Received: 20 March 2021 | Revised: 12 May 2021 | Accepted: 4 June 2021 | Published online: 13 August 2021

DOI: 10.1002/dad2.12227

Alzheimer's & Dementia
Diagnosis, Assessment & Disease Monitoring

RESEARCH ARTICLE

Routine magnetoencephalography in memory clinic patients: A machine learning approach

Alida A. Gouw[1,2] | **Arjan Hillebrand**[2] | **Deborah N. Schoonhoven**[1,2] | **Matteo Demuru**[1,2] | **Peterjan Ris**[2] | **Philip Scheltens**[1] | **Cornelis J. Stam**[2]

[1] Alzheimer Center and Department of Neurology, VU University medical center, Amsterdam UMC, Amsterdam, The Netherlands

[2] Department of Clinical Neurophysiology and MEG Center, Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam UMC, Amsterdam, The Netherlands

Correspondence
Alida A. Gouw, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands.
Email: AA.Gouw@amsterdamumc.nl

Abstract

Introduction: We report the routine application of magnetoencephalography (MEG) in a memory clinic, and its value in the discrimination of patients with Alzheimer's disease (AD) dementia from controls.

Methods: Three hundred sixty-six patients visiting our memory clinic underwent MEG recording. Source-reconstructed MEG data were visually assessed and evaluated in the context of clinical findings and other diagnostic markers. We analyzed the diagnostic accuracy of MEG spectral measures in the discrimination of individual AD dementia patients (n = 40) from subjective cognitive decline (SCD) patients (n = 40) using random forest models.

Results: Best discrimination was obtained using a combination of relative theta and delta power (accuracy 0.846, sensitivity 0.855, specificity 0.837). The results were validated in an independent cohort. Hippocampal and thalamic regions, besides temporal-occipital lobes, contributed considerably to the model.

Discussion: MEG has been implemented successfully in the workup of memory clinic patients and has value in diagnostic decision-making.

KEYWORDS
Alzheimer's disease, diagnostic biomarker, machine learning, magnetoencephalography, random forest classifier

1 | BACKGROUND

Diagnostic evaluation of memory clinic patients has shifted from the assessment of clinical symptoms toward biomarker-supported diagnoses.[1,2] Commonly used biomarkers include structural brain imaging, cerebrospinal fluid (CSF) examination, and positron emission tomography (PET). Their diagnostic accuracies, however, vary across dementia types. For Alzheimer's disease (AD), for example, CSF biomarkers and PET imaging supportive for AD pathology have greatly increased the specificity of the diagnosis,[3] but for Lewy body dementia (LBD) and frontotemporal dementia (FTD), specific molecular biomarkers are not (yet) available. These diagnoses therefore rely on biomarkers that are not based directly on underlying pathology.[2,4] Other supportive biomarkers may be helpful to further improve the diagnostic process.

Normal cognition relies on optimal neuronal functioning and information transfer via synapses. A non-invasive method to assess synaptic functioning is electroencephalography (EEG). It has been established that slowing of the posterior dominant rhythm is supportive of dementia due to AD.[5–7] In patients with subjective cognitive decline (SCD) or

HIGHLIGHTS

- Routine application of magnetoencephalography (MEG) is feasible in the workup of memory clinic patients.
- MEG reliably discriminates patients with Alzheimer's disease dementia from controls.
- The thalamus may be more important in Alzheimer's disease than currently recognized.

RESEARCH IN CONTEXT

1. Systematic review: We reviewed the literature using scientific databases (eg, PubMed) focusing on magnetoencephalography (MEG) as a diagnostic marker for Alzheimer's disease (AD).
2. Interpretation: We have shown that routine application of MEG is feasible in the diagnostic workup of memory clinic patients and report demographics and MEG findings of the prospective cohort (n = 366). We further demonstrated that MEG spectral characteristics have value in the discrimination of individual patients with AD dementia from patients with subjective cognitive decline, especially in the temporal-occipital lobe, hippocampus, and thalamus.
3. Future directions: Multiclass classification of multiple dementia diagnoses will bring MEG closer to its implementation in clinical practice. Our findings also indicate that a more in-depth study of subcortical structures, for example, the thalamus, may lead to novel insights in the pathogenesis of AD.

mild cognitive impairment (MCI), oscillatory slowing, reflected by an increase of relative theta power, was found to be predictive of future cognitive decline.[8–10] In LBD, EEG changes are generally more severe than in dementia due to AD[11,12] and frontal intermittent rhythmic delta activity (FIRDA)[13] is much more common. FTD is generally associated with a remarkably normal EEG.[14] Few EEG studies have been performed in other types of dementia.[15]

An alternative method for measuring synaptic activity is magnetoencephalography (MEG), which detects the magnetic fields induced by postsynaptic currents. Compared to standard clinically used EEG, MEG recording is more patient friendly and time efficient. Technically, it has a higher spatial resolution, it is reference free, and it is more sensitive to deeper sources such as the hippocampus in AD.[16–20] Indeed, in AD patients, MEG oscillatory activity originating specifically from the hippocampus correlated stronger with cognitive decline than cortical activity.[21] These findings point toward a potentially valuable place for MEG in both research and clinical care.

A growing number of studies have moved beyond group differences toward the individual level. These classification studies are essential to evaluate whether EEG/MEG can provide biomarkers for clinical purposes. Methods varied widely with respect to the used classification method and input features, and performance ranged from poor to almost perfect.[22,23] The wide variability of results may be explained by the use of small samples and the lack of validation in an independent data set,[22] so the true validity of EEG/MEG features as biomarkers has not yet been fully established.

The objective of this article is twofold. First, we describe the implementation of MEG as a routine investigation during the 1-day diagnostic workup in our memory clinic. Second, we evaluate its value in the discrimination of AD dementia patients and SCD subjects using a random forest classifier. We expected that especially higher relative theta power, as an early marker for AD, in both cortical and subcortical structures would be most important in the classification of individual patients.

2 | MATERIALS AND METHODS

First, the implementation of MEG is described, followed by the methodology of the classification between AD dementia and SCD patients. (An overview is given in Figure 1.)

2.1 | Methods I: implementation of MEG in the memory clinic

2.1.1 | Diagnostic workup

The standardized 1-day diagnostic screening of the Alzheimer Center Amsterdam consists of a neurological and neuropsychological examination, magnetic resonance imaging (MRI), EEG, standard laboratory work, and lumbar puncture. The EEG protocol has been described previously.[7,24] Since April 2015, a MEG recording has replaced the EEG in an unselected subset of the memory clinic patients (first two patients each Monday). During the multidisciplinary meeting on the Friday of the same week, diagnoses are made using standard criteria after consideration of the clinical findings and diagnostic investigations.[2–4,25] Results of the MEG recording (or EEG) and other diagnostic tests are weighed with respect to the clinical differential diagnosis of the neurologist to come to a final diagnosis. Subjects gave informed consent for the use of their clinical data for research purposes.[24] The ethical review board of the VU University medical center approved this protocol.

