Maestú, Fernando; Peña, Jose-Maria; Garcés, Pilar; González, Santiago; Bajo, Ricardo; Bagic, Anto; Cuesta, Pablo; Funke, Michael; Mäkelä, Jyrki P.; Menasalvas, Ernestina; Nakamura, Akinori; Parkkonen, Lauri; López, Maria E.; del Pozo, Francisco; Sudre, Gustavo; Zamrini, Edward; Pekkonen, Eero; Henson, Richard N.; Becker, James T.

A multicenter study of the early detection of synaptic dysfunction in Mild Cognitive Impairment using magnetoencephalography-derived functional connectivity

Published in: NEUROIMAGE. CLINICAL

DOI: 10.1016/j.nicl.2015.07.011

Published: 01/01/2015

Document Version
Publisher's PDF, also known as Version of record

Published under the following license: CC BY-NC-ND

Please cite the original version: Maestú, F., Peña, J.-M., Garcés, P., González, S., Bajo, R., Bagic, A., Cuesta, P., Funke, M., Mäkelä, J. P., Menasalvas, E., Nakamura, A., Parkkonen, L., López, M. E., del Pozo, F., Sudre, G., Zamrini, E., Pekkonen, E., Henson, R. N., & Becker, J. T. (2015). A multicenter study of the early detection of synaptic dysfunction in Mild Cognitive Impairment using magnetoencephalography-derived functional connectivity. NEUROIMAGE. CLINICAL, 9, 103-109. https://doi.org/10.1016/j.nicl.2015.07.011

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.
A multicenter study of the early detection of synaptic dysfunction in Mild Cognitive Impairment using Magnetoencephalography-derived functional connectivity

Fernando Maestú PhD, Jose-María Peña PhD, Pilar García MS, Santiago González PhD, Ricardo Bajo PhD, Anto Bagić MD, PhD, Pablo Cuesta MS, Michael Funke MD, Jyrki P. Mäkelä MD, Ernestina Menasalvas PhD, Akinori Nakamura MD, Lauri Parkkonen PhD, Maria E. López PhD, Francisco del Pozo PhD, Gustavo Sudre PhD, Edward Zamrini MD, Eero Pekkonen MD, Richard N. Henson PhD, James T. Becker PhD, Magnetoencephalography International Consortium of Alzheimer’s Disease

1. Introduction

Dementia of the Alzheimer’s type (DAT) is the major cause of clinical dementia in the elderly (Qiu et al., 2009), and is characterized by the accumulation of the Beta amyloid protein, the phosphorylation of the Tau protein, and the loss of synapses. Amyloid deposition impairs normal inter-neuronal connectivity (Garcia-Marin et al., 2009), whereas Tau results in disruption of axonal microtubule bundle organization (Taniguchi et al., 2001). The progressive loss of the number and efficiency of synapses disrupts inter- and intra-regional communication, leading to the idea that the DAT is a disconnection syndrome (Delbeuck et al., 2003;
Selkoe, 2002; Morrison et al., 1991). As pathological changes associated with DAT start decades before the clinical symptoms appear, it is important to determine whether the pathophysiological changes, especially those at the level of the synapse, can be detected prior to the development of DAT.

Synaptic dysfunction and disruption of connectivity can be studied with Magnetoencephalography (MEG). MEG records the magnetic fields induced by intracellular postsynaptic activity (Hämäläinen, 1993), providing a direct measure of neuronal field potentials that can be used to assess the organization of brain functional architecture in DAT (Stam et al., 2009). The physiopathological characteristics of this disease could manifest differently at different stages of the disease. Thus, while advanced stages of DAT may be associated with functional disconnection, earlier stages may be apparent in terms of communication disruption (Buldu et al., 2011). Indeed, MEG studies of patients with Mild Cognitive Impairment (MCI), the intermediate clinical stage between normal cognition and dementia (Petersen, 2004), find that alterations in neuronal organization across the cortex seem to precede clinical dementia. MCI patients have increased synchronization over prefrontal and posterior regions (Bajo et al., 2010), and those who develop dementia within 2 years have higher synchronization than those who remain with MCI (Bajo et al., 2012; Lopez et al., 2014). Thus, hypersynchronization could be a hallmark of network disruption at early clinical stages of the disease.

