# Signals of Decline: Machine Learning driven Biomarkers for Alzheimer's Disease

# **Anonymous Author(s)**

Affiliation Address email

#### Abstract

Alzheimer's disease (AD) is a growing global health challenge, with pathological changes beginning decades before clinical symptoms. Identifying non-invasive and interpretable biomarkers is critical for early intervention. Magnetoencephalography (MEG) provides access to brain oscillatory dynamics and connectivity patterns that are disrupted in mild cognitive impairment (MCI), a prodromal stage of AD. We evaluate five families of MEG-derived features and train an ensemble of 200 feature models, achieving MCI classification with F1 72.43%. We use Shapley Additive Explanations (SHAP) to highlight discriminative regions and connections, offering interpretable insights and pointing to potential new markers. Beyond binary detection, model scores correlate with Mini-Mental State Examination (MMSE) scores, suggesting potential for continuous disease staging. Together, these results establish MEG-based machine learning as a promising avenue for robust and clinically meaningful biomarkers.

## 4 1 Introduction

3

5

6

8

10

11

12

13

28

29

30

31

The blessing of longer life prospects, driven by biomedical advances and better access to healthcare, 15 has brought with it a rise in age-related neurological disorders such as Alzheimer's disease (AD). By 17 2050, more than **130 million** people [1] are expected to live with AD, imposing a global economic burden of over \$9 trillion [2, 3, 4, 5]. AD is a spectrum disorder in which pathological changes emerge 18 long before measurable cognitive decline [6, 7, 8], making early biomarkers essential for intervention. 19 Current clinical biomarkers rely on cerebrospinal fluid (CSF) measurements or PET imaging [7, 9]. 20 While accurate, these methods are costly, invasive, and difficult to scale. Recent progress in plasma 21 biomarkers [10, 11] offers new possibilities, but electrophysiological measures such as EEG and 22 MEG provide an attractive alternative: they are non-invasive, relatively inexpensive, and directly 23 capture neural dynamics known to be disrupted in AD [12, 13, 14]. Particularly, alterations in spectral power and functional connectivity have been linked to mild cognitive impairment (MCI), a prodromal 25 26 stage of AD. Detecting MCI is therefore critical, as it reveals early disease processes and creates an opportunity for timely intervention, even if not all individuals progress to AD. 27

In this work, we present a systematic framework to evaluate MEG-derived features for MCI detection and interpretation. We show that an ensemble of 200 models trained across five feature families achieves an F1 score of 72.43%. We demonstrate that SHAP analysis [15] can provide detailed insights into early AD processes, allowing for interpretable and clinically meaningful biomarkers. Beyond binary classification, we find that model scores correlate with MMSE, pointing to the potential for regression-based disease staging.

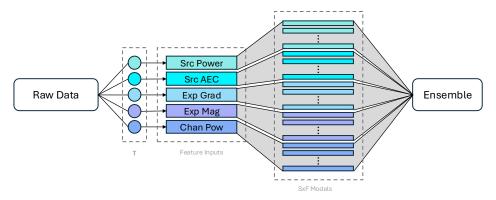


Figure 1: 5 different feature inputs (F) are transformed (T) from the raw data (see section 2.1). We train each feature model with 40 seeds (S). Finally we ensemble the 200 (SxF) models.

#### 4 2 Method

36

37

39

40

41

42

43

44

45

46

47

48

49

50

51

63

### 2.1 Dataset and Feature Extraction

We use the BioFIND dataset [16], a multi-site MEG study comprising 324 resting-state recordings (158 MCI, 166 controls). Data was collected at the University of Cambridge (UK) and the Centre for Biomedical Technology (Spain). Each subject underwent 3-5 minutes of eye-closed resting-state MEG and individual MRI. We make use only of the MEG data. Preprocessing followed Vaghari et al. [16] with recordings epoched into 2-second windows, artifacts removed, and spectra estimated using the Welch method. It is important to note that the MEG system had two types of sensor, namely magnetometers (MAG) and gradiometers (GRD) which capture complementary aspects of the magnetic field. Using this data, we derived five feature sets:

- Source Power (Src\_Power): Spectral power for  $\alpha$  and  $\beta$  frequency bands in 38 cortical parcels (Harvard-Oxford atlas [17]).
- Source Amplitude Envelope Correlations (Src\_AEC): Connectivity in  $\alpha$  (8–12 Hz) and  $\beta$  (12–30 Hz) bands, orthogonalized to mitigate field spread [18].
- Exponentiated Spectral Slopes (Exp\_Grad, Exp\_Mag): Aperiodic slope of the spectrum, extracted from GRD and MAG [19], corresponding to two feature sets.
- Channel Power (Chan\_Pow): Band-limited power for 102 MAG and 204 GRD sensors across eight frequency bands.