2.1.2 | Data acquisition

A standardized recording protocol and pre-processing pipeline was developed so it approaches EEG procedures and it fits in the fast diagnostic routine of the memory clinic. MEG data are recorded using a 306-channel Elekta system (Elekta Neuromag Oy, Helsinki, Finland) inside a magnetically shielded room (Vacuumschmelze, Hanau,

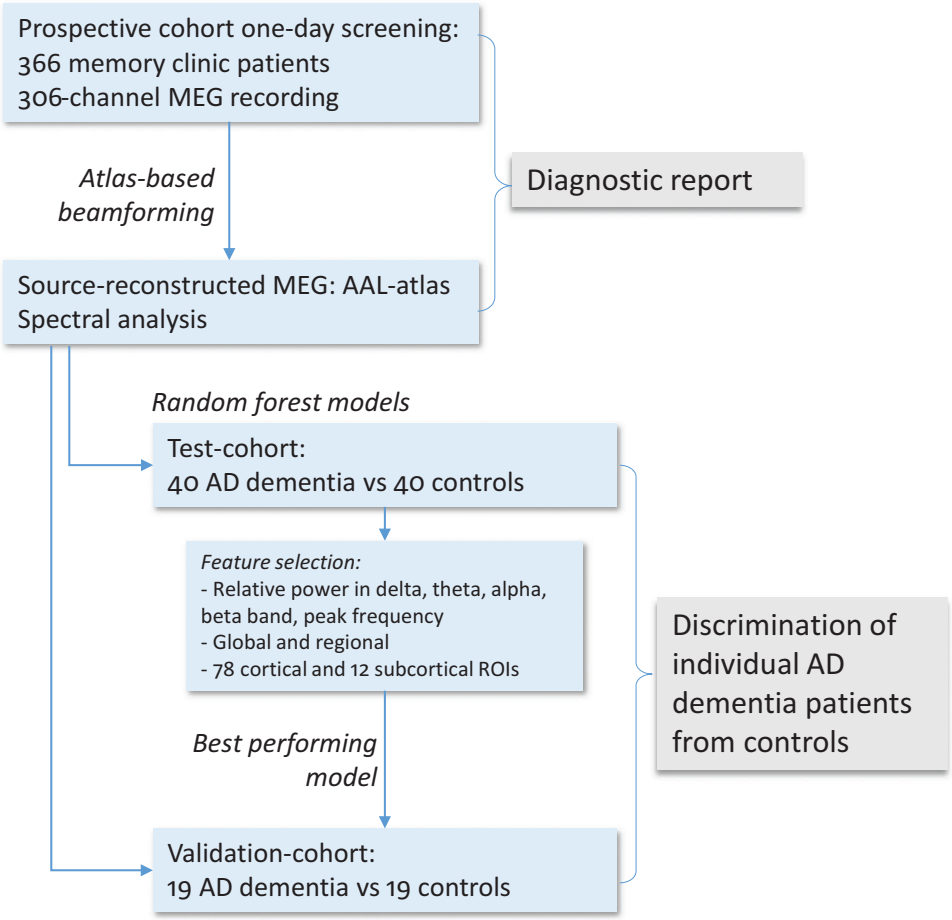

FIGURE 1 Flow chart: Overview of patients and methods

Germany) at a sample frequency of 1250 Hz, with an anti-aliasing filter of 410 Hz and a high-pass filter of 0.1 Hz. Two data sets of 5-minute duration are recorded in a task-free and mainly eyes-closed condition. During the recording, patients are asked to open their eyes one to two times for several seconds. Finally, a working memory test is performed, during which the patient looks at and memorizes 12 pictures during 10 seconds and tries to recall the pictures after 1 minute, providing a picture test score.[26] Technicians closely monitor both the recording for artifacts and the patient via a video system, and alert patients when they get drowsy. The head position relative to the MEG sensors is recorded continuously using signals from five head-localization coils. The positions of the head-localization coil and the outline of the participants scalp (≈500 points) are digitized using a three-dimensional (3D) digitizer (Fastrak, Polhemus, Colchester, VT, USA). Two horizontal and one vertical electro-oculography channel and an ECG channel are recorded.

2.1.3 | Data pre-processing

The pre-processing pipeline consists of spatial filtering and source-reconstruction of the MEG data. A detailed description is given in the Supplementary Materials. In short, the MEG time series were filtered using the temporal extension of Signal Space Separation (tSSS). Then, the digitized scalp surface was co-registered to a best fitting MRI template. After co-registration, the voxels of the MRI were labeled using the automated anatomical labeling (AAL)-atlas,[27] consisting of 116 regions (78 cortical, 12 subcortical, and 26 cerebellar regions). For each region, neuronal activity at each centroid voxels was reconstructed using a scalar beamforming approach.[28]

2.1.4 | Diagnostic MEG report

For diagnostic purposes, time series of 80 regions (78 cortical and 2 hippocampal regions) were assessed. In BrainWave software (developed by CS and freely available via https://home.kpn.nl/stam7883/brainwave.html), 13.12 seconds of MEG data per page are depicted by downsampling to 312 Hz. In addition, sensor-space data were reviewed in Graph (Elekta Neuromag Oy, version 2.94). MEG data were assessed by certified clinical neurophysiologists without knowledge of clinical information, except for age, gender, and medication use.

The diagnostic report consists of (1) a description of background rhythms, focal abnormalities, transients, and/or epileptiform activity,