However, some of the existing MEG studies of MCI and DAT evaluated differences at the group level, but did not use blinded design or used only a small local sample of patients. For MEG to have the greatest utility in clinical practice, it must be able to detect changes in network dysfunction at the individual level, regardless of patient sample. The purpose of this study was to determine whether MEG could accurately classify individual MCI patients relative to cognitively normal elders.

To accomplish this goal, we designed a blinded study that combined rest state MEG recordings from 78 MCI patients and 54 controls from a single laboratory in Madrid. Datasets 2 and 3 contained data from four other MEG centers, and used advanced data mining methods in order to extract features of MEG connectivity that best differentiated the patients from controls.

2. Methods

2.1. Study design

This study was executed in two training and two testing stages, using three separate datasets. Dataset 1 consisted of resting state MEG recordings from 78 MCI patients and 54 controls from a single laboratory in Madrid. Datasets 2 and 3 contained data from four other MEG centers (Dataset 2: 13 MCI patients and 15 controls, Dataset 3: 11 MCI patients and 13 controls) (see Table 1). In a first training stage, Dataset 1 was used to characterize the functional links that best discriminated MCI patients from controls, and the resulting model was tested with Datasets 2 and 3. This tested how a classifier trained on data from one site generalized to data from other sites. We then evaluated whether a classifier trained on multiple datasets was superior to one trained on a single dataset using Datasets 1 and 2, and tested on Dataset 3.

2.2. Subjects

102 MCI patients (mean age 73.7 ± 5.1; 55% are female) and 82 age-matched controls (mean age 69.6 ± 4.6; 72% are female) participated in the study. They were recruited from five sites (Pittsburgh, Cambridge, Helsinki, Madrid, and Obu-Nagoya) as part of the activities of the MAGIC-AD group (Magnetoencephalography International Consortium for Alzheimer’s Disease; see Zamrini et al., 2011).

Albert et al. (2011) proposed a terminology for classifying individuals with “MCI due to Alzheimer’s Disease” with varying levels of certainty. When no biomarkers (e.g., amyloid PET or Tau values) are available, or in amyloid negative cases, the category of “MCI-core clinical criteria” can be still used. Thus, in this study we use this definition of MCI, as well the standard research criteria (Albert et al., 2011; Petersen, 2004), which require: 1) cognitive complaints corroborated by an informant; 2) objective cognitive impairment; 3) normal general cognitive function; 4) relatively preserved activities of daily living; and 5) not meeting the criteria for dementia. All participants in this group were classified with amnestic MCI. Participants were excluded if they had a history of any significant neurological disease, psychiatric disorders, or any significant systemic illness (e.g., advanced cancer or acute heart disease). The study was approved by the ethics committee at each MEG center, and all subjects gave written informed consent prior to participation.

2.3. Procedures

2.3.1. MEG acquisition

Each center used the same MEG protocol under similar conditions. Three to 5 minutes of eyes-closed resting state data were recorded while the participants were seated in a 306-channel Vectorview system (Elekta Oy, Helsinki, Finland) housed in a magnetically shielded room. MEG data were recorded at a sampling rate of 1000 Hz in Cambridge, Obu-Nagoya, Pittsburgh and Madrid, and at 1001.6 Hz in Helsinki (except for 4 subjects at 600–643 Hz). An online bandpass filter at 0.03–330 Hz was applied to all data at each site. The position of the head relative to the sensor array was monitored by four head position indicator coils attached to the scalp. For most subjects (108/132 in Dataset 1, 25/28 in Dataset 2, 24/24 in Dataset 3), head position was monitored continuously during the MEG recording, while for the remaining subjects only the initial head position was estimated. Electrooculograms were used to monitor eye movements (except for 4 subjects). A temporal signal space separation (tSSS) was applied to the MEG raw data in order to eliminate the contribution of non-brain sources to the MEG data, as proven useful in previous studies (Taulu and Simola, 2006; Nonenon et al., 2012; Gonzalez-Moreno et al., 2014). Maxfilter software (version 2.2, Elekta Neuromag) was used to perform the tSSS, along with a coordinate transformation into a common sensor space. We compared the level of magnetic noise in empty room recordings (i.e., without a subject present) and found that none of the centers had a noise level more than two standard deviations higher than the mean of the others. Thus recording sites were considered comparable.

