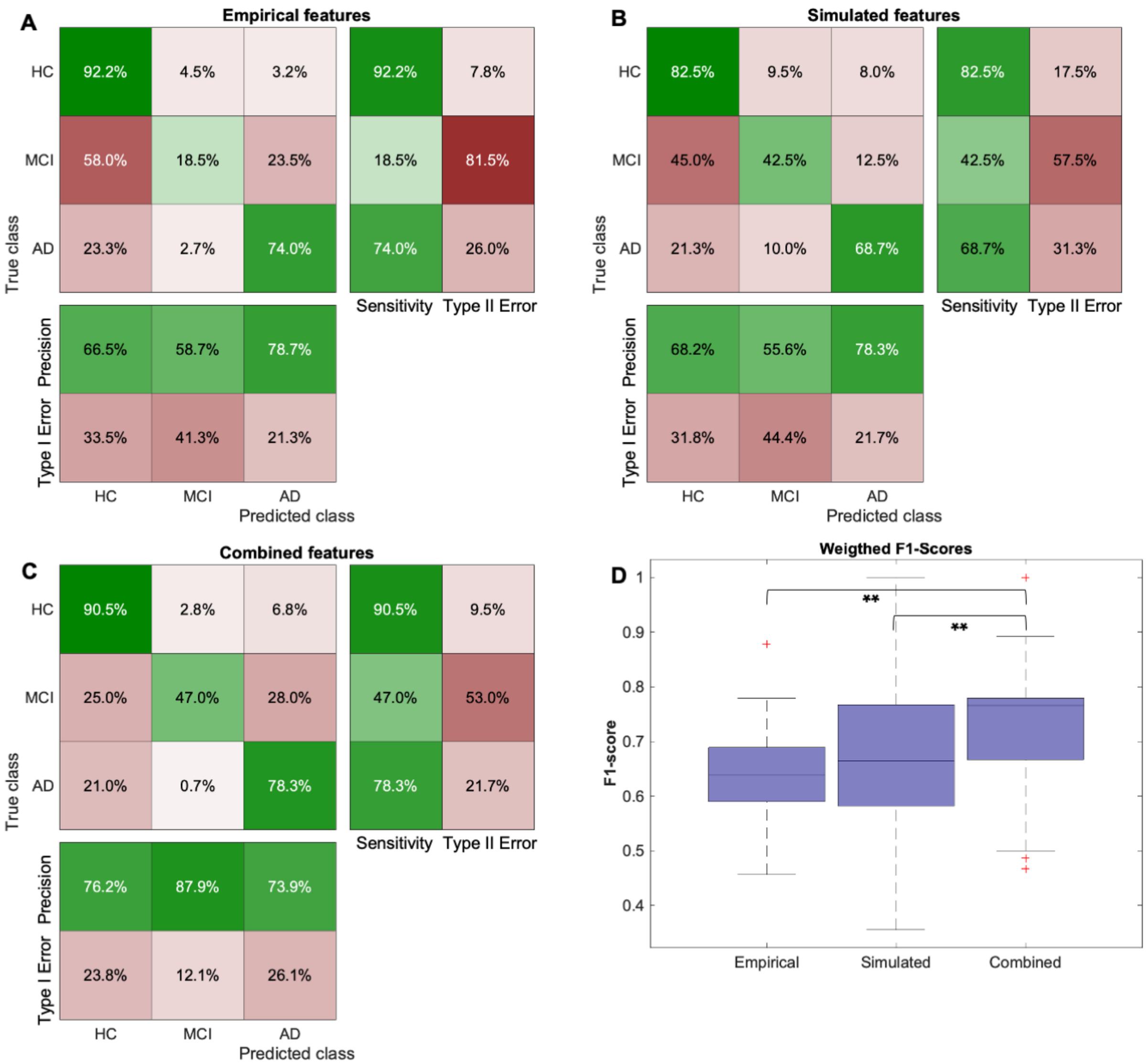

Figure 4. Results of the nested cross-validation classification approach. (**A-C**) Confusion matrices are computed by summing the confusion matrices across all 100 cross-validation runs and normalizing per class. In particular, the combined approach improved the prediction of MCI participants, as AD and HC were already quite well distinguishable by the empirical features. (**D**) Boxplots of mean F1-scores for three different feature spaces. The combined approach (0.7428) outperformed the empirical one (0.6434) by about 0.1. Significance assessment with the Wilcoxon Signed-Rank test from 100 cross-validation runs: combined vs. empirical: $p = 2.52 \times 10^{-11}$; combined vs. simulated: $p = 1.6 \times 10^{-7}$, empirical vs. simulated $p = 0.34$.

As a further analysis to understand this classification improvement, we calculated the feature importance. **Figure 5A** shows the mean entropy-based feature importance given by the Random Forest classifier for 100 outer cross-validation runs. This is used just to show that there is a decreasing curve, as we would expect if meaningful features are found (as opposed to a more uniform distribution). Many of the more important features seem to be biologically plausible in context of AD (**Figure 5B**), as, e.g., Tau in entorhinal cortex (Braak stage 1), frequencies in the thalamus (as important rhythm generator) and putamen, and volumes in putamen and hippocampus (as signs of atrophy). Moreover, we observed that features related to important AD-related functional networks (as the default-mode network (Grothe, Teipel et al. 2016) and the visual ventral stream (Grill-Spector and Malach 2004, Kravitz, Saleem et al. 2013)) are involved with a higher feature importance rank. A table with the full name and corresponding functional network according to Rosen and Halgren (Rosen and Halgren 2020) for the 50 highest ranked features is given in **Table 2**.

We also showed that feature relevance is dependent on the structural degree of the regions in the underlying SC network (**Figure 5C**). This is an indicator of network effects contributing to the improved classification and another indicator for meaningful classification results.

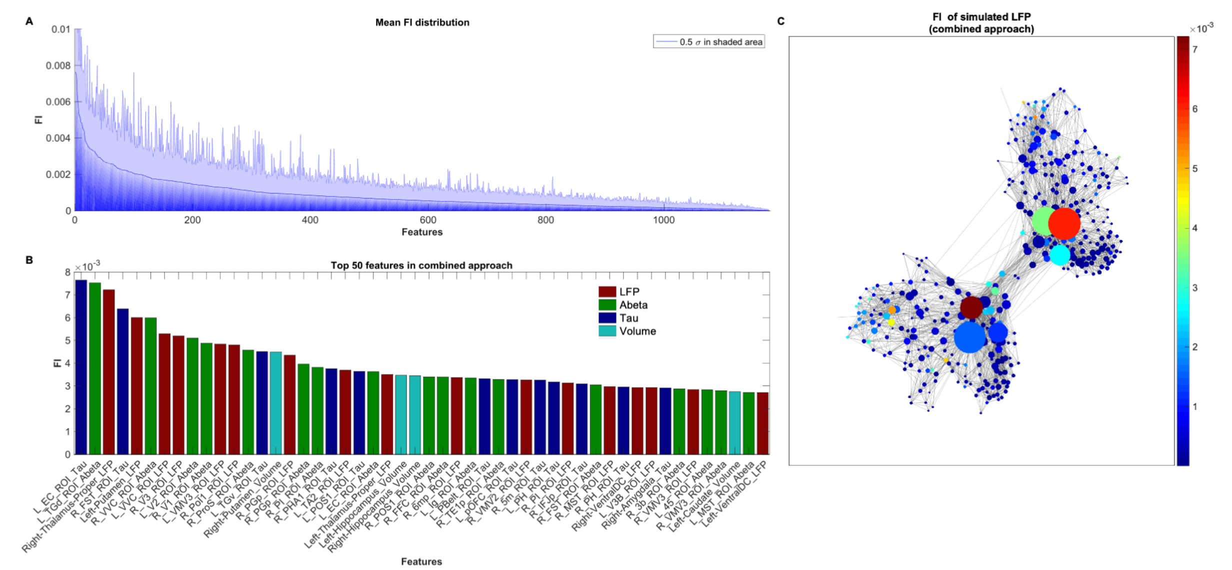

Figure 5. Feature importance (FI) distribution. (**A**) Mean RF-derived feature importance from 100 outer cross-validation runs. Entropy criterion with combined feature types shown here. Feature importance values are normalized, so all features sum to 1. In shaded blue, half standard deviation σ is displayed for each feature. (**B**) Top 50 features across all cross-validation runs. Both empirical (Tau in dark blue, Abeta in green, Volume in light blue), as well as simulated frequencies (red), contributed to the improved classification. Many features seem moreover to be biologically plausible in the context of AD, e.g., Tau in the entorhinal cortex, thalamic and putaminal frequencies, and putaminal as well as hippocampal volumes. (**C**) Visualization of the SC graph with color indicating FI of the regional LFP frequencies, while vertex diameter reflects the structural degree. It shows a network dependency of the LFP FI. Only edges with connection strength above the 95th percentile are shown.

4. DISCUSSION

In this study, we show that the involvement of virtual, simulated TVB features into ML classification can lead to an improved classification between HC, MCI, and AD.