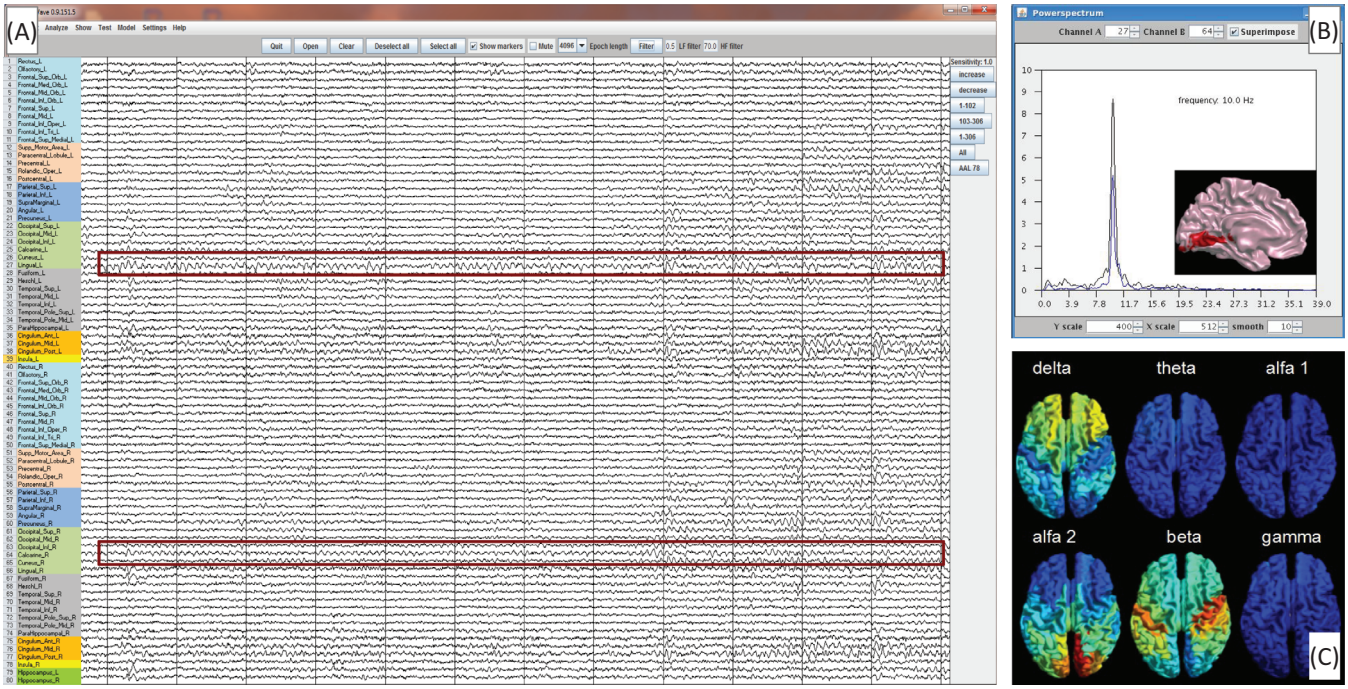

FIGURE 2 Illustration of the standard figures in the diagnostic magnetoencephalography (MEG) report. (A) a representative page of 13.12 second source-space eyes-closed MEG time-series (80 automated anatomical labeling [AAL]-regions). (B) Power spectrum of posterior dominant rhythm, calculated from a single time-series (13.12 seconds) of each occipital lobe. (C) top-view head plots, with globally scaled color-coded relative power in six frequency bands at 78 cortical regions of the AAL-atlas. A warmer color represents higher relative power

based on visual assessment; (2) a table with graphoelements; (3) a semi-quantitative MEG score (explained below); (4) a table with relative band power (delta 0.5–4 Hz, theta 4–8 Hz, alpha1 8–10 Hz, alpha2 10–13 Hz, beta 13–30 Hz, gamma 30–48 Hz) and peak frequency (dominant frequency between 4 and 13 Hz) calculated from representative artifact-free epochs by fast Fourier transformation in BrainWave; and (5) a conclusion.

The semi-quantitative MEG score, similar to the score used in EEG assessment, consists of a four-digit visual rating scheme: (1) severity of abnormalities (range 1–5: normal, mild, moderate, severe abnormalities, isoelectric, respectively), the presence = 1 or absence = 0 of (2) diffuse abnormalities, (3) focal abnormalities, and (4) epileptiform abnormalities. Focal abnormalities are (transient) slow or sharp wave activity in one or more time-series. Diffuse abnormalities are defined as slowing of the rhythmic background activity. Epileptiform abnormalities were defined as (poly)spikes or (poly)spike-slow wave discharges. Inter-observer agreement was previously evaluated with kappa values of 0.60 and 0.87 for detection of focal and diffuse abnormalities, respectively.[29] Figure 2 shows the three standard illustrations that accompany the report.

2.1.5 | Statistical analyses

Diagnosis distributions, patient demographics, and MEG visual scores of the cohort were described and compared between diagnosis groups using chi-square tests and Student *t* tests, where appropriate.

2.2 | Methods II: Classification of AD and SCD patients

We investigated the discriminative ability of MEG spectral features for individual AD dementia patients and SCD subjects using the random forest algorithm in the machine learning module of BrainWave. A subset of 40 consecutive patients with AD dementia were included when they had AD biomarker evidence defined by a CSF t-tau/amyloid beta (Aβ)42 ratio >0.52, and/or a positive fluoride-18 florbetaben-PET.[30] Forty SCD patients were selected on the basis of absence of amyloid pathology (CSF or PET) when available. For each patient, 20 epochs of 4096 samples (3.27 seconds) each (total ≈65.5 seconds) were visually selected based on the absence of artifacts, eyes-closed condition, and alertness of the patient. The following spectral measures were used as features: regional (90 cortical and subcortical AAL regions, excluding cerebellum; supplementary Table 1) and global average of relative power in the delta, theta, full alpha (8–13 Hz), and beta band, and peak frequency.

2.2.1 | Random forest classifier

The classifier was built with the random forest approach[31]. This method builds a decision tree using a random sample of the original data set, with replacement (bootstrapped dataset), and using a random subsample of features at each split of the tree. This process is repeated to create a group of decision trees (a "forest"; the number of trees was

TABLE 1 Patients' characteristics by diagnosis group

	N (%)	Age, mean (SD)	Gender, M / F	Picture test, median (IQR)*
Subjective cognitive decline	97 (26.5%)	57.9 (9.1)	53 / 44	5 (5–7)
Dementia due to Alzheimer's disease	89 (24.3%)	65.7 (7.7)	44 / 45	2 (1–3)
Psychiatric disorder	43 (11.7%)	55.3 (9.0)	27 / 16	5 (4–6)
Mild cognitive impairment	41 (11.2%)	66.2 (7.6)	30 / 11	4 (3–4)
Inconclusive diagnosis	34 (9.3%)	62.7 (8.4)	21 / 13	3.5 (1.5–4.5)
Other dementia or neurological disease	25 (6.9%)	64.2 (7.4)	15 / 10	4 (2–5)
Lewy body dementia	15 (4.1%)	68.4 (5.5)	14 / 1	3 (2–4)
Frontotemporal lobe dementia	14 (3.8%)	63.6 (7.6)	7 / 7	4.5 (3–6)
Vascular dementia	8 (2.2%)	72.6 (5.6)	6 / 2	3 (1.5–4.5)

Picture test (range 0–12), * 3 missing.

set at 500). The number of input variables randomly chosen at each split was set as the square root of the number of features. Ten percent white noise was added to the feature values. The performance of the model is tested using the samples not used in building the decision tree (out-of-bag data set). These samples are run down each decision tree and eventually assigned to one of the classes by majority vote.