2.3.2. MEG signal processing

MEG data preprocessing was performed with FieldTrip (Oostenveld et al., 2011), and was performed blind to diagnosis prior to the application of the classification algorithm. MEG recordings were filtered into classic frequency bands (Theta (4–8 Hz), Alpha (8–12 Hz), Beta (12–30 Hz), Gamma (30–45 Hz) and broadband (2–45 Hz)) with Finite Impulse Response (FIR) filters of order 1500. As delta activity (2–4 Hz) contains little time-related information in 2-s time windows (~8 oscillations), delta band was not employed for the present functional connectivity analysis. Then, the continuous resting state data were split into 2-s epochs. Ocular, muscular and jump artifacts were detected with

Table 1

| Number of participants from each MEG center for each dataset. | Dataset 1 | Dataset 2 | Dataset 3 |
|---|---|---|---|
| | NC | MCI | NC | MCI | NC | MCI |
| Madrid | 54 | 78 | – | – | – | – |
| Cambridge | – | – | 3 | 3 | 9 | 6 |
| Helsinki | – | – | 6 | 4 | 1 | 2 |
| Obu | – | – | 3 | 3 | 3 | 3 |
| Pittsburgh | – | – | 3 | 3 | 3 | 3 |
| Total | 54 | 78 | 15 | 13 | 13 | 11 |

MCI: Mild Cognitive Impairment; NC: elderly control subjects.
FieldTrip functions for automatic artifact detection, using the recommended parameter values in the tutorial; all epochs that contained artifacts were discarded from the analysis. All subjects had a minimum of 40 clean epochs, and the number of clean epochs did not differ between controls (M ± SD = 79 ± 40) and MCI (M ± SD = 72 ± 31) cases (p = 0.27). The extent of head movement during these epochs did not differ between controls (M ± SD = 13 ± 0.22) mm/s and MCI cases (M ± SD = 17 ± 0.22) mm/s (p = 0.76). Synchronization metrics were restricted to the 102 magnetometers, because this approach facilitates the comparability of the results with other MEG systems (planar gradiometers are unique to Elekta systems).

2.3.3. MEG synchronization metrics

Mutual Information (MI) captures both linear and nonlinear dependencies (Hlaváčková-Schindler, 2007), and was used to estimate functional connectivity between all pairs of magnetometers. First, for every epoch and magnetometer, a histogram (with 15 bins) was computed from the corresponding time series. Then, the probability of each pair of sensors where epoch and magnetometer, a histogram (with 15 bins) was computed for each pair of sensors, and the joint probability was calculated, leading to 20,604 (i.e., 5151 × 4) features per subject and frequency band.

The MI between each pair of sensors was calculated as:

$$MI_{xy} = \sum_{i=1}^{N} \sum_{j=1}^{N} p(x \in i, y \in j) \log \left( \frac{p(x \in i, y \in j)}{p(x \in i)p(y \in j)} \right)$$

and was normalized with:

$$MI_{xy, norm} = \frac{MI_{xy}}{\sqrt{H_{xx}H_{yy}}}$$

where

$$H_{xx} = -\sum_{i=1}^{N} p(x \in i) \log(p(x \in i))$$

and

$$H_{yy} = -\sum_{j=1}^{N} p(y \in j) \log(p(y \in j))$$

are the entropies of x and y.

This process yielded 102 × 102 symmetrical matrices that contained coupling for each subject, frequency band, and 2-s epoch. For each of the 5151 sensor pairs (i.e., (102 × 101)/2), four summary values of interquartile range, maximum, minimum, and median across epochs were calculated, leading to 20,604 (i.e., 5151 × 4) features per subject and frequency band.