The diagnostic value of the underlying empirical features can be improved by integrating the features into a multi-scale brain simulation framework in TVB. We showed for ML algorithms a superiority of feature sets that contain both the empirical and the virtual derived metrics. The absolute gain of accuracy was 10 %. Keeping in mind that all differences between the subjects have to be derived from their amyloid PET signal (because all other factors, e.g. the underlying SC, are the same) this provides evidence that TVB is able to decode the information that is contained in empirical data like the amyloid PET. More specific for the PET and its usage in diagnostics, it highlights the relevance of spatial distribution, which is often not considered in its analysis.

The main reason for this improvement seems to be a better classification of MCI subjects. without the simulated features, the models frequently misclassify MCI subjects as HC.

In contrast, the simulated features alone result in more misclassification of healthy controls as either MCI or AD subjects compared to using the empirical features alone. However, combining both the empirical features with the simulated features appears to add their strengths in a clinically beneficial way; these models retain all or most of the ability to correctly classify healthy controls with the empirical Features and retain much of the ability of the simulated Features to classify MCI patients. The processing inside TVB seems to reorganize the existing data beneficially.

In theory, a larger number of available features could provide a machine learning algorithm greater flexibility in finding useful combinations, simply due to there being a higher degree of freedom during feature selection and weighting. However, the equal empirical data source in combination with the used nested cross-validation method protects from an overfitting bias due to the larger feature space. If the explanation for the improvement in classification accuracy was simply the presence of additional noisy features, we would see a flatter feature importance distribution than shown in **Figure 5**, and therefore a more random distribution of selected features across the 100 cross-validation iterations. Instead, we see that only a few features with high importance are consistently guiding classification, indicating that they in fact provide useful discriminative information. Preventing this kind of overfitting via feature selection is a key motivation behind our use of the nested cross-validation approach (**Figure 2**): since the features are selected on the training and validation (test) set in the inner loop, any overfitting due to feature selection should not be transferred to the test set in the outer loop.

We have shown that only a few, selected features seem to play a crucial role in classification throughout the cross-validation iterations (**Figure 5**). In the following, we will further evaluate the plausibility of these features with a detailed interpretation of these features in a biological context.

Following, we take a closer look at the 50 top features (which represent 4.24% of the total) comprised of 33 empirical features (17 Abeta, 12 Tau, 4 Volume) and 17 simulated features (**Table 2**). The anatomical distribution of feature relevance is summarized in **Figure 6**, for each, Abeta, Tau, and LFP. We will therefore elucidate in how far the top features are biologically plausible in context of AD, and if this offers a possible explanation for the improved classification performance with simulated features.

Table 2. Full name description and functional network association of the 50 top features with the highest feature importance in the classification problem. Parcellation adapted from (Glasser, Coalson et al. 2016). Functional networks adapted from (Rosen and Halgren 2020). DS: Dorsal Stream, VS: Ventral Stream, SMA: Supplementary motor area

Rank	Feature Name	Full-Parcel-Name	Network
1	L_EC_ROI_Tau:	Entorhinal cortex	Default mode
2	L_TGd_ROI_Abeta	Dorsal temporal gyrus	Default mode
19	R_PHA1_ROI_Tau	Parahippocampal area 1	Default mode
21	L_POS1_ROI_Tau	Parieto-occipital sulcus area 1	Default mode
22	L_EC_ROI_Abeta	Entorhinal cortex	Default mode
26	R_POS1_ROI_Abeta	Parieto-occipital sulcus area 1	Default mode
32	L_pOFC_ROI_Tau	Posterior orbitofrontal cortex	Default mode
4	R_FST_ROI_Tau	Fundus of superior temporal sulcus	Visual
8	R_V3_ROI_LFP	Visual area 3	Visual
9	L_V2_ROI_Abeta	Visual area 2	Visual
10	R_V1_ROI_Abeta	Visual area 1	Visual
13	R_ProS_ROI_Abeta	Prostriate region	Visual
38	R_FST_ROI_Abeta	Fundus of superior temporal sulcus	Visual
39	R_MST_ROI_LFP	Medial superior temporal area	Visual
49	L_MST_ROI_Abeta	Medial superior temporal area	Visual
6	R_VVC_ROI_Abeta	Ventral visual complex	Visual (VS)
7	L_VVC_ROI_LFP	Ventral visual complex	Visual (VS)
11	L_VMV3_ROI_LFP	Ventromedial visual complex 3	Visual (VS)
27	R_FFC_ROI_Abeta	Fusiform face complex	Visual (VS)
33	R_VMV2_ROI_LFP	Ventromedial visual area 2	Visual (VS)
35	L_PH_ROI_Tau	Area PH in lateral occipital lobe	Visual (VS)
40	R_PH_ROI_Tau	Area PH in lateral occipital lobe	Visual (VS)
45	R_VMV3_ROI_LFP	Ventromedial visual complex 3	Visual (VS)
47	R_VMV3_ROI_Abeta	Ventromedial visual complex 3	Visual (VS)
42	L_V3B_ROI_LFP	Visual area 3b	Visual (DS)
16	R_PGp_ROI_LFP	Parietal area G posterior	Dorsal attention
17	R_PGp_ROI_Abeta	Parietal area G posterior	Dorsal attention
12	R_PoI1_ROI_LFP	Posterior insula 1	Cingulo-opercular
18	R_PI_ROI_Abeta	Parainsular cortex	Cingulo-opercular
36	R_PI_ROI_LFP	Parainsular cortex	Cingulo-opercular
20	L_TA2_ROI_LFP	Temporal region A	Auditory

30	L_PBelt_ROI_Tau	Parabelt complex (Auditory cortex)	Auditory
28	R_6mp_ROI_LFP	Area 6 medial posterior (SMA)	Somatomotor
29	L_Ig_ROI_Abeta	Insula granular cortex	Somatomotor
34	R_5m_ROI_Tau	Area 5 medial of paracentral lobule	Somatomotor
44	R_3b_ROI_Abeta	Area 3b of postcentral gyrus	Somatomotor
14	L_TGv_ROI_Tau	Ventral temporal gyrus	Language
46	L_45_ROI_Abeta	Area 45 of inferior frontal gyrus	Language
31	R_TE1p_ROI_Abeta	Temporal area 1 posterior	Frontoparietal
37	R_IFJp_ROI_Tau	Inferior frontal junction posterior	Frontoparietal
3	Right-Thalamus-Proper_LFP	Thalamus proper	Subcortical
5	Left-Putamen_LFP	Putamen	Subcortical
23	Left-Thalamus-Proper_LFP	Thalamus proper	Subcortical
41	Right-VentralDC_LFP	Ventral diencephalon	Subcortical
43	Right-Amygdala_Tau	Amygdala	Subcortical
50	Left-VentralDC_LFP	Ventral diencephalon	Subcortical
15	Right-Putamen_Volume	Putamen	Volume
24	Left-Hippocampus_Volume	Hippocampus volume	Volume
25	Right-Hippocampus_Volume	Hippocampus volume	Volume
48	Left-Caudate_Volume	Nucleus caudatus	Volume

Most of the tau top features can be allocated to the temporal lobule, which is also the location of early Tau deposition according to the neuropathological BRAAK and Braak stages I-III (Braak and Braak 1991, Braak and Braak 1997, Braak, Alafuzoff et al. 2006) and the location of increased in-vivo binding of ^{18}F-AV-1451 in AD (Cho, Choi et al. 2016, Hansson, Grothe et al. 2017). In particular, the entorhinal cortex is a consistent starting point of the sequential spread of Tau through the brain (Braak and Braak 1991, Cho, Choi et al. 2016) and also showed the most robust relationship between flortaucipir and memory scores in a recent machine learning study (Knopman, Lundt et al. 2019).

The Abeta Top features showed a more disseminated allocation mostly in the temporal, occipital, frontal, and insular cortices, which is also in line with typical amyloid deposition and locations of increased AV45 uptake in AD (Thal, Rüb et al. 2002, Grothe, Barthel et al. 2017). Interestingly there is some regional overlap between Abeta and Tau top-features, namely the left entorhinal cortex and the right fundus of the superior temporal sulcus. This overlap could represent a potential synergistic effect between Abeta and Tau in these regions. Synergism effects between Abeta and Tau in the temporal lobe are an important element in the progression from MCI to AD. They have been described in voxel-based analyses, as well as in molecular studies (Vossel, Zhang et al. 2010, Khan, Liu et al. 2014, Pascoal, Mathotaarachchi et al. 2017). In a study by Halawa, Gatchel et al. (2019) both temporal tau and amyloid burden were associated with an impairment of activities of daily living. Still, the combination of both pathologic markers showed an association, that by far surpassed the mere additive effect.

As atrophy is a known phenomenon in NDD, we interpret the volume loss as such. The volume top features, derived from structural MRI, correspond to the most prominent atrophic regions in AD. Both hippocampal (Jack, Petersen et al. 1997, Killiany, Gomez-Isla et al. 2000), as well as striatal atrophy (Madsen, Ho et al. 2010), have been described as prognostic markers of AD. Volumes of bilateral hippocampi and the right putamen were related to a loss of memory function (Jack, Petersen et al. 2000, Zhao, Li et al. 2015).