The performance of each random forest model was described in terms of accuracy (ratio of number of correctly classified subjects/number of total subjects), sensitivity (ratio of correctly classified SCD patients/all SCD patients), and specificity (ratio of correctly classified AD patients/all AD patients). The contribution of each feature is quantified with the variable importance (VIMP) score (range 0–1), based on the number of times a feature was selected for a split in the tree.

2.2.2 | Validation

We made several models using various combinations of features, risking overfitting. We therefore validated the best-performing classifier in an independent within-center data set. In view of future implementation in the clinic, we used unselected downsampled MEG data (first four epochs/≈52.5 seconds of four times downsampled data) of source-reconstructed time-series for this validation step. This validation set consisted of 19 AD dementia patients (17 biomarker positive by CSF or amyloid-PET; two biomarkers unavailable) and 19 SCD patients (regardless of AD biomarkers) from the prospective cohort.

3 | RESULTS

3.1 | I: Prospective cohort

In the period from April 2015 to September 2019, a total of 370 patients have undergone MEG recording during the diagnostic screening. Four patients did not give consent for the use of their data for research purposes. The remaining patients (n = 366) had a mean age of 62.3 ± 9.1 years, and consisted of 149 (41%) female and 217 (59%) male patients (Table 1). The most frequent diagnosis was SCD, followed by AD dementia. Other diagnoses by descending prevalence were psychiatric disorder, MCI, inconclusive diagnosis, other dementia/neurological disorder, DLB, FTD, and vascular dementia (VaD). The category "other dementia/neurological disorder" included dementia types such as progressive supranuclear palsy, corticobasal degeneration, or other neurological disorders (eg, limbic encephalitis, hydrocephalus, multiple sclerosis). An inconclusive diagnosis was given when more information was awaited, for example, results of amyloid-PET scans or follow-up over time.

3.1.1 | Diagnostic reporting of MEG

All MEG recordings were performed on Monday with the multidisciplinary diagnostic meeting on Friday. MEG preprocessing for each data set took around 2 hours for trained technicians. For all cases, except one, a diagnostic report was completed on time. The MEG of this one case contained continuous diffuse slow waves without a recognizable posterior dominant rhythm. It was regarded as uninterpretable due to (movement) artifacts. In retrospect, this was probably not the case as in the multidisciplinary meeting it became clear that the patient had severe AD dementia based on clinical and MRI findings, and the severe slowing was probably congruent with the clinical diagnosis.

MEG visual scores are reported in Figure 3 and supplementary Table 2. The visual scores varied considerably between diagnoses, ranging from normal (score 1) to moderately abnormal (score 3). None of the patients had a severity score of 4 or 5. In both SCD and psychiatric disorders, the majority had score 1, whereas score 3 was rare. In dementia due to AD, the majority of patients were scored as moderately abnormal, but a considerable proportion had mild abnormalities or even a normal visual score. None of the DLB patients had a normal MEG, and the majority had score 3. FTD patients mostly had a score 1 or 2.

Distributions of diffuse and focal abnormalities per diagnosis group are described in Table 2. DLB and AD patients showed the highest proportion of diffuse abnormalities, in contrast to FTD patients who rarely showed diffuse abnormalities. Two patients (0.5%) had epileptiform abnormalities. Of those, one was diagnosed with moderately advanced AD and the other with cognitive decline due to high dosed antipsychotic medication.

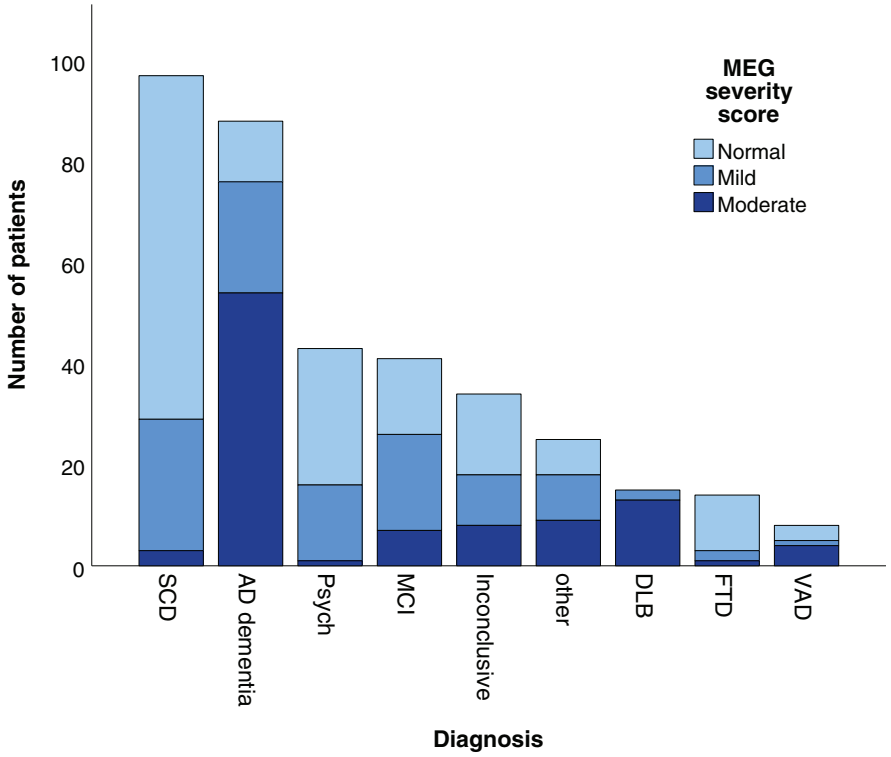

FIGURE 3 Bar graph with distribution of magnetoencephalography (MEG) severity scores by diagnosis group. Groups are ordered by decreasing prevalence

TABLE 2 Diffuse and focal abnormalities by diagnosis group

Presence of abnormalities	Diffuse	Focal	Both diffuse and focal
Total	151 (42%)	138 (38%)	83 (23%)
AD dementia*	66 (75%)	49 (56%)	39 (44%)
Subjective cognitive decline	13 (13%)	22 (23%)	6 (6%)
Psychiatric disorder	12 (28%)	10 (23%)	6 (14%)
Mild cognitive impairment	18 (44%)	16 (39%)	8 (20%)
Frontotemporal lobe dementia	2 (14%)	2 (14%)	1 (7%)
Lewy body dementia	13 (87%)	8 (53%)	6 (40%)
Vascular dementia	5 (63%)	4 (50%)	4 (50%)
Other dementia or neurological disease	11 (44%)	16 (64%)	9 (36%)
Inconclusive diagnosis	11 (32%)	11 (32%)	4 (12%)

Chi-square $P < .001$ for presence of diffuse abnormalities. Chi-square $P < .001$ for presence of focal abnormalities. Chi-square $P < .001$ for presence of both diffuse and focal abnormalities. *MEG of one patient with AD dementia was considered uninterpretable.