# 52 2.2 System Architecture

The base architecture used for all models is a 3 linear layer block with dropout followed by a self attention block. This results in a total of 415K trainable parameters. The input feature datasets were all divided into 80%, 10%, 10% splits for training, validation and testing respectively. We use Adam [20] as optimizer and we perform hyperparameter finetuning on learning rate and batch size per model, applying early stopping with a patience of 20. For each feature set, we trained 40 classifiers with different seeds to ensure robustness across subjects. Ensembles were then constructed both per feature and across features, yielding 200 models in the final ensemble following the architecture in Figure 1. We report our F1, accuracy, AUC, recall, and precision results for MCI classification.

#### 3 Results

## 3.1 Classifying Mild Cognitive Impairment

Our method trains different models on different input features corresponding to distinct pieces of information relating to the patients neural dynamics. By transforming the raw data into a single feature we can reduce noise but also relevant information. It then makes sense that individual feature

models on average perform significantly lower than performances reported in previous work (the best single feature model Src AEC performing a full 7% lower than Vaghari et al. s 68% Acc [21]). However, soft voting ensembles across seeds for the same feature already begin to show improvements with feature models achieving a boost of up to 4%. Table 1 shows results for feature ensembles as well as the final ensemble of all 200 models resulting in an F1 Score of 72.43%. We achieve state of the art results using a simple and modular architecture that enables for explainable methods such as SHAP.

	N Models	F1	AUC	Acc.	Recall	Precision
Source Power	40	67.95	69.65	63.55	79.49	59.33
Source AEC	40	64.12	66.56	61.99	69.87	59.24
Exp Gradiometer	40	65.08	58.35	54.21	87.82	51.7
Exp Magnetometer	40	64.62	53.13	49.84	94.23	49.16
Channel Power	40	66.86	68.33	63.86	75.0	60.31
Ensemble	200	72.43	73.88	68.22	85.9	62.62

Table 1: Results for model ensemble per Feature and across Features.

## 73 3.2 Interpretability through SHAP Values

Our individual feature models with shorter dimensionality inputs allow us to use SHAP explanations effectively. Shapley Additive Values can quantify contributions per input, which provides insights on numerous aspects. For Source AECs they indicate which connections between regions are most informative in identifying MCI, while for Exponentiated Spectral Slopes it shows which sensors report most differing values between controls and patients with MCI. To exemplify the interpretability of our method we present SHAP values for the  $\alpha$  frequency band on our Src\_AEC and Source\_Power model ensembles in Figure 2. As can be observed, our SHAP analysis revealed discriminative regions in the  $\alpha$  source power band, with occipital parcels emerging as particularly informative (Fig. 2a, 2c).

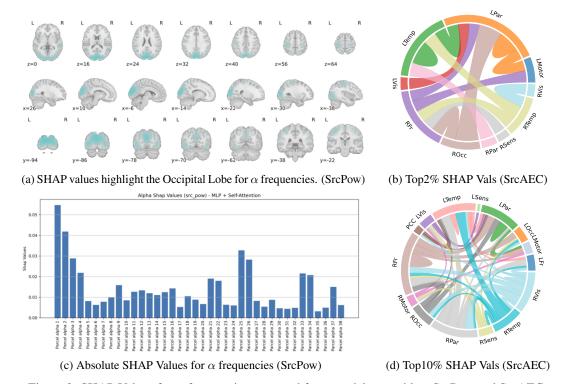


Figure 2: SHAP Values for  $\alpha$  frequencies extracted from model ensembles: SrcPow and SrcAEC

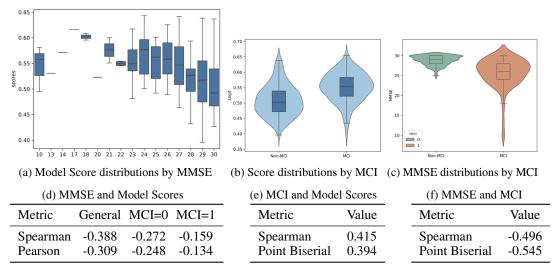


Figure 3: Correlations across our Model Ensemble Scores, MCI labels and MMSE scores.

In the connectivity domain, fronto-parietal and inter-temporal connections showed strong contributions (Fig. 2b, 2d) consistent with established AD pathology [22].

#### 84 3.3 Correlation with Levels of Cognitive Function

Beyond achieving state-of-the-art performance in MCI classification with interpretable models, our analysis shows that the model's predictive scores capture clinically meaningful information about cognitive severity. As shown in 3d, 3e and 3f, model outputs are moderately correlated with MMSE scores and align with the expected association between MMSE and MCI status. Notably, while correlations within diagnostic subgroups are weaker, the overall trend demonstrates that higher model scores correspond to lower MMSE performance, suggesting that the model is sensitive not only to categorical impairment but also to gradations of decline. This positions our approach as more than a binary classifier: it provides a signal that could be leveraged to quantify disease severity or track progression. Although the current model is not optimized for regression, these findings highlight its translational potential and motivate future work on continuous modeling of cognitive decline, with implications for staging, monitoring, and personalized dementia interventions.