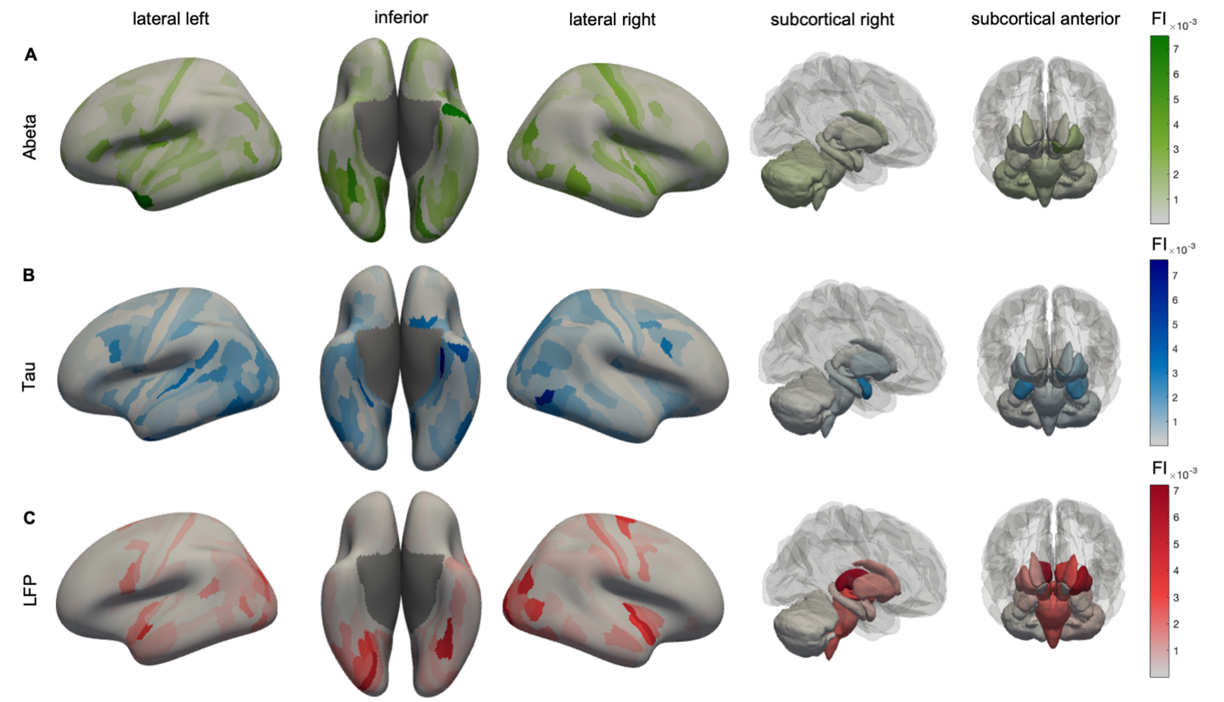

Figure 6. Anatomical representation of feature importance (FI) distribution. Displayed are cortical regions from left, right, and inferior as well as subcortical regions. The color indicates the feature importance. **(A)** Abeta FI. The anatomical patterns reveal high importance of left-temporal regions, as well as the left dorsal stream in the parietal and occipital cortex. (**B**) Tau features show a similar pattern as Abeta, but with a higher focus on typical Braak-stage-1 regions (as the entorhinal cortex). (**C**) Simulated frequencies do not show strong laterality as the empirical features but seem to have a focus in both occipital lobes, where typically alpha oscillations occur. Besides this, subcortical areas like the thalami play a more crucial role than for Abeta and Tau. For more detailed information, see the discussion of this paper and **Table 2**.

The interpretation of the simulated features is more complicated and needs to be treated carefully. While LFP and EEG differ in size and location of source recording, the prevalent model is that both are generated by synchronized synaptic currents of cortical pyramidal cells, and both show the same type of oscillations during states of wakefulness and sleep (Steriade 2003, Nunez and Srinivasan 2005, Nunez and Srinivasan 2006). Hence it is reasonable to assume that AD typical disturbances in EEG signal can in some part be related to LFP and vice versa.

In general, the simulated features depend on network information. Concurrently, in this work, all subjects used the same SC. As a consequence, the difference between subjects can be attributed to the spatial distribution of their respective Abeta PET.

Many of the top simulated features can be allocated to two functional areas. Firstly, the visual cortex including the ventral and dorsal stream, and, secondly, subcortical structures like the thalamus and the putamen.

The occipito-temporal and occipito-parietal regions of the first area are typical alpha-rhythm generators in resting-state EEG (Barzegaran, Vildavski et al. 2017). Alteration of these posterior alpha sources is a typical phenomenon in AD and MCI compared to HC (Babiloni, Del Percio et al. 2018). These alterations could represent an alteration of memory processes in the visuospatial and episodic compartment and were also used to classify between MCI and HC (Babiloni, Del Percio et al. 2018, Smailovic and Jelic 2019).

The ventral or 'what' stream and the dorsal or 'where' stream have been implicated in object recognition and spatial localization and are essential for accurate visuospatial navigation (Grill-Spector and Malach 2004, Kravitz, Saleem et al. 2013). Impairment in visuospatial navigation is a potential cognitive marker in early AD/MCI that could be more specific than episodic memory or attention deficits (Coughlan, Laczó et al. 2018, Williams, An et al. 2020). This impairment seems to be caused by a complex interplay of Abeta and Tau burden and disturbances in the underlying functional network, which is represented well by our top features. Higher Abeta-burden in older adults was associated with decreased neural activity in the ventral visual stream (Rieck, Rodrigue et al. 2015) and Abeta-PET positive MCI patients perform worse in a navigation task than Abeta-PET negative MCI patients (Schöberl, Pradhan et al. 2020). Other essential contributors to spatial navigation in the brain's navigation circuit are the grid cells in the entorhinal cortex, which is primarily affected by tau deposition (Rowland, Roudi et al. 2016). Tau-related disruption of grid-cell firing in the entorhinal cortex of transgenic mice expressing human tau leads to increased theta oscillations that were associated with spatial memory deficits (Fu, Rodriguez et al. 2017). In humans, the performance in an entorhinal cortex-based immersive virtual reality navigation task helped to differentiate between MCI patients at low and high risk of developing AD (Howett, Castegnaro et al. 2019).

On a network level, theta-band oscillations in high-density EEG in a network consisting of the temporal lobe, striatum, inferior occipital lobe, and cerebellum correlated with the performance in a spatial memory task (Bauer, Buckley et al. 2020) and the reduction in visual processing network complexity correlated with the stage of AD (Huang, Beach et al. 2020). Interestingly, a recent review focused on spatial navigation measured with LFPs revealed that the neural representation of spatial features occurs on a mesoscopic level of mainly theta oscillations (Kunz, Maidenbaum et al. 2019) and thus the same level that TVB operates (Ritter, Schirner et al. 2013).

In recent years there has been considerable evidence that thalamic dysfunction. Mainly the anterior thalamus has been identified to play a crucial role in early disease progression (Aggleton, Pralus et al. 2016). The anterior Thalamus is part of the Papez circuit and as such strongly connected to episodic memory (Aggleton 2014). While the relevance of the (unparcellated) thalamus is represented by the simulated features, it did not play an important role as an empirical feature. This disparity makes sense for Tau, considering that only some of the thalamic nuclei have been shown to contain NFT even in late stages of the disease (Rüb, Stratmann et al. 2016) and therefore a strong PET-signal is unlikely. In MCI patients the thalamus displays both increased and decreased functional connectivity, which could be a hint towards loss of function or compensation mechanisms (Cai, Huang et al. 2015). A change in the LFP of the thalamus supports a shift in function, as different frequency bands can be attributed to various brain functions (Buzsáki and Draguhn 2004). According to Schnitzler and Gross (2005), faster rhythmic activities correspond to more local neuronal communication, while slower oscillations likely arise from larger populations within wider-range networks. A change in the function of the thalamus and other subcortical hubs could therefore have possible network effects detectable in LFP/EEG. Reduced complexity of the EEG signal and perturbations in EEG synchrony are significant effects of AD on EEG (Dauwels, Vialatte et al. 2010, Rossini, Di Iorio et al. 2020).

The default mode network (DMN) was exclusively represented by empirical top features. The DMN is consistently affected in AD (Grothe, Teipel et al. 2016) and shows the earliest accumulation of Abeta (Palmqvist, Schöll et al. 2017). Nevertheless, the empirical features poorly performed when classifying MCI, which can in some cases be interpreted as a prodromal stage of AD (Dubois, Feldman et al. 2010). This could partly be explained by the missing network information of the empirical features. The breakdown of functional connectivity of the DMN occurs concurrently with Abeta deposition in preclinical subjects (Veitch, Weiner et al. 2019) and in the early stage of MCI (Lee, Yoo et al. 2016). It is possible that network disturbances, detectable by the simulated features, occur even earlier than Abeta deposition in the DMN, but simulated features do not represent the DMN. Alves et al. (Alves, Foulon et al. 2019) proposed a new neuroanatomical model of the DMN that includes subcortical structures like the thalamus, which is represented by the simulated top features. Following this approach, the combination of empirical and simulated top features yields a more comprehensive description of this functional network that is an important marker in the disease progression of AD.