3.1.2 | MEG as part of the diagnostic process in individual patients

The results from MEG visual analysis are used as a supportive diagnostic tool during the multidisciplinary meeting alongside other diagnostic tests, by relating the results to the clinical differential diagnosis of the neurologist. For example, when after clinical evaluation a diagnosis of LBD or dementia due to AD is considered and the MEG shows no diffuse abnormalities (ie, no global slowing of the posterior dominant rhythm) with mild focal temporal slow waves (congruent with a MEG visual score of 2001), the findings point toward a dementia due to AD diagnosis. For LBD, one would have expected more severe abnormalities and diffuse slowing, possibly with FIRDA. Another example involves patients with changes in behavior and frontal symptoms, which may be caused by the behavioral variant of FTD or the frontal subtype of AD. When the MEG shows diffuse abnormalities, it strongly supports a diagnosis of AD, as diffuse abnormalities are uncommon in FTD.

3.2 | II: Classification of AD and SCD patients

3.2.1 | Patient cohort for classification

All 40 patients with dementia due to AD were biomarker positive based on CSF and/or PET. Thirty-three SCD patients (83%) had negative amyloid biomarkers. The remaining SCD patients had no available AD biomarkers. Of those, four patients were stable at yearly follow-up evaluations (1–2 years), whereas three patients were not followed-up. The SCD group was significantly younger (57.2 ± 8.4 years) than the AD group (66.0 ± 7.5 years; $t(78) = 4.98$, $P < .001$). Gender distribution did not differ (19 [48%] female vs 23 [58%] for SCD and AD respectively; chi-square (1,N = 80) = 0.80, $P = .37$).

TABLE 3 Performance of random forest classifier to discriminate individual AD dementia patients from controls

Random forest models	Accuracy	Sensitivity	Specificity
Single MEG measures (91 features)			
Delta power	0.714	0.676	0.753
Theta power	0.844	0.860	0.827
Alfa power	0.700	0.736	0.663
Beta power	0.751	0.751	0.750
Peak frequency	0.808	0.820	0.795
Combination of two MEG measures (182 features)			
Delta + theta power	***0.851***	***0.868***	***0.835***
Theta + alfa power	0.811	0.795	0.828
Theta + beta power	0.783	0.754	0.813
Theta + peak frequency	0.824	0.832	0.815
Combination of three MEG measures (273 features)			
Delta + theta + alpha power	0.828	0.825	0.832
Delta + theta + beta power	0.821	0.829	0.812
Theta + alfa + beta power	0.798	0.782	0.815
Delta + theta power + peak frequency	0.838	0.855	0.822
Theta + alfa power + peak frequency	0.806	0.805	0.808
Theta + beta power + peak frequency	0.810	0.810	0.810
Combination four MEG measures (364 features)			
Delta + theta + alpha power + beta power	0.816	0.814	0.818
Delta + theta + alpha power + peak frequency	0.820	0.820	0.820
Delta + theta + beta power + peak frequency	0.832	0.855	0.810
Theta + alpha + beta power + peak frequency	0.792	0.784	0.800
Combination of all five MEG measures (455 features)			
Delta + theta + alpha + beta power + peak frequency	0.821	0.835	0.808

Model with highest accuracy, sensitivity, and specificity values is depicted in bold.

Several models were systematically constructed (Table 3). First, each MEG spectral measure was entered separately (1 global mean and 90 regional values = 91 features). Relative theta power yielded the highest accuracy of 0.844, sensitivity of 0.860, and specificity of 0.827. Peak frequency also showed an accuracy higher than 0.8. Then, models were repeated with combinations of theta power and, respectively, one, two, and three other spectral measures. Finally, all five spectral measures were entered simultaneously. The combination of delta and theta power yielded overall highest discriminatory value (accuracy 0.851, sensitivity 0.868, and specificity 0.835). Figure 4 illustrates the output of this model. The VIMP scores demonstrate that relative theta power contributed more than relative delta power. With respect to cortical regions, occipital and temporal lobes were more important than frontal and parietal lobes. Furthermore, subcortical regions contributed significantly to the model. In particular, the bilateral hippocampus and thalamus had the highest VIMP scores of all features.

3.2.2 | Validation in an independent cohort

Validation of the best-performing random forest model yielded an accuracy of 0.842, a sensitivity of 0.789, and a specificity of 0.895. These results indicate that the model is robust against overfitting. In other words, it is generalizable and does not lose validity when using unselected and downsampled MEG data.

4 | DISCUSSION

We described the implementation of MEG in the diagnostic workup of memory clinic patients by reporting our recording, pre-processing, and reporting procedures, designed to fit the swift diagnostic process of our memory clinic. We further evaluated the ability of MEG to discriminate individual AD dementia patients from SCD patients using a random forest classifier. With widely used and simple spectral features, a model with relative theta and delta power resulted in an accuracy of 0.851. Validation of this model in an independent cohort yielded comparable results, demonstrating the generalizability of the classifier. Temporal and occipital cortical regions, as well as hippocampal and thalamic regions, contributed considerably to the model.

4.1 | MEG: from research to clinic

Several articles have advocated the use of MEG as a diagnostic tool in dementia[20,32–34] but, in the neurodegenerative field, MEG has up to now been used only in research settings. Possible reasons include the limited availability of MEG to memory clinics, the time-consuming (pre-)processing steps, and difficulties in the extraction of MEG features that are translatable to the clinic. The availability of a MEG system next door to the memory clinic and the long-standing experience with the use of EEG in memory clinic patients have been supportive for the implementation of MEG in our center.

4.2 | MEG visual analysis

The most prevalent diagnoses in our unselected cohort were dementia due to AD and SCD. As it is well-recognized in literature, one would expect diffuse oscillatory slowing in AD dementia patients. This was indeed the case in the majority of these patients, but a normal MEG based on visual assessment was found in 14%. This is a remarkable