# 4 Discussion and Conclusion

This work shows that MEG-based machine learning can yield robust and interpretable biomarkers for Alzheimer's disease. By evaluating five MEG-derived feature families, we achieved state-of-the-art MCI classification (F1 = 72.43%), while also demonstrating that model scores correlate with MMSE, capturing gradations of cognitive decline. SHAP analyses revealed candidate regions and connections that may reflect early disease processes, though further work is required to validate their clinical significance.

These findings highlight MEG as a scalable, non-invasive alternative to current biomarkers, with potential for both early detection and disease staging. Limitations include reliance on a single dataset and the need for longitudinal validation, particularly to support continuous modeling of decline. Nonetheless, our framework establishes MEG as a promising foundation for clinically meaningful, non-invasive, interpretable biomarkers for early detection and staging of cognitive decline.

# 108 5 Acknowledgements

We are grateful for access to the BioFIND database on UK dementia portal (DPUK). This work was undertaken using resources in the Dementias Platform UK (DPUK) Data Portal: DPUK Portal.

## References

111

126

127

128

129

130

- 112 [1] Martin James Prince et al. "World Alzheimer Report 2015 The Global Impact of Demen113 tia: An analysis of prevalence, incidence, cost and trends". English. In: (Aug. 2015). Pub114 lisher: Alzheimer's Disease International. URL: https://kclpure.kcl.ac.uk/portal/
  115 en/publications/world-alzheimer-report-2015--the-global-impact-of116 dementia(ae525fda-1938-4892-8daa-a2222a672254).html (visited on 02/21/2022).
- 117 [2] Amir Abbas Tahami Monfared et al. "The Humanistic and Economic Burden of Alzheimer's
  118 Disease". en. In: *Neurology and Therapy* 11.2 (June 2022), pp. 525–551. ISSN: 2193-6536.
  119 DOI: 10.1007/s40120-022-00335-x. URL: https://doi.org/10.1007/s40120-022120 00335-x (visited on 08/05/2024).
- [3] Jianping Jia et al. "The cost of Alzheimer's disease in China and re-estimation of costs worldwide". en. In: *Alzheimer's & Dementia* 14.4 (2018). \_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1016/j.jalz.2017.12.006, pp. 483-491. ISSN: 1552-5279. DOI: 10.1016/j.jalz.2017.12.006. URL: https://onlinelibrary.wiley.com/doi/abs/10.1016/j.jalz.2017.12.006 (visited on 08/05/2024).
  - [4] Richard Cimler et al. "Predictions of Alzheimer's disease treatment and care costs in European countries". en. In: *PLOS ONE* 14.1 (Jan. 2019). Publisher: Public Library of Science, e0210958. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0210958. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0210958 (visited on 09/13/2024).
- [5] Emilie Clay et al. "Economic Burden for Alzheimer's Disease in China from 2010 to 2050:
   A Modelling Study". en. In: Journal of Market Access & Health Policy 7.1 (Jan. 2019).
   Number: 1 Publisher: Multidisciplinary Digital Publishing Institute, p. 1667195. ISSN: 2001-6689. DOI: 10.1080/20016689.2019.1667195. URL: https://www.mdpi.com/2001-6689/7/1/1667195 (visited on 09/13/2024).
- Reisa A. Sperling et al. "Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease". In: *Alzheimer's & Dementia* 7.3 (May 2011), pp. 280–292. ISSN: 1552-5260. DOI: 10.1016/j.jalz.2011.03.003. URL: https://www.sciencedirect.com/science/article/pii/S1552526011000999 (visited on 08/05/2024).
- Clifford R. Jack Jr. et al. "Revised criteria for diagnosis and staging of Alzheimer's disease:
  Alzheimer's Association Workgroup". en. In: *Alzheimer's & Dementia* 20.8 (2024). \_eprint:
  https://onlinelibrary.wiley.com/doi/pdf/10.1002/alz.13859, pp. 5143–5169. ISSN: 1552-5279.
  DOI: 10.1002/alz.13859. URL: https://onlinelibrary.wiley.com/doi/abs/10.
  1002/alz.13859 (visited on 09/26/2024).
- Bruno Dubois et al. "Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria". In: *Alzheimer's & Dementia* 12.3 (Mar. 2016), pp. 292–323. ISSN: 1552-5260. DOI: 10.1016/j.jalz.2016.02.002. URL: https://www.sciencedirect.com/science/article/pii/S1552526016000509 (visited on 09/26/2024).
- [9] Clifford R. Jack et al. "Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers". English. In: *The Lancet Neurology* 12.2 (Feb. 2013). Publisher: Elsevier, pp. 207–216. ISSN: 1474-4422, 1474-4465. DOI: 10.1016/S1474-4422(12)70291-0. URL: https://www.thelancet.com/journals/laneur/article/PIIS1474-4422(12)70291-0/abstract (visited on 12/16/2024).
- James D. Doecke et al. "Total A42/A40 ratio in plasma predicts amyloid-PET status, independent of clinical AD diagnosis". In: *Neurology* 94.15 (Apr. 2020). Publisher: Wolters Kluwer, e1580–e1591. DOI: 10.1212/WNL.000000000009240. URL: https://www-neurologyorg.ezp-prod1.hul.harvard.edu/doi/10.1212/WNL.0000000000009240 (visited on 04/10/2025).
- Thomas K. Karikari et al. "Blood phospho-tau in Alzheimer disease: analysis, interpretation, and clinical utility". en. In: *Nature Reviews Neurology* 18.7 (July 2022). Publisher: Nature Publishing Group, pp. 400–418. ISSN: 1759-4766. DOI: 10.1038/s41582-022-00665-2. URL: https://www.nature.com/articles/s41582-022-00665-2 (visited on 12/16/2024).