2.4. Data partitioning

We used an approach inspired by the Clinical Data Partitioning (CliDaPa) algorithm (González et al., 2010) which has shown robustness and accuracy in the classification of complex datasets coming from multiple data sources. CliDaPa effectively exploits the availability of having clinical data combined to any massive data records (e.g., links), as shown in Table S4. Moreover, using bootstrapping as an internal validation method was also motivated by the fact that it provides less variance than other cross-validation methods.

Although the accuracy was computed from the bootstrap validation, once the algorithm stops the final model is computed. CliDaPa then provides the partitioning criterion and which machine-learning algorithm to use on each partition. Then these machine-learning algorithms are re-run on the entire partition data (no bootstrap) to get the final model (including the partitioning tree and the corresponding model per partition).

2.5. Feature selection, analysis process and internal validation

The aforementioned combined process was performed on the links (records of the synchronization pairs of channels) of the subjects included in each partition. The CliDaPa process involves a pipeline consisting of: (1) a feature selection filtering (performing a Chi-Square ($\chi^2$) filter with threshold 0.0) and (2) the evaluation of all the following machine-learning classifiers: Random Forest (Breiman, 2001), Bayesian Network (Buntine, 1991), C4.5 induction tree (Quinlan, 1993), K-nearest Neighbor (Cover and Hart, 1967), Logistic Regression (Ng and Jordan, 2002) and Support Vector Machine (Platt, 1999).

The parameters used for each machine-learning algorithm is included in the Supplementary material (Table S2).

CliDaPa evaluated the feature selection with each of the classifiers and selected the one with the best accuracy. In order to avoid overfitting, accuracy is calculated using a bootstrap validation (0.632 bootstrap variant as described by (Efron and Tibshirani, 1993) with 100 random sampling folds with replacement). This bootstrap validation process is referred to as CliDaPa’s internal validation (see Fig. 1).

We selected bootstrapping as the validation method based on the findings that it is the best validation method in datasets with more attributes than instances (e.g., links), as occurs in gene expression analysis (Braga-Neto and Dougherty, 2004). Moreover, using bootstrapping as an internal validation method was also motivated by the fact that it provides less variance than other cross-validation methods.

Although the accuracy was computed from the bootstrap validation, once the algorithm stops the final model is computed. CliDaPa then provides the partitioning criterion and which machine-learning algorithm to use on each partition. Then these machine-learning algorithms are re-run on the entire partition data (no bootstrap) to get the final model (including the partitioning tree and the corresponding model per partition).

2.6. External validation

The outcome of the CliDaPa algorithm was a partitioning tree (a series of conditions based on the demographic attributes to partition the subjects), plus a selected set of features and a particular machine-learning algorithm. The validation of the overall analysis algorithm was performed by using a separate set of unseen data (both demographic and link data), as shown in Fig. 2.

When training with Dataset 1 (single-site model), external validation was performed with Datasets 2 and 3. When combining Datasets 1 and 2 for training (multi-site model), external validation was performed with Dataset 3.

3. Results

Table 2 shows the results of the segmentation process of the CliDaPa algorithm when trained with Dataset 1 (single site data), and the subsequent bootstrap validation. We initially used Alpha, Beta and Theta bands as well as the broadband data. For the internal validation (training with Dataset 1 and validating with the same Dataset 1 using bootstrap), all of these bands produced classification results higher than 75%, so we decided to proceed with all of them for the external validation. When using Dataset 2 to validate the model obtained with Dataset 1, we observed that broadband was the most accurate for classification, which is why we used the broadband for the final step of using data from Datasets 1 and 2 to classify new subjects from Dataset 3. Nonetheless, for completeness, Table S4 shows the results from this final blind validation with Dataset 3 for each band. Broadband was clearly the most robust for classification, whereas Beta and Theta were prone to overfitting.
The results of the internal cross-classification analysis revealed an overall accuracy of 79% (see Table 2). The sensitivity was 83% and the specificity 72%. The links between sensor pairs that best classified the groups were bilateral antero-posterior and mid-anterior interhemispheric connections. Fig. 3 presents all those links that appear in all of the 100 iterations of the bootstrap validation. These are the intersection of all the sets of pairs of synchronized channels (the common core dataset) selected by the feature selection algorithm along all these iterations. The columns shown in Fig. 4 represent the links in the connectivity map representation in Fig. 3.