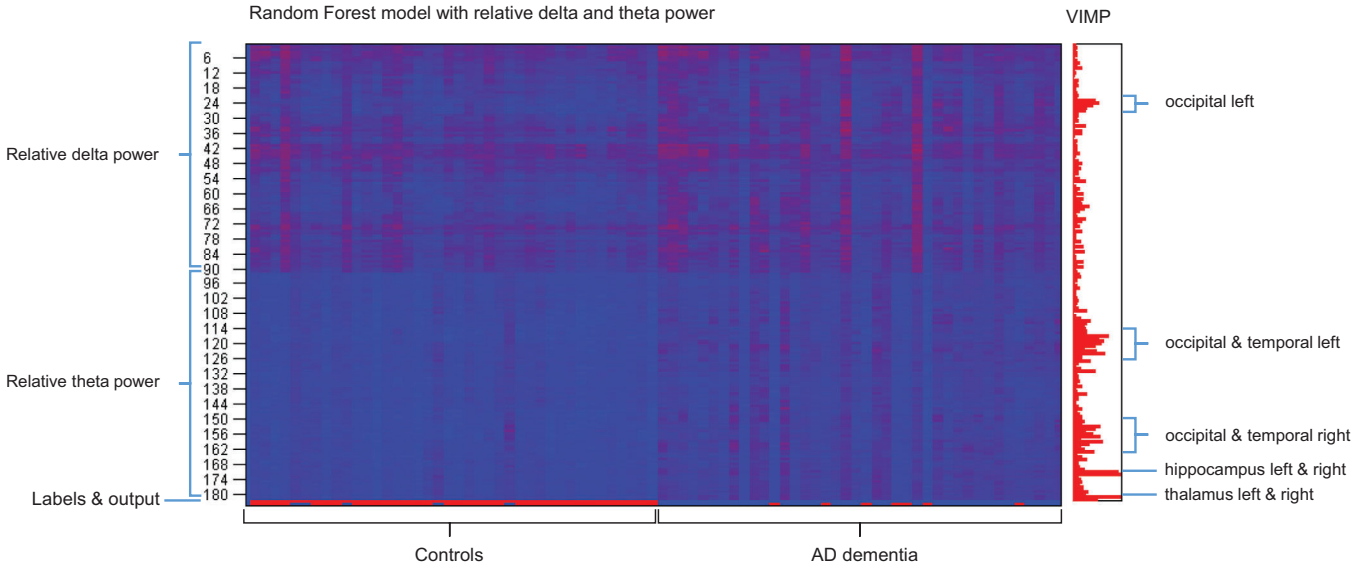

FIGURE 4 Output of the machine learning module of BrainWave, after running a random forest model with global and regional relative delta and theta power (feature nr 1–91 for delta power and 92 to 182 for theta power on the vertical axis) on a data set of 40 controls and 40 AD dementia patients (on the horizontal-axis). Values of the spectral measures are depicted as a blue to red scale (higher values are deeper red). Patient labels (diagnosis: red = SCD; blue = AD) are set in the row below the MEG features (row 183). The output of the model, that is, classification of individual patients, is given in the bottom row. Variable importance (VIMP) scores for each MEG feature are shown in a tilted histogram on the right side of the figure. The length of each bar indicates the relative importance of the corresponding feature to the classification

finding and may hypothetically represent a disease mechanism with little synaptic damage and a relatively "benign" disease course. In contrast, none of the LBD patients had a normal MEG, but mostly moderate diffuse abnormalities. These findings are consistent with previous EEG reports.[11,12,35] Together with the existing literature and the fact that there are no pathology-specific markers for LBD, we propose that MEG/EEG should have a more prominent position in the diagnosis of LBD.[36]

In our memory clinic, each patient receives a standard set of diagnostic tests, as we strive for a design where patient care and research go hand in hand.[24] This standardized design, where MEG is performed and analyzed without *a priori* knowledge of clinical differential diagnosis, means that the clinical value of MEG in our cohort varies from patient to patient. When the diagnosis is already obvious from clinical assessment or due to strong support by another test (eg, in the case of MRI findings in vascular dementia), MEG will not add much to the diagnostic process, but in other cases MEG may provide crucial information, as highlighted in the example cases in the results section and previous paragraphs. Hence, other centers may consider a scenario where MEG is only performed in those cases where other investigations did not lead to a decisive diagnosis.

4.3 | Classification of AD patients from SCD subjects

We assessed the ability of quantitative MEG to classify individual AD dementia patients and SCD subjects. Using relative theta and delta power, adequate values of accuracy of 0.851, sensitivity 0.868, and specificity 0.835 were achieved. Few studies have classified AD dementia patients from controls with MEG. One study[37] used decomposed MEG epochs with a blind source separation (BSS) technique. Median frequency resulted in an accuracy of 83.33%, whereas spectral entropy reached an accuracy of 73.81%. A study by[38] reported ROC-based classification performance of sample entropy and Lempel-Ziv complexity, showing an accuracy of 70.7% and 78.05%, respectively, increasing to 85.37% when the features were entered in an adaptive network–based fuzzy interference system. Another study[39] used several measures from information theory, a disequilibrium measure and three measures of statistical complexity, yielding an accuracy of 83.9% with linear discriminant analysis. Finally, a recent study reported the performance of functional connectivity using support vector machine classification, which resulted in a sensitivity of 76.7% and a specificity of 65.3%.[40] Compared to these reports, the classifier in the present study yielded a similar or higher accuracy in a larger and better characterized data set, and perhaps most importantly, with validation in an independent data set. Especially studies with small sample sizes in combination with powerful machine learning algorithms and feature selection may suffer from "overfitting."[22] Moreover, in contrast to the complex nature of the features of some previous studies, the present study used easily interpretable features, which are better suitable for clinical purposes.

Of note are MEG studies that focused on the MCI stage of AD with larger sample sizes in a multi-site design and/or external validation. MEG characteristics entered into the classifiers included mutual information measures, oscillatory peak frequency, relative alpha power, or functional network topology measures. Performance of the classifiers varied widely, from just above a chance level to very high.[41–43] It is important to note that none of these studies used biomarker evidence

of AD pathology. It is becoming increasingly recognized that AD is defined by its underlying pathological processes, and biomarker proof is therefore essential for research purposes.[1,34]

EEG studies focusing on the classification of AD dementia patients are reviewed in a recent paper.[22] A large study evaluated AD and MCI patients with biomarker support and controls using absolute and relative band power.[9] Pairwise classification of AD dementia patients and controls resulted in a moderate accuracy of 72.9%, whereas the classification of MCI patients and controls showed an accuracy of 63.1%. Several other mostly smaller studies reported results that ranged from poor to almost perfect.[35,44–48]

4.4 | Subcortical brain regions in AD

It has been demonstrated that source-reconstructed MEG can detect neuronal activity from deeper structures.[16,17,19] We combined source-space MEG with a random forest classifier because, unlike many other machine learning methods, it provides the contribution of each feature to the final model. This gives important insights into whether the classifier is based on solid pathophysiological differences between patients, making the classifier more easily accepted for clinical purposes than "black-box" machine learning methods. It may in turn lead to new hypotheses on pathophysiological mechanisms. Indeed, our results indicated that besides the cortical temporal and occipital regions, the hippocampus and thalamus contributed strongly to the model. These findings confirm the well-known relevance of the hippocampus in AD, but in AD research, the thalamus has received relatively little attention. Several recent (preclinical) studies indicate early changes in the thalamus and dysregulation of the corticothalamic network in AD.[49,50] Together with our results, these reports suggest that a more in-depth study of the thalamus may provide novel insights into pathophysiology of AD.