Network disturbances (as seen in the subcortical hub regions) may come apparent earlier than stronger Abeta or Tau uptake in the PET-Imaging leading to a better classification of the MCI group. At the same time, simulated features only use the amyloid PET and cannot get the information from the Tau-PET, which seems to better correlate with the cognitive decline of patients (Wennberg, Whitwell et al. 2019). This missing information could therefore be a reason for the misclassification of HC by the simulated features. Hence, it would be of great interest how a simulation using Tau-PET performs in this classification problem.

As a limitation of our study, we see that the used simulated feature, the mean simulated LFP frequency (averaged across a wide range of the large-scale coupling parameter G), is not directly equivalent to a biophysical measurement like empirically measured LFP. LFP frequency was averaged across the global scaling factor G. G scales the strength of long-range connections in the brain network model and is a crucial factor in the simulation. Across the dimension of G, many different dynamics can develop, from which some are similar to empirically observed phenomena, but others are not. In our former work, we found that particular areas of G with non-plausible frequency patterns hold the potential to differentiate between diagnosis groups (Stefanovski, Triebkorn et al. 2019). This is mainly because of the underlying mathematics of the Jansen-Rit model: besides two limit cycles that are producing alpha-like and theta-like activity, the local dynamic model has an area of stable focus, wherein no oscillations are produced in the absence of noise. Technically, this stable focus is represented as a zero-line artifact that appears mainly in the HC group, because only Abeta values above a critical value led to the presence of the slower theta-limit-cycle. By averaging LFP frequencies across the whole spectrum of G, we incorporate this zero-line information, which leads to apparently higher mean LFP frequencies for the AD group. In contrast, in the area of plausible results, AD has lower frequencies – as it would be expected (Stefanovski, Triebkorn et al. 2019). This can also be seen as another advantage of TVB. It shows how TVB is not just reproducing data that we could also obtain with EEG or intracranial electrodes but delivers 'artificial' data that is still informative.

Another limitation of this study is the limited population size and heterogeneity. Although our results were obtained undergoing careful cross-validation, future studies will have to reproduce the results using a more extensive cohort. Ideally, external validation with a dataset outside of ADNI could be performed.

We used ML as an approach for the comparison of classifier performance with empirical data against simulated data, which is wholly derived from the empirical data. Improvement in classification is then strong evidence for successful processing of the empirical data in TVB – TVB decodes the information, that is embedded within the empirical data, that cannot be detected by statistics or ML classifiers. We showed in ADNI data that TVB can derive additional information out of the spatial distribution pattern in PET images.

Our work provides novel evidence that TVB can act as a biophysical brain model - and not just like a *black box*. Complex multi-scale brain simulation in TVB can lead to additional information, that goes beyond the implemented empirical data. Our analysis of feature importance supports this hypothesis, as the features with the highest relevance are already well-known AD factors and hence, biologically plausible surrogates for clinically relevant information in the data. Moreover, in this pilot study, we demonstrate that TVB simulation can lead to an improved diagnostic value of empirical data and might become a clinically relevant tool.

AUTHOR CONTRIBUTIONS.

All authors have made substantial intellectual contributions to this work and approved it for publication. PT and LS had equal contributions to this work. Particular roles according CrediT (Brand, Allen et al. 2015):

Paul Triebkorn: Conceptualization, Data curation, Investigation, Methodology, Visualization, Writing - original draft : **Leon Stefanovski**: Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing - original draft. **Kiret Dhindsa**: Formal analysis, Methodology, Software, Writing - review & editing. **Margarita-Arimatea Diaz-Cortes**: Methodology, Software, Writing - review & editing, **Patrik Bey**: Methodology, Software, Writing - review & editing. **Konstantin Bülau**: Validation, Writing - review & editing. **Roopa Pai**: Data curation, Writing - review & editing. **Andreas Spiegler**: Methodology, Writing - review & editing. **Ana Solodkin**: Writing - review & editing. **Viktor Jirsa**: Writing - review & editing. **Anthony Randal McIntosh**: Writing - review & editing. **Petra Ritter**: Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing - review & editing.

DATA AVAILABILITY.

The raw data for this study is available in ADNI. The codes used in this study are available on request to the corresponding author.

ETHICS STATEMENT.

The study providing the data for this work has been approved from the Ethics Board of the Charité - Universitätsmedizin Berlin under the approval number EA2/100/19.

The work described in this article has been carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans (http://www.wma.net/en/30publications/10policies/b3/index.html; the EC Directive 86/609/EEC for animal experiments (http://ec.europa.eu/environment/chemicals/lab_animals/legislation_en.html); and the Uniform Requirements for manuscripts submitted to Biomedical journals (http://www.icmje.org).

CONFLICT OF INTEREST STATEMENT.

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ACKNOWLEDGMENTS.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Computation of underlying data has been performed on the HPC for Research cluster of the Berlin Institute of Health.

FUNDING.

PR acknowledges support by EU H2020 Virtual Brain Cloud 826421, Human Brain Project SGA2 785907; Human Brain Project SGA3 945539, ERC Consolidator 683049; German Research Foundation SFB 1436 (project ID 425899996); SFB 1315 (project ID 327654276); SFB 936 (project ID 178316478; SFB-TRR 295 (project ID 424778381); SPP Computational Connectomics RI 2073/6-1, RI 2073/10-2, RI 2073/9-1; Berlin Institute of Health & Foundation Charité, Johanna Quandt Excellence Initiative.

REFERENCES.

Aggleton, J. P. (2014). "Looking beyond the hippocampus: old and new neurological targets for understanding memory disorders." Proceedings of the Royal Society B: Biological Sciences **281**(1786): 20140565.

Aggleton, J. P., A. Pralus, A. J. Nelson and M. Hornberger (2016). "Thalamic pathology and memory loss in early Alzheimer's disease: moving the focus from the medial temporal lobe to Papez circuit." Brain **139**(Pt 7): 1877-1890.

Aksu, Y., D. J. Miller, G. Kesidis, D. C. Bigler and Q. X. Yang (2011). "An MRI-derived definition of MCI-to-AD conversion for long-term, automatic prognosis of MCI patients." PLoS One **6**(10): e25074.

Alves, P. N., C. Foulon, V. Karolis, D. Bzdok, D. S. Margulies, E. Volle and M. Thiebaut de Schotten (2019). "An improved neuroanatomical model of the default-mode network reconciles previous neuroimaging and neuropathological findings." Communications Biology **2**(1): 370.

Alzheimer's Association (2019). "2019 Alzheimer's disease facts and figures." Alzheimer's & Dementia **15**(3): 321-387.

Babiloni, C., C. Del Percio, R. Lizio, G. Noce, S. Lopez, A. Soricelli, R. Ferri, M. T. Pascarelli, V. Catania, F. Nobili, D. Arnaldi, F. Famà, D. Aarsland, F. Orzi, C. Buttinelli, F. Giubilei, M. Onofrj, F. Stocchi, L. Vacca, P. Stirpe, P. Fuhr, U. Gschwandtner, G. Ransmayr, H. Garn, L. Fraioli, M. Pievani, G. B. Frisoni, F. D'Antonio, C. De Lena, B. Güntekin, L. Hanoğlu, E. Başar, G. Yener, D. D. Emek-Savaş, A. I. Triggiani, R. Franciotti, J. P. Taylor, M. F. De Pandis and L. Bonanni (2018). "Abnormalities of Resting State Cortical EEG Rhythms in Subjects with Mild Cognitive Impairment Due to Alzheimer's and Lewy Body Diseases." Journal of Alzheimer's Disease **62**: 247-268.

Barzegaran, E., V. Y. Vildavski and M. G. Knyazeva (2017). "Fine Structure of Posterior Alpha Rhythm in Human EEG: Frequency Components, Their Cortical Sources, and Temporal Behavior." Scientific Reports **7**(1): 8249.

Bauer, M., M. G. Buckley and T. Bast (2020). "Individual differences in theta-band oscillations in a spatial memory network revealed by EEG predict rapid place learning." bioRxiv: 2020.2006.2005.134346.

Beheshti, I., N. Maikusa, M. Daneshmand, H. Matsuda, H. Demirel and G. Anbarjafari (2017). "Classification of Alzheimer's Disease and Prediction of Mild Cognitive Impairment Conversion Using Histogram-Based Analysis of Patient-Specific Anatomical Brain Connectivity Networks." J Alzheimers Dis **60**(1): 295-304.

Binkhonain, M. and L. Zhao (2019). "A review of machine learning algorithms for identification and classification of non-functional requirements." Expert Systems with Applications: X **1**: 100001.