Fig. 4 shows a HeatMap representation of the average channel synchronization values for each of the four groups (True Positive: MCI correctly classified, True Negative: Control subjects correctly classified, False Positive: type I error or control subjects misclassified as MCI, and False Negative: type II error or MCI cases misclassified as controls) for all subjects in Datasets 1 and 2. Columns represent links identified as representative biomarkers (those that appear in all and every 100 iterations of the bootstrap validation), and color indicates the z-score value of the average synchronization for that group considering the mean and standard deviation of the entire dataset. This color is red when the average value of that group for the given channel is higher than the mean value (hypersynchronization), and blue when this average value is lower than the mean of the dataset (hyposynchronization). This representation shows that the model selects specific links to compute the synchronization differences to perform the prediction. The errors occur in the cases when the synchronization pattern is high for some pair of channels but low in others, although these low values are particularly low (false positive errors). As shown in Fig. 4 most of the classifying links presented higher synchronization in MCI patients.

The model developed with the single-site Dataset 1 was then tested with Dataset 2 (external validation), resulting in an overall accuracy of 82%, with a sensitivity of 92% and a specificity of 73% (see Table 2).

To confirm these results, members of the consortium sent a third group of data (Dataset 3) and the model was once again tested (see Table 3, left columns). A total accuracy of 79%, with a sensitivity of 91% and specificity of 69%, was achieved. Finally, to evaluate whether a model trained on data from multiple sites generalized better, we reran the classification algorithms after combining Datasets 1 and 2, and tested it on Dataset 3. This new model yielded an accuracy of 83%, a sensitivity of 100% and a specificity of 69% (see Table 3, right columns).

4. Discussion

The era of biomarker analysis of the natural history of DAT has provided a wealth of new data and insights regarding the pathophysiology of the disease. In particular, in vivo amyloid imaging has shown that while Beta amyloid may be a necessary precondition for the development of DAT, it is not the only factor involved in the expression of the clinical syndrome, and, furthermore, amyloid deposition can occur without cognitive impairment (Jack et al., 2014). New biomarkers that...
sites suggests that the pro-
utility across different clinical and research environments. When our
dementia from MCI (Bajo et al., 2010, 2012). The fact that this
was described in previous studies, and is a predictor of incident
of the hyper-synchronization pro-
ually being controls. This could be due to the homogeneity of the MCI
trols, with approximately one third of individuals classi-
ed as MCI patients. This pro-
elated to the release of Beta amyloid protein
causing neuronal hyperexcitability (Cirrito et al., 2008). Indeed, it has
been described that the accumulation of neuritic plaques produces the
loss of inhibitory neurons, producing an excess of excitability (Garcia-
Marin et al., 2009). Therefore, as this neuronal hyperexcitability could
cause spurious synchronization and even hypersynchronization, we
speculate that our findings are a consequence of increasing interstitial
amyloid. Obviously, both amyloid imaging and MEG imaging are re-
quired to test this hypothesis, since other alternative hypothesis such as
compensation could be plausible to explain this phenomenon. Never-
theless, our data add to the evidence that hypersynchronization is a re-
liable feature of MCI. The breakdown of the normal neuronal networks
in an individual patient can be used to identify the presence of neuro-
pathological changes that may lead to Alzheimer’s dementia.

Both anterior/posterior and interhemispheric connections were found relevant: these links allowed the most accurate classification.
This network profile was described previously (Bajo et al., 2010) and at least the antero-posterior hyperconnectivity seems to predict who
will develop dementia (Bajo et al., 2012). However, no conclusions
about specific brain regions can be drawn from the present sensor-
space analysis. In fact, functional connectivity estimates with MI in
do not rely on a specific underlying neuropathology, but only on the
consequences of that pathology – in this case synaptic dysfunction –
may prove to be useful, especially in the earliest stages of the disease. Our report, and the other studies demonstrating sensitivity of MEG im-
aging to DAT and MCI (Zamrini et al., 2011), suggest that analysis of
neuronal population interactions using this technology may be a fruitful
avenue for this type of research.