4.5 | Strengths and limitations

The strengths of the present study are the use of a relatively large, well-characterized data set from a prospective clinical cohort. The AD dementia patients all have biomarker proof of underlying AD pathology. We also confirmed generalizability of the classifier in an independent data set. Among limitations is the use of MRI templates rather than individual native MRI scans. We therefore recently implemented the use of individual MRI scans instead of the templates (beginning in 2021), possibly increasing diagnostic accuracy further. The most important limitation is that only pairwise classification between the two most prevalent diagnosis groups, AD dementia versus SCD patients, was performed. To be readily useful as a supportive marker in the memory clinic, a multi-class classifier differentiating multiple dementia diagnoses should be built, that is, "a MEG-based differential diagnosis." With the steady growth of the prospective cohort, we hope that this aim will be within reach in the near future.

5 | CONCLUSION

This study describes the implementation of MEG in the standardized diagnostic workup of memory clinic patients. MEG spectral features discriminate individual patients with AD dementia from SCD subjects with high accuracy and good generalizability. Relative theta and delta power in both cortical and subcortical regions, including the hippocampi, contributed strongly to the classifier. Our results also indicated that thalamic dysfunction is potentially more important in the pathogenesis of AD than currently acknowledged.

ACKNOWLEDGMENTS

Research of the Alzheimer Centre Amsterdam is part of the neurodegeneration research program of Amsterdam Neuroscience. The Alzheimer Centre Amsterdam is supported by Stichting Alzheimer Nederland and Stichting VUmc fonds. The authors would like to acknowledge C. Plugge, M. van den Hoek, and N. Zwagerman for preprocessing of the MEG recordings.

CONFLICTS OF INTERESTS

A.A. Gouw has received past (2014-2018) research support from Probiodrug and Boehringer Ingelheim via the VUmc Alzheimer Center. P. Scheltens has received consultancy fees (paid to the institution) from AC Immune, Alkermes, Alnylam, Alzheon, Anavex, Biogen, Brainstorm Cell, Cortexyme, Denali, EIP, ImmunoBrain Checkpoint, GemVax, Genentech, Green Valley, Novartis, Novo Nordisk, PeopleBio, Renew LLC, and Roche. He is principal investigator of studies with AC Immune, CogRx, FUJI-film/Toyama, IONIS, UCB, and Vivoryon. He is a part-time employee of Life Sciences Partners Amsterdam. He serves on the board of Brain Research Center and New Amsterdam Pharma.

D.N. Schoonhoven, C.J. Stam, A. Hillebrand, M. Demuru, and P. Ris declare no conflict of interest.

REFERENCES

1. Jack CR Jr, Bennett DA, Blennow K, Research Framework NIA-AA. Toward a biological definition of Alzheimer's disease. *Alzheimers Dement*. 2018;14:535-562.
2. McKeith IG, Boeve BF, Dickson DW, et al. Diagnosis and management of dementia with Lewy bodies: fourth consensus report of the DLB Consortium. *Neurology*. 2017;89:88-100.
3. McKhann GM, Knopman DS, Chertkow H, et al. 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. *Alzheimers Dement*. 2011;7:263-269.
4. Rascovsky K, Hodges JR, Knopman D, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*. 2011;134:2456-2477.
5. Jeong J. EEG dynamics in patients with Alzheimer's disease. *Clin Neurophysiol*. 2004;115:1490-1505.
6. Babiloni C, Vecchio F, Lizio R, et al. Resting state cortical rhythms in mild cognitive impairment and Alzheimer's disease: electroencephalographic evidence. *J Alzheimers Dis*. 2011;26 Suppl 3:201-214.
7. de Waal H, Stam CJ, de Haan W, van Straaten EC, Scheltens P, van der Flier WM. Young Alzheimer patients show distinct regional changes of oscillatory brain dynamics. *Neurobiol Aging*. 2012;33:1008 e25-31.