Blennow, K., M. J. de Leon and H. Zetterberg (2006). "Alzheimer's disease." Lancet **368**(9533): 387-403.

Blinowska, K. J., F. Rakowski, M. Kaminski, F. De Vico Fallani, C. Del Percio, R. Lizio and C. Babiloni (2017). "Functional and effective brain connectivity for discrimination between Alzheimer's patients and healthy individuals: A study on resting state EEG rhythms." Clinical Neurophysiology **128**(4): 667-680.

Bloom, G. S. (2014). "Amyloid-β and tau: The trigger and bullet in alzheimer disease pathogenesis." JAMA Neurology **71**(4): 505-508.

Bojak, I., T. F. Oostendorp, A. T. Reid and R. Kotter (2010). "Connecting mean field models of neural activity to EEG and fMRI data." Brain Topogr **23**(2): 139-149.

Boshra, R., K. Dhindsa, O. Boursalie, K. I. Ruiter, R. Sonnadara, R. Samavi, T. E. Doyle, J. P. Reilly and J. F. Connolly (2019). "From group-level statistics to single-subject prediction: machine learning detection of concussion in retired athletes." IEEE Transactions on Neural Systems and Rehabilitation Engineering **27**(7): 1492-1501.

Braak, H., I. Alafuzoff, T. Arzberger, H. Kretzschmar and K. Del Tredici (2006). "Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry." Acta Neuropathol **112**(4): 389-404.

Braak, H. and E. Braak (1991). "Neuropathological stageing of Alzheimer-related changes." Acta Neuropathol **82**(4): 239-259.

Braak, H. and E. Braak (1997). "Frequency of stages of Alzheimer-related lesions in different age categories." Neurobiol Aging **18**(4): 351-357.

Brand, A., L. Allen, M. Altman, M. Hlava and J. Scott (2015). "Beyond authorship: attribution, contribution, collaboration, and credit." Learned Publishing **28**(2): 151-155.

Breiman, L. (2001). "Random Forests." Machine Learning **45**(1): 5-32.

Buzsáki, G. and A. Draguhn (2004). "Neuronal Oscillations in Cortical Networks." Science **304**(5679): 1926-1929.

Cai, S., L. Huang, J. Zou, L. Jing, B. Zhai, G. Ji, K. M. von Deneen, J. Ren, A. Ren and I. for the Alzheimer's Disease Neuroimaging (2015). "Changes in Thalamic Connectivity in the Early and Late Stages of Amnestic Mild Cognitive Impairment: A Resting-State Functional Magnetic Resonance Study from ADNI." PLOS ONE **10**(2): e0115573.

Cawley, G. C. and N. L. Talbot (2010). "On over-fitting in model selection and subsequent selection bias in performance evaluation." The Journal of Machine Learning Research **11**: 2079-2107.

Cho, H., J. Y. Choi, M. S. Hwang, J. H. Lee, Y. J. Kim, H. M. Lee, C. H. Lyoo, Y. H. Ryu and M. S. Lee (2016). "Tau PET in Alzheimer disease and mild cognitive impairment." Neurology **87**(4): 375-383.

Conover, W. J. and R. L. Iman (1982). "Analysis of covariance using the rank transformation." Biometrics: 715-724.

Cortes, C. and V. Vapnik (1995). "Support-vector networks." Machine learning **20**(3): 273-297.

Coughlan, G., J. Laczó, J. Hort, A.-M. Minihane and M. Hornberger (2018). "Spatial navigation deficits — overlooked cognitive marker for preclinical Alzheimer disease?" Nature Reviews Neurology **14**(8): 496-506.

Cui, Y., B. Liu, S. Luo, X. Zhen, M. Fan, T. Liu, W. Zhu, M. Park, T. Jiang and J. S. Jin (2011). "Identification of conversion from mild cognitive impairment to Alzheimer's disease using multivariate predictors." PLoS One **6**(7): e21896.

Cuingnet, R., E. Gerardin, J. Tessieras, G. Auzias, S. Lehéricy, M.-O. Habert, M. Chupin, H. Benali and O. Colliot (2011). "Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database." NeuroImage **56**(2): 766-781.

Dauwels, J., F. Vialatte and A. Cichocki (2010). "Diagnosis of Alzheimers Disease from EEG Signals: Where Are We Standing?" Current Alzheimer Research **7**(6): 487-505.

Dhall, D., R. Kaur and M. Juneja (2020). Machine Learning: A Review of the Algorithms and Its Applications. Proceedings of ICRIC 2019, Springer**:** 47-63.

Dubois, B., H. H. Feldman, C. Jacova, J. L. Cummings, S. T. Dekosky, P. Barberger-Gateau, A. Delacourte, G. Frisoni, N. C. Fox, D. Galasko, S. Gauthier, H. Hampel, G. A. Jicha, K. Meguro, J. O'Brien, F. Pasquier, P. Robert, M. Rossor, S. Salloway, M. Sarazin, L. C. de Souza, Y. Stern, P. J. Visser and P. Scheltens (2010). "Revising the definition of Alzheimer's disease: a new lexicon." Lancet Neurol **9**(11): 1118-1127.

Falahati, F., E. Westman and A. Simmons (2014). Multivariate Data Analysis and Machine Learning in Alzheimer's Disease with a Focus on Structural Magnetic Resonance Imaging.

Fan, Y., N. Batmanghelich, C. M. Clark and C. Davatzikos (2008). "Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline." NeuroImage **39**(4): 1731-1743.

Farina, F. R., D. D. Emek-Savaş, L. Rueda-Delgado, R. Boyle, H. Kiiski, G. Yener and R. Whelan (2020). "A comparison of resting state EEG and structural MRI for classifying Alzheimer's disease and mild cognitive impairment." NeuroImage **215**: 116795.

Ferri, R., C. Babiloni, V. Karami, A. I. Triggiani, F. Carducci, G. Noce, R. Lizio, M. T. Pascarelli, A. Soricelli, F. Amenta, A. Bozzao, A. Romano, F. Giubilei, C. Del Percio, F. Stocchi, G. B. Frisoni, F. Nobili, L. Patanè and P. Arena (2020). "Stacked autoencoders as new models for an accurate Alzheimer's disease classification support using resting-state EEG and MRI measurements." Clinical Neurophysiology.

Fischl, B., D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen and A. M. Dale (2002). "Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain." Neuron **33**(3): 341-355.

Forlenza, O. V., B. S. Diniz, A. L. Teixeira, F. Stella and W. Gattaz (2013). "Mild cognitive impairment (part 2): biological markers for diagnosis and prediction of dementia in Alzheimer's disease." Brazilian Journal of Psychiatry **35**(3): 284-294.

Fu, H., G. A. Rodriguez, M. Herman, S. Emrani, E. Nahmani, G. Barrett, H. Y. Figueroa, E. Goldberg, S. A. Hussaini and K. E. Duff (2017). "Tau Pathology Induces Excitatory Neuron Loss, Grid Cell Dysfunction, and Spatial Memory Deficits Reminiscent of Early Alzheimer's Disease." Neuron **93**(3): 533-541.e535.

Gauthier, S., H. Zhang, K. P. Ng, T. A. Pascoal and P. Rosa-Neto (2018). "Impact of the biological definition of Alzheimer's disease using amyloid, tau and neurodegeneration (ATN): what about the role of vascular changes, inflammation, Lewy body pathology?" Transl Neurodegener **7**: 12.

Glasser, M. F., T. S. Coalson, E. C. Robinson, C. D. Hacker, J. Harwell, E. Yacoub, K. Ugurbil, J. Andersson, C. F. Beckmann, M. Jenkinson, S. M. Smith and D. C. Van Essen (2016). "A multi-modal parcellation of human cerebral cortex." Nature **536**(7615): 171-178.

Glasser, M. F., S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, D. C. Van Essen and M. Jenkinson (2013). "The minimal preprocessing pipelines for the Human Connectome Project." NeuroImage **80**: 105-124.

Grassi, M., G. Perna, D. Caldirola, K. Schruers, R. Duara and D. A. Loewenstein (2018). "A clinically-translatable machine learning algorithm for the prediction of Alzheimer's disease conversion in individuals with mild and premild cognitive impairment." Journal of Alzheimer's Disease **61**(4): 1555-1573.

Gray, K. R., P. Aljabar, R. A. Heckemann, A. Hammers and D. Rueckert (2013). "Random forest-based similarity measures for multi-modal classification of Alzheimer's disease." NeuroImage **65**: 167-175.

Grill-Spector, K. and R. Malach (2004). "The human visual cortex." Annu Rev Neurosci **27**: 649-677.

Grothe, M. J., H. Barthel, J. Sepulcre, M. Dyrba, O. Sabri and S. J. Teipel (2017). "In vivo staging of regional amyloid deposition." Neurology **89**(20): 2031-2038.