- Martí Colom-Cadena et al. "The clinical promise of biomarkers of synapse damage or loss in
   Alzheimer's disease". en. In: *Alzheimer's Research & Therapy* 12.1 (Mar. 2020), p. 21. ISSN:
   1758-9193. DOI: 10.1186/s13195-020-00588-4. URL: https://doi.org/10.1186/s13195-020-00588-4 (visited on 08/05/2024).
- Delshad Vaghari, Ehsanollah Kabir, and Richard N. Henson. "Late combination shows that MEG adds to MRI in classifying MCI versus controls". en. In: *NeuroImage* 252 (May 2022), p. 119054. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2022.119054. URL: https: //www.sciencedirect.com/science/article/pii/S1053811922001835 (visited on 06/27/2022).
- Fernando Maestú et al. "The Importance of the Validation of M/EEG With Current Biomarkers in Alzheimer's Disease." In: *Frontiers in Human Neuroscience* 13 (Feb. 2019). MAG ID: 2912389312, pp. 17–17. DOI: 10.3389/fnhum.2019.00017.
- 178 [15] Scott M. Lundberg and Su-In Lee. "A unified approach to interpreting model predictions". In:

  179 Proceedings of the 31st International Conference on Neural Information Processing Systems.

  180 NIPS'17. Red Hook, NY, USA: Curran Associates Inc., Dec. 2017, pp. 4768–4777. ISBN:

  181 978-1-5108-6096-4. (Visited on 08/19/2025).
- 182 [16] Delshad Vaghari et al. A multi-site, multi-participant magnetoencephalography resting-state dataset to study dementia: The BioFIND dataset. 2022. DOI: doi:10.1016/j.neuroimage. 2022.119344.
- Giles L. Colclough et al. "How reliable are MEG resting-state connectivity metrics?" In:
   NeuroImage 138 (Sept. 2016). MAG ID: 2418779657, pp. 284-293. DOI: 10.1016/j.
   neuroimage.2016.05.070.
- Giles L. Colclough et al. "A symmetric multivariate leakage correction for MEG connectomes."
  In: NeuroImage 117 (Aug. 2015). MAG ID: 2021285908, pp. 439–448. DOI: 10.1016/j.neuroimage.2015.03.071.
- Thomas Donoghue et al. "Parameterizing neural power spectra into periodic and aperiodic components". en. In: *Nature Neuroscience* 23.12 (Dec. 2020). Publisher: Nature Publishing Group, pp. 1655–1665. ISSN: 1546-1726. DOI: 10.1038/s41593-020-00744-x. URL: https://www.nature.com/articles/s41593-020-00744-x (visited on 08/18/2025).
- [20] Diederik P Kingma and Jimmy Ba. "Adam: A method for stochastic optimization". In: arXiv preprint arXiv:1412.6980 (2014).
- Delshad Vaghari et al. A multi-site, multi-participant magnetoencephalography resting-state dataset to study dementia: The BioFIND dataset. en. Tech. rep. Type: article. medRxiv, May 2021, p. 2021.05.19.21257330. DOI: 10.1101/2021.05.19.21257330. (Visited on 03/19/2025).
- Pablo Martínez-Cañada et al. "Combining aperiodic 1/f slopes and brain simulation: An EEG/MEG proxy marker of excitation/inhibition imbalance in Alzheimer's disease". In: Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 15.3 (2023), e12477.