8. Gouw AA, Alsema AM, Tijms BM, et al. EEG spectral analysis as a putative early prognostic biomarker in nondemented, amyloid positive subjects. *Neurobiol Aging*. 2017;57:133-142.
9. Musaeus CS, Engedal K, Høgh P, et al. EEG theta power is an early marker of cognitive decline in dementia due to Alzheimer's disease. *J Alzheimers Dis*. 2018;64:1359-1371.
10. Moretti DV, Frisoni GB, Fracassi C, et al. MCI patients' EEGs show group differences between those who progress and those who do not progress to AD. *Neurobiol Aging*. 2011;32:563-571.
11. van der Zande JJ, Gouw AA, van Steenoven I, Scheltens P, Stam CJ, Lemstra AW. EEG characteristics of dementia with Lewy bodies, Alzheimer's disease and mixed pathology. *Front Aging Neurosci*. 2018;10:190.
12. Bonanni L, Thomas A, Tiraboschi P, Perfetti B, Varanese S, Onofrj M. EEG comparisons in early Alzheimer's disease, dementia with Lewy bodies and Parkinson's disease with dementia patients with a 2-year follow-up. *Brain*. 2008;131:690-705.
13. Lee H, Brekelmans GJ, Roks G. The EEG as a diagnostic tool in distinguishing between dementia with Lewy bodies and Alzheimer's disease. *Clin Neurophysiol*. 2015;126:1735-1739.
14. Lindau M, Jelic V, Johansson SE, Andersen C, Wahlund LO, Almkvist O. Quantitative EEG abnormalities and cognitive dysfunctions in frontotemporal dementia and Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2003;15:106-114.
15. Gawel M, Zalewska E, Szmidt-Salkowska E, Kowalski J. The value of quantitative EEG in differential diagnosis of Alzheimer's disease and subcortical vascular dementia. *J Neurol Sci*. 2009;283:127-133.
16. Ruzich E, Crespo-García M, Dalal SS, Schneiderman JF. Characterizing hippocampal dynamics with MEG: a systematic review and evidence-based guidelines. *Hum Brain Mapp*. 2019;40:1353-1375.
17. Hillebrand A, Nissen IA, Ris-Hilgersom I, et al. Detecting epileptiform activity from deeper brain regions in spatially filtered MEG data. *Clin Neurophysiol*. 2016;127:2766-2769.
18. Hillebrand A, Gaetz W, Furlong PL, Gouw AA, Stam CJ. Practical guidelines for clinical magnetoencephalography - Another step towards best practice. *Clin Neurophysiol*. 2018;129:1709-1711.
19. Attal Y, Schwartz D. Assessment of subcortical source localization using deep brain activity imaging model with minimum norm operators: a MEG study. *PLoS One*. 2013;8:e59856.
20. Hari R, Baillet S, Barnes G, et al. IFCN-endorsed practical guidelines for clinical magnetoencephalography (MEG). *Clin Neurophysiol*. 2018;129:1720-1747.
21. Engels MM, Hillebrand A, van der Flier WM, Stam CJ, Scheltens P, van Straaten EC. Slowing of hippocampal activity correlates with cognitive decline in early onset Alzheimer's disease. An MEG study with virtual electrodes. *Front Hum Neurosci*. 2016;10:238.
22. Yang S, Bornot JMS, Wong-Lin K, Prasad G. M/EEG-based bio-markers to predict the MCI and Alzheimer's disease: a review from the ML perspective. *IEEE transactions on bio-medical engineering*. 2019;66:2924-2935.
23. Mandal PK, Banerjee A, Tripathi M, Sharma A. A comprehensive review of magnetoencephalography (MEG) studies for brain functionality in healthy aging and Alzheimer's disease (AD). *Front Comput Neurosci*. 2018;12:60.
24. van der Flier WM, Pijnenburg YA, Prins N, et al. Optimizing patient care and research: the Amsterdam Dementia Cohort. *J Alzheimers Dis*. 2014;41:313-327.
25. Sachdev P, Kalaria R, O'Brien J, et al. Diagnostic criteria for vascular cognitive disorders: a VASCOG statement. *Alzheimer disease and associated disorders*. 2014;28:206-218.
26. Stam CJ, van Cappellen van Walsum AM, Micheloyannis S. Variability of EEG synchronization during a working memory task in healthy subjects. *Int J Psychophysiol*. 2002;46:53-66.
27. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*. 2002;15:273-289.
28. Hillebrand A, Tewarie P, van Dellen E, et al. Direction of information flow in large-scale resting-state networks is frequency-dependent. *Proc Natl Acad Sci U S A*. 2016;113:3867-3872.
29. Liedorp M, van der Flier WM, Hoogervorst EL, Scheltens P, Stam CJ. Associations between patterns of EEG abnormalities and diagnosis in a large memory clinic cohort. *Dement Geriatr Cogn Disord*. 2009;27:18-23.
30. Duits FH, Teunissen CE, Bouwman FH, et al. The cerebrospinal fluid "Alzheimer profile": easily said, but what does it mean?. *Alzheimers Dement*. 2014;10:713-723 e2.
31. Breiman L. Random Forests. *Machine Learning*. 2001;45:5-32.
32. Williams MA, Sachdev PS. Magnetoencephalography in neuropsychiatry: ready for application?. *Curr Opin Psychiatry*. 2010;23:273-277.
33. Nakamura A, Cuesta P, Fernández A, et al. Electromagnetic signatures of the preclinical and prodromal stages of Alzheimer's disease. *Brain*. 2018;141:1470-1485.
34. Maestú F, Cuesta P, Hasan O, Fernandéz A, Funke M, Schulz PE. The importance of the validation of M/EEG with current biomarkers in Alzheimer's disease. *Front Hum Neurosci*. 2019;13:17.
35. Dauwan M, van der Zande JJ, van Dellen E, et al. Random forest to differentiate dementia with Lewy bodies from Alzheimer's disease. *Alzheimers Dement (Amst)*. 2016;4:99-106.
36. van der Zande JJ, Gouw AA, van Steenhoven I, et al. Diagnostic and prognostic value of EEG in prodromal dementia with Lewy bodies. *Neurology*. 2020. in press.
37. Escudero J, Hornero R, Poza J, Abásolo D, Fernández A. Assessment of classification improvement in patients with Alzheimer's disease based on magnetoencephalogram blind source separation. *Artif Intell Med*. 2008;43:75-85.
38. Gómez C, Hornero R, Abásolo D, Fernández A, Escudero J. Analysis of MEG background activity in Alzheimer's disease using nonlinear methods and ANFIS. *Ann Biomed Eng*. 2009;37:586-594.
39. Bruña R, Poza J, Gómez C, García M, Fernández A, Hornero R. Analysis of spontaneous MEG activity in mild cognitive impairment and Alzheimer's disease using spectral entropies and statistical complexity measures. *J Neural Eng*. 2012;9:036007.
40. Koelewijn L, Lancaster TM, Linden D, et al. Oscillatory hyperactivity and hyperconnectivity in young APOE-ε4 carriers and hypoconnectivity in Alzheimer's disease. *Elife*. 2019;8.
41. Maestu F, Pena JM, Garces P, et al. A multicenter study of the early detection of synaptic dysfunction in mild cognitive impairment using magnetoencephalography-derived functional connectivity. *Neuroimage Clin*. 2015;9:103-109.
42. Hughes LE, Henson RN, Pereda E, et al. Biomagnetic biomarkers for dementia: a pilot multicentre study with a recommended methodological framework for magnetoencephalography. *Alzheimers Dement (Amst)*. 2019;11:450-462.
43. López-Sanz D, Bruña R, Delgado-Losada ML, et al. Electrophysiological brain signatures for the classification of subjective cognitive decline: towards an individual detection in the preclinical stages of dementia. *Alzheimers Res Ther*. 2019;11:49.
44. Ommundsen N, Engedal K, Oksengard AR. Validity of the quantitative EEG statistical pattern recognition method in diagnosing Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2011;31:195-201.
45. Kanda PA, Trambaiolli LR, Lorena AC, et al. Clinician's road map to wavelet EEG as an Alzheimer's disease biomarker. *Clin EEG Neurosci*. 2014;45:104-112.
46. Trambaiolli LR, Spolaor N, Lorena AC, Anghinah R, Sato JR. Feature selection before EEG classification supports the diagnosis of Alzheimer's disease. *Clin Neurophysiol*. 2017;128:2058-2067.

47. Babiloni C, Triggiani AI, Lizio R, et al. Classification of single normal and Alzheimer's disease individuals from cortical sources of resting state EEG rhythms. *Front Neurosci*. 2016;10:47.
48. Triggiani AI, Bevilacqua V, Brunetti A, et al. Classification of healthy subjects and Alzheimer's disease patients with dementia from cortical sources of resting state EEG rhythms: a study using artificial neural networks. *Front Neurosci*. 2016;10:604.
49. Aggleton JP, Pralus A, Nelson AJ, Hornberger M. Thalamic pathology and memory loss in early Alzheimer's disease: moving the focus from the medial temporal lobe to Papez circuit. *Brain*. 2016;139:1877-1890.
50. Busche MA, Kekuš M, Adelsberger H, et al. Rescue of long-range circuit dysfunction in Alzheimer's disease models. *Nat Neurosci*. 2015;18:1623-1630.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Gouw AA, Hillebrand A, Schoonhoven DN, et al. Routine magnetoencephalography in memory clinic patients: A machine learning approach. *Alzheimer's Dement*. 2021;13:e12227. https://doi.org/10.1002/dad2.12227