Grothe, M. J., S. J. Teipel and f. t. A. s. D. N. Initiative (2016). "Spatial patterns of atrophy, hypometabolism, and amyloid deposition in Alzheimer's disease correspond to dissociable functional brain networks." Human Brain Mapping **37**(1): 35-53.

Halawa, O. A., J. R. Gatchel, R. E. Amariglio, D. M. Rentz, R. A. Sperling, K. A. Johnson and G. A. Marshall (2019). "Inferior and medial temporal tau and cortical amyloid are associated with daily functional impairment in Alzheimer's disease." Alzheimers Res Ther **11**(1): 14.

Haller, S., K.-O. Lövblad and P. Giannakopoulos (2011). Principles of Classification Analyses in Mild Cognitive Impairment (MCI) and Alzheimer Disease.

Hansson, O., M. J. Grothe, T. O. Strandberg, T. Ohlsson, D. Hägerström, J. Jögi, R. Smith and M. Schöll (2017). "Tau Pathology Distribution in Alzheimer's disease Corresponds Differentially to Cognition-Relevant Functional Brain Networks." Frontiers in Neuroscience **11**(167).

Heneka, M. T., M. J. Carson, J. E. Khoury, G. E. Landreth, F. Brosseron, D. L. Feinstein, A. H. Jacobs, T. Wyss-Coray, J. Vitorica, R. M. Ransohoff, K. Herrup, S. A. Frautschy, B. Finsen, G.

C. Brown, A. Verkhratsky, K. Yamanaka, J. Koistinaho, E. Latz, A. Halle, G. C. Petzold, T. Town, D. Morgan, M. L. Shinohara, V. H. Perry, C. Holmes, N. G. Bazan, D. J. Brooks, S. Hunot, B. Joseph, N. Deigendesch, O. Garaschuk, E. Boddeke, C. A. Dinarello, J. C. Breitner, G. M. Cole, D. T. Golenbock and M. P. Kummer (2015). "Neuroinflammation in Alzheimer's disease." The Lancet Neurology **14**(4): 388-405.

Howett, D., A. Castegnaro, K. Krzywicka, J. Hagman, D. Marchment, R. Henson, M. Rio, J. A. King, N. Burgess and D. Chan (2019). "Differentiation of mild cognitive impairment using an entorhinal cortex-based test of virtual reality navigation." Brain **142**(6): 1751-1766.

Huang, J., P. Beach, A. Bozoki and D. C. Zhu (2020). "Alzheimer's Disease Progressively Alters the Face-Evoked Visual-Processing Network." J Alzheimers Dis.

Ithapu, V. K., V. Singh, O. C. Okonkwo, R. J. Chappell, N. M. Dowling and S. C. Johnson (2015). "Imaging-based enrichment criteria using deep learning algorithms for efficient clinical trials in mild cognitive impairment." Alzheimer's & Dementia **11**(12): 1489-1499.

Jack, C. R., Jr., D. A. Bennett, K. Blennow, M. C. Carrillo, B. Dunn, S. B. Haeberlein, D. M. Holtzman, W. Jagust, F. Jessen, J. Karlawish, E. Liu, J. L. Molinuevo, T. Montine, C. Phelps, K. P. Rankin, C. C. Rowe, P. Scheltens, E. Siemers, H. M. Snyder, R. Sperling, C. Elliott, E. Masliah, L. Ryan and N. Silverberg (2018). "NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease." Alzheimer's & Dementia: The Journal of the Alzheimer's Association **14**(4): 535-562.

Jack, C. R., Jr., R. C. Petersen, Y. Xu, P. C. O'Brien, G. E. Smith, R. J. Ivnik, B. F. Boeve, E. G. Tangalos and E. Kokmen (2000). "Rates of hippocampal atrophy correlate with change in clinical status in aging and AD." Neurology **55**(4): 484-489.

Jack, C. R., Jr., R. C. Petersen, Y. C. Xu, S. C. Waring, P. C. O'Brien, E. G. Tangalos, G. E. Smith, R. J. Ivnik and E. Kokmen (1997). "Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease." Neurology **49**(3): 786-794.

Jackson, J., E. Jambrina, J. Li, H. Marston, F. Menzies, K. Phillips and G. Gilmour (2019). Targeting the synapse in Alzheimer's disease.

Jadhav, S., J. Avila, M. Scholl, G. G. Kovacs, E. Kovari, R. Skrabana, L. D. Evans, E. Kontsekova, B. Malawska, R. de Silva, L. Buee and N. Zilka (2019). "A walk through tau therapeutic strategies." Acta Neuropathol Commun **7**(1): 22.

Jenkinson, M., C. F. Beckmann, T. E. Behrens, M. W. Woolrich and S. M. Smith (2012). "FSL." Neuroimage **62**(2): 782-790.

Jiang, H., W.-K. Ching, W.-S. Cheung, W. Hou and H. Yin (2017). "Hadamard Kernel SVM with applications for breast cancer outcome predictions." BMC systems biology **11**(7): 138.

Jie, B., D. Zhang, B. Cheng and D. Shen (2013). Manifold Regularized Multi-Task Feature Selection for Multi-Modality Classification in Alzheimer's Disease. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013, Berlin, Heidelberg, Springer Berlin Heidelberg.

Jirsa, V. K., K. J. Jantzen, A. Fuchs and J. A. S. Kelso (2002). "Spatiotemporal forward solution of the EEG and MEG using network modeling." IEEE Transactions on Medical Imaging **21**(5): 493-504.

Khan, U. A., L. Liu, F. A. Provenzano, D. E. Berman, C. P. Profaci, R. Sloan, R. Mayeux, K. E. Duff and S. A. Small (2014). "Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical Alzheimer's disease." Nat Neurosci **17**(2): 304-311.

Killiany, R. J., T. Gomez-Isla, M. Moss, R. Kikinis, T. Sandor, F. Jolesz, R. Tanzi, K. Jones, B. T. Hyman and M. S. Albert (2000). "Use of structural magnetic resonance imaging to predict who will get Alzheimer's disease." Ann Neurol **47**(4): 430-439.

Kloppel, S., C. Stonnington, C. Chu, B. Draganski, R. Scahill, J. Rohrer, N. C. Fox, C. Jack, J. Ashburner and R. S. J. Frackowiak (2008). Automatic classification of MR scans in Alzheimers disease.

Knopman, D. S., E. S. Lundt, T. M. Therneau, P. Vemuri, V. J. Lowe, K. Kantarci, J. L. Gunter, M. L. Senjem, M. M. Mielke, M. M. Machulda, B. F. Boeve, D. T. Jones, J. Graff-Radford, S. M. Albertson, C. G. Schwarz, R. C. Petersen and C. R. Jack (2019). "Entorhinal cortex tau, amyloid-β, cortical thickness and memory performance in non-demented subjects." Brain **142**(4): 1148-1160.

Kotsiantis, S. B., I. D. Zaharakis and P. E. Pintelas (2006). "Machine learning: a review of classification and combining techniques." Artificial Intelligence Review **26**(3): 159-190.

Kravitz, D. J., K. S. Saleem, C. I. Baker, L. G. Ungerleider and M. Mishkin (2013). "The ventral visual pathway: an expanded neural framework for the processing of object quality." Trends in cognitive sciences **17**(1): 26-49.

Kunz, L., S. Maidenbaum, D. Chen, L. Wang, J. Jacobs and N. Axmacher (2019). "Mesoscopic Neural Representations in Spatial Navigation." Trends in Cognitive Sciences **23**(7): 615-630.

Lee, E. S., K. Yoo, Y. B. Lee, J. Chung, J. E. Lim, B. Yoon and Y. Jeong (2016). "Default Mode Network Functional Connectivity in Early and Late Mild Cognitive Impairment: Results From the Alzheimer's Disease Neuroimaging Initiative." Alzheimer Dis Assoc Disord **30**(4): 289-296.

Litvak, V., J. Mattout, S. Kiebel, C. Phillips, R. Henson, J. Kilner, G. Barnes, R. Oostenveld, J. Daunizeau, G. Flandin, W. Penny and K. Friston (2011). "EEG and MEG data analysis in SPM8." Comput Intell Neurosci **2011**: 852961.

Liu, M., D. Zhang and D. Shen (2012). "Ensemble sparse classification of Alzheimer's disease." NeuroImage **60**(2): 1106-1116.

Madsen, S. K., A. J. Ho, X. Hua, P. S. Saharan, A. W. Toga, C. R. Jack, M. W. Weiner and P. M. Thompson (2010). "3D maps localize caudate nucleus atrophy in 400 Alzheimer's disease, mild cognitive impairment, and healthy elderly subjects." Neurobiology of Aging **31**(8): 1312-1325.

Martí-Juan, G., G. Sanroma-Guell and G. Piella (2020). "A survey on machine and statistical learning for longitudinal analysis of neuroimaging data in Alzheimer's disease." Computer Methods and Programs in Biomedicine **189**: 105348.

McKhann, G. M., D. S. Knopman, H. Chertkow, B. T. Hyman, C. R. Jack, Jr., C. H. Kawas, W. E. Klunk, W. J. Koroshetz, J. J. Manly, R. Mayeux, R. C. Mohs, J. C. Morris, M. N. Rossor, P. Scheltens, M. C. Carrillo, B. Thies, S. Weintraub and C. H. Phelps (2011). "The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease." Alzheimer's & dementia : the journal of the Alzheimer's Association **7**(3): 263-269.

Moradi, E., A. Pepe, C. Gaser, H. Huttunen and J. Tohka (2015). "Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects." Neuroimage **104**: 398-412.

Nunez, P. L. and R. Srinivasan (2005). Electric Fields of the Brain. Oxford, UK, Oxford university press.

Nunez, P. L. and R. Srinivasan (2006). "A theoretical basis for standing and traveling brain waves measured with human EEG with implications for an integrated consciousness." Clinical Neurophysiology **117**(11): 2424-2435.

Nunomura, A., G. Perry, G. Aliev, K. Hirai, A. Takeda, E. K. Balraj, P. K. Jones, H. Ghanbari, T. Wataya and S. Shimohama (2001). "Oxidative damage is the earliest event in Alzheimer disease." Journal of Neuropathology & Experimental Neurology **60**(8): 759-767.

Oltu, B., M. F. Akşahin and S. Kibaroğlu (2021). "A novel electroencephalography based approach for Alzheimer's disease and mild cognitive impairment detection." Biomedical Signal Processing and Control **63**: 102223.

Palmqvist, S., M. Schöll, O. Strandberg, N. Mattsson, E. Stomrud, H. Zetterberg, K. Blennow, S. Landau, W. Jagust and O. Hansson (2017). "Earliest accumulation of β-amyloid occurs within the default-mode network and concurrently affects brain connectivity." Nature Communications **8**(1): 1214.

Pascoal, T. A., S. Mathotaarachchi, M. Shin, A. L. Benedet, S. Mohades, S. Wang, T. Beaudry, M. S. Kang, J.-P. Soucy, A. Labbe, S. Gauthier and P. Rosa-Neto (2017). "Synergistic interaction between amyloid and tau predicts the progression to dementia." Alzheimer's & Dementia **13**(6): 644-653.

Petersen, R. C., B. Caracciolo, C. Brayne, S. Gauthier, V. Jelic and L. Fratiglioni (2014). "Mild cognitive impairment: a concept in evolution." Journal of internal medicine **275**(3): 214-228.

Pimenova, A. A., T. Raj and A. M. Goate (2018). "Untangling Genetic Risk for Alzheimer's Disease." Biol. Psychiatry **83**(4): 300-310.

Rathore, S., M. Habes, M. A. Iftikhar, A. Shacklett and C. Davatzikos (2017). "A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages." NeuroImage **155**: 530-548.

Rebala, G., A. Ravi and S. Churiwala (2019). Classification. An Introduction to Machine Learning. G. Rebala, A. Ravi and S. Churiwala. Cham, Springer International Publishing: 57-66.

Reuter, M., N. J. Schmansky, H. D. Rosas and B. Fischl (2012). "Within-subject template estimation for unbiased longitudinal image analysis." Neuroimage **61**(4): 1402-1418.

Rieck, J. R., K. M. Rodrigue, K. M. Kennedy, M. D. Devous, Sr. and D. C. Park (2015). "The effect of beta-amyloid on face processing in young and old adults: A multivariate analysis of the BOLD signal." Hum Brain Mapp **36**(7): 2514-2526.

Ritter, P., M. Schirner, A. R. McIntosh and V. K. Jirsa (2013). "The virtual brain integrates computational modeling and multimodal neuroimaging." Brain Connectivity **3**(2): 121-145.

Ritter, P., M. Schirner, A. R. McIntosh and V. K. Jirsa (2013). "The virtual brain integrates computational modeling and multimodal neuroimaging." Brain Connect **3**(2): 121-145.

Rondina, J. M., L. K. Ferreira, F. L. de Souza Duran, R. Kubo, C. R. Ono, C. C. Leite, J. Smid, R. Nitrini, C. A. Buchpiguel and G. F. Busatto (2018). "Selecting the most relevant brain regions to discriminate Alzheimer's disease patients from healthy controls using multiple kernel learning: A comparison across functional and structural imaging modalities and atlases." Neuroimage Clin **17**: 628-641.

Rosen, B. Q. and E. Halgren (2020). "A whole-cortex probabilistic diffusion tractography connectome." bioRxiv: 2020.2006.2022.166041.

Rossini, P. M., R. Di Iorio, F. Vecchio, M. Anfossi, C. Babiloni, M. Bozzali, A. C. Bruni, S. F. Cappa, J. Escudero, F. J. Fraga, P. Giannakopoulos, B. Guntekin, G. Logroscino, C. Marra, F. Miraglia, F. Panza, F. Tecchio, A. Pascual-Leone and B. Dubois (2020). "Early diagnosis of Alzheimer's disease: the role of biomarkers including advanced EEG signal analysis. Report from the IFCN-sponsored panel of experts." Clinical Neurophysiology **131**(6): 1287-1310.

Rowland, D. C., Y. Roudi, M.-B. Moser and E. I. Moser (2016). "Ten Years of Grid Cells." Annual Review of Neuroscience **39**(1): 19-40.

Rüb, U., K. Stratmann, H. Heinsen, D. Del Turco, E. Ghebremedhin, K. Seidel, W. den Dunnen and H.-W. Korf (2016). "Hierarchical Distribution of the Tau Cytoskeletal Pathology in the Thalamus of Alzheimer's Disease Patients." Journal of Alzheimer's Disease **49**: 905-915.

Sadigh-Eteghad, S., B. Sabermarouf, A. Majdi, M. Talebi, M. Farhoudi and J. Mahmoudi (2015). "Amyloid-Beta: A Crucial Factor in Alzheimer's Disease." Medical Principles and Practice **24**(1): 1-10.

Sajda, P. (2006). "Machine Learning for Detection and Diagnosis of Disease." Annual Review of Biomedical Engineering **8**(1): 537-565.

Samper-González, J., N. Burgos, S. Bottani, S. Fontanella, P. Lu, A. Marcoux, A. Routier, J. Guillon, M. Bacci, J. Wen, A. Bertrand, H. Bertin, M.-O. Habert, S. Durrleman, T. Evgeniou and O. Colliot (2018). "Reproducible evaluation of classification methods in Alzheimer's disease: framework and application to MRI and PET data." bioRxiv.

Samuel, A. L. (1959). "Some studies in machine learning using the game of checkers." IBM Journal of research and development **3**(3): 210-229.

Sanz Leon, P., S. A. Knock, M. M. Woodman, L. Domide, J. Mersmann, A. R. McIntosh and V. Jirsa (2013). "The Virtual Brain: a simulator of primate brain network dynamics." Frontiers in Neuroinformatics **7**(10).

Sarica, A., A. Cerasa and A. Quattrone (2017). "Random Forest algorithm for the classification of neuroimaging data in Alzheimer's disease: A systematic review." Frontiers in aging neuroscience **9**: 329.

Schnitzler, A. and J. Gross (2005). "Normal and pathological oscillatory communication in the brain." Nature Reviews Neuroscience **6**(4): 285-296.

Schöberl, F., C. Pradhan, S. Irving, K. Buerger, G. Xiong, G. Kugler, S. Kohlbecher, J. Engmann, P. Werner, M. Brendel, E. Schneider, R. Perneczky, K. Jahn, C. la Fougère, P. Bartenstein, T. Brandt, M. Dieterich and A. Zwergal (2020). "Real-space navigation testing differentiates between amyloid-positive and -negative aMCI." Neurology **94**(8): e861-e873.

Schöll, M., Samuel N. Lockhart, Daniel R. Schonhaut, James P. O'Neil, M. Janabi, R. Ossenkoppele, Suzanne L. Baker, Jacob W. Vogel, J. Faria, Henry D. Schwimmer, Gil D. Rabinovici and William J. Jagust (2016). "PET Imaging of Tau Deposition in the Aging Human Brain." Neuron **89**(5): 971-982.

Selkoe, D. J. (2019). Early network dysfunction in Alzheimer's disease.

Selkoe, D. J. and J. Hardy (2016). "The amyloid hypothesis of Alzheimer's disease at 25 years." EMBO Mol Med **8**(6): 595-608.

Shaikh, T. A. and R. Ali (2019). "Automated atrophy assessment for Alzheimer's disease diagnosis from brain MRI images." Magnetic resonance imaging **62**: 167-173.

Shannon, C. E. (1948). "A mathematical theory of communication." The Bell system technical journal **27**(3): 379-423.

Smailovic, U. and V. Jelic (2019). "Neurophysiological Markers of Alzheimer's Disease: Quantitative EEG Approach." Neurology and Therapy **8**(2): 37-55.

Smith, R. E., J.-D. Tournier, F. Calamante and A. Connelly (2012). "Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information." NeuroImage **62**(3): 1924-1938.

Smith, R. E., J. D. Tournier, F. Calamante and A. Connelly (2015). "SIFT2: Enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography." Neuroimage **119**: 338-351.

Smith, S. M., M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney, R. K. Niazy, J. Saunders, J.

Vickers, Y. Zhang, N. De Stefano, J. M. Brady and P. M. Matthews (2004). "Advances in functional and structural MR image analysis and implementation as FSL." Neuroimage **23 Suppl 1**: S208-219.

Soumaya, Z., B. D. Taoufiq, N. Benayad, K. Yunus and A. Abdelkrim (2021). "The detection of Parkinson disease using the genetic algorithm and SVM classifier." Applied Acoustics **171**: 107528.

Spiegler, A., S. J. Kiebel, F. M. Atay and T. R. Knösche (2010). "Bifurcation analysis of neural mass models: Impact of extrinsic inputs and dendritic time constants." NeuroImage **52**(3): 1041-1058.

Stefanovski, L., P. Triebkorn, A. Spiegler, M.-A. Diaz-Cortes, A. Solodkin, V. Jirsa, A. R. McIntosh, P. Ritter and for the Alzheimer's Disease Neuroimaging Initiative (2019). "Linking Molecular Pathways and Large-Scale Computational Modeling to Assess Candidate Disease Mechanisms and Pharmacodynamics in Alzheimer's Disease." Frontiers in Computational Neuroscience **13**(54).

Stefanovski, L., P. Triebkorn, A. Spiegler, M. A. Diaz-Cortes, A. Solodkin, V. Jirsa, A. R. McIntosh and P. Ritter (2019). "Linking Molecular Pathways and Large-Scale Computational Modeling to Assess Candidate Disease Mechanisms and Pharmacodynamics in Alzheimer's Disease." Frontiers in Computational Neuroscience.

Steriade, M. (2003). Neuronal Substrates of Sleep and Epilepsy. Cambridge, UK, Cambridge University Press.

Stone, M. (1974). "Cross‐validatory choice and assessment of statistical predictions." Journal of the Royal Statistical Society: Series B (Methodological) **36**(2): 111-133.

Storck Steffen, E. and U. Pietrzik Claus (2018). The Blood brain-barrier and its role in Alzheimer's disease. Neuroforum. **24:** A197.

Suk, H. I., S. W. Lee and D. Shen (2014). "Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis." Neuroimage **101**: 569-582.

Sweeney, M. D., A. P. Sagare and B. V. Zlokovic (2018). "Blood-brain barrier breakdown in Alzheimer disease and other neurodegenerative disorders." Nat Rev Neurol **14**(3): 133-150.

Swerdlow, R. H. and S. M. Khan (2009). "The Alzheimer's disease mitochondrial cascade hypothesis: an update." Exp Neurol **218**(2): 308-315.

Tadel, F., S. Baillet, J. C. Mosher, D. Pantazis and R. M. Leahy (2011). "Brainstorm: A User-Friendly Application for MEG/EEG Analysis." Computational Intelligence and Neuroscience **2011**: 1-13.

Takatori, S., W. Wang, A. Iguchi and T. Tomita (2019). "Genetic Risk Factors for Alzheimer Disease: Emerging Roles of Microglia in Disease Pathomechanisms." Adv Exp Med Biol **1118**: 83-116.

Tapia-Rojas, C., F. Cabezas-Opazo, C. A. Deaton, E. H. Vergara, G. V. W. Johnson and R. A. Quintanilla (2019). "It's all about tau." Prog Neurobiol **175**: 54-76.

Teipel, S. J., J. Kurth, B. Krause and M. J. Grothe (2015). "The relative importance of imaging markers for the prediction of Alzheimer's disease dementia in mild cognitive impairment — Beyond classical regression." NeuroImage: Clinical **8**: 583-593.

Thal, D. R., U. Rüb, M. Orantes and H. Braak (2002). "Phases of A beta-deposition in the human brain and its relevance for the development of AD." Neurology **58**(12): 1791-1800.

Thomas, B. A., V. Cuplov, A. Bousse, A. Mendes, K. Thielemans, B. F. Hutton and K. Erlandsson (2016). "PETPVC: a toolbox for performing partial volume correction techniques in positron emission tomography." Phys Med Biol **61**(22): 7975-7993.

Tong, T., R. Wolz, Q. Gao, R. Guerrero, J. V. Hajnal and D. Rueckert (2014). "Multiple instance learning for classification of dementia in brain MRI." Medical Image Analysis **18**(5): 808-818.

Tonnies, E. and E. Trushina (2017). "Oxidative Stress, Synaptic Dysfunction, and Alzheimer's Disease." J Alzheimers Dis.

Tournier, J.-D., F. Calamante and A. Connelly (2010). Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions.

Tournier, J. D., F. Calamante and A. Connelly (2007). "Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution." Neuroimage **35**(4): 1459-1472.

Tournier, J. D., F. Calamante and A. Connelly (2013). "Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging." NMR Biomed **26**(12): 1775-1786.

Van Cauwenberghe, C., C. Van Broeckhoven and K. Sleegers (2016). "The genetic landscape of Alzheimer disease: clinical implications and perspectives." Genet Med **18**(5): 421-430.

van Rossum, I. A., S. Vos, R. Handels and P. J. Visser (2010). "Biomarkers as predictors for conversion from mild cognitive impairment to Alzheimer-type dementia: implications for trial design." Journal of Alzheimer's Disease **20**(3): 881-891.

Veitch, D. P., M. W. Weiner, P. S. Aisen, L. A. Beckett, N. J. Cairns, R. C. Green, D. Harvey, C. R. Jack, Jr., W. Jagust, J. C. Morris, R. C. Petersen, A. J. Saykin, L. M. Shaw, A. W. Toga and J. Q. Trojanowski (2019). "Understanding disease progression and improving Alzheimer's disease clinical trials: Recent highlights from the Alzheimer's Disease Neuroimaging Initiative." Alzheimers Dement **15**(1): 106-152.

Veraart, J., D. S. Novikov, D. Christiaens, B. Ades-aron, J. Sijbers and E. Fieremans (2016). "Denoising of diffusion MRI using random matrix theory." NeuroImage **142**: 394-406.

Vossel, K. A., K. Zhang, J. Brodbeck, A. C. Daub, P. Sharma, S. Finkbeiner, B. Cui and L. Mucke (2010). "Tau Reduction Prevents Aβ-Induced Defects in Axonal Transport." Science **330**(6001): 198-198.

Wennberg, A. M., J. L. Whitwell, N. Tosakulwong, S. D. Weigand, M. E. Murray, M. M. Machulda, L. Petrucelli, M. M. Mielke, C. R. Jack, D. S. Knopman, J. E. Parisi, R. C. Petersen,

D. W. Dickson and K. A. Josephs (2019). "The influence of tau, amyloid, alpha-synuclein, TDP-43, and vascular pathology in clinically normal elderly individuals." Neurobiology of Aging **77**: 26-36.

Williams, O. A., Y. An, N. M. Armstrong, M. Kitner-Triolo, L. Ferrucci and S. M. Resnick (2020). "Profiles of Cognitive Change in Preclinical and Prodromal Alzheimer's Disease Using Change-Point Analysis." Journal of Alzheimer's Disease **75**: 1169-1180.

Wolpert, D. H. and W. G. Macready (1997). "No free lunch theorems for optimization." IEEE transactions on evolutionary computation **1**(1): 67-82.

Woolrich, M. W., S. Jbabdi, B. Patenaude, M. Chappell, S. Makni, T. Behrens, C. Beckmann, M. Jenkinson and S. M. Smith (2009). "Bayesian analysis of neuroimaging data in FSL." Neuroimage **45**(1 Suppl): S173-186.

Young, J., M. Modat, M. J. Cardoso, A. Mendelson, D. Cash and S. Ourselin (2013). Accurate multimodal probabilistic prediction of conversion to Alzheimer's disease in patients with mild cognitive impairment.

Yun, H. J., K. Kwak, J.-M. Lee and I. Alzheimer's Disease Neuroimaging (2015). "Multimodal Discrimination of Alzheimer's Disease Based on Regional Cortical Atrophy and Hypometabolism." PLOS ONE **10**(6): e0129250.

Zetterberg, H. and J. M. Schott (2019). "Biomarkers for Alzheimer's disease beyond amyloid and tau." Nat Med **25**(2): 201-203.

Zhang, D., Y. Wang, L. Zhou, H. Yuan and D. Shen (2011). "Multimodal classification of Alzheimer's disease and mild cognitive impairment." NeuroImage **55**(3): 856-867.

Zhao, H., X. Li, W. Wu, Z. Li, L. Qian, S. Li, B. Zhang and Y. Xu (2015). "Atrophic Patterns of the Frontal-Subcortical Circuits in Patients with Mild Cognitive Impairment and Alzheimer's Disease." PLOS ONE **10**(6): e0130017.

Zhao, J., X. Ding, Y. Du, X. Wang and G. Men (2019). "Functional connectivity between white matter and gray matter based on fMRI for Alzheimer's disease classification." Brain and behavior **9**(10): e01407.