We conclude that it is possible to construct a machine learning
model using functional connectivity data derived from MEG imaging to
discriminate between MCI patients and cognitively normal elderly
subjects at the individual subject level. The model has a high sensitivity
at the cost of a somewhat lower specificity. The model is very good at
identifying individuals with MCI but is less accurate at identifying con-
trols, with approximately one third of individuals classified as “MCI” ac-
tually being controls. This could be due to the homogeneity of the MCI
group and the variability of the control group. In fact, the control
group could have included some subjects that will go on to develop
MCI later.

One important finding of this study was the cross-center stability of
the hyper-synchronization profile of the MCI patients. This profile
was described in previous studies, and is a predictor of incident
dementia from MCI (Bajo et al., 2010, 2012). The fact that this
hypersynchronization was observed in MCI patients across all five
sites suggests that the profile has some degree of reliability, and hence
utility across different clinical and research environments. When our
data are viewed in the context of other MEG studies related to MCI
and dementia (Zamrini et al., 2011), we can see a nonlinear change in
the neural networks as individuals proceed from normal cognition,
through one of subjective memory complaints, to MCI, and ultimately
to dementia. As individuals develop MCI, the functional network be-
comes less organized and hyper-synchronized (Bajo et al., 2010; Buldú et al., 2011), whereas patients with dementia have significantly
less synchronization, a lack of network organization, and significant
cognitive impairment (Stam et al., 2009). The hypersynchronization might be viewed as a compensatory response; if so, the compensation
is not adaptive, because the patient’s network organization is closer to
random (Buldú et al., 2011). This random network structure is less effi-
cient in information-processing terms and may arise from synaptic loss,
particularly if that loss occurs at hubs in the normal (non-random) net-
work. Thus, a more random network structure following synaptic loss in
DAT would lead to a reduced efficiency of normal information flow,
which is consistent with the idea of a disconnection syndrome. Thus,
the network becomes more highly synchronized as the clinical syn-
rome progresses until at some point the loss of organization in the net-
work, and perhaps disruption of the postsynaptic potentials, results in a
significant decrease in synchronization leading to dementia (Stam et al.,
2009).

This hypersynchronization profile may be related to the underlying
neuropathology, and specifically to the release of Beta amyloid protein
causing neuronal hyperexcitability (Cirrito et al., 2008). Indeed, it has
been described that the accumulation of neuritic plaques produces the
loss of inhibitory neurons, producing an excess of excitability (Garcia-
Marin et al., 2009). Therefore, as this neuronal hyperexcitability could
cause spurious synchronization and even hypersynchronization, we
speculate that our findings are a consequence of increasing interstitial
amyloid. Obviously, both amyloid imaging and MEG imaging are re-
quired to test this hypothesis, since other alternative hypothesis such as
compensation could be plausible to explain this phenomenon. Never-
theless, our data add to the evidence that hypersynchronization is a re-
liable feature of MCI. The breakdown of the normal neuronal networks
in an individual patient can be used to identify the presence of neuro-
pathological changes that may lead to Alzheimer’s dementia.

Table 2
Results of the validation of the first model. Dataset 1 is used to build a model that is vali-
dated with Dataset 1 (bootstrap validation, two columns of the left-hand side), and with
Dataset 2 (hold-out or external validation, two columns on the right side).

| Data to construct model | Dataset 1 (CiliDaPa) | Dataset 2 (hold-out) |
|-------------------------|---------------------|---------------------|
| Data to test model      | Dataset 1 (bootstrap) | Dataset 2 (hold-out) |

| Predicted class | Real class | Real class |
|----------------|-----------|-----------|
| MCI            | Sensitivity | Specificity |
| Normal         | 65        | 15        |
|                | 13        | 39        |
| Accuracy       | 0.83      | 0.72      |


Fig. 3. Graphical representation of the synchronization links selected as classifier features in model 1. Interhemispheric and antero-posterior links are shown in green and yellow, respectively.
sensor-space could be affected by volume conduction, but classification rates should not change as both groups could be equally affected by this effect. While analysis in source space might shed light on these brain regions – for example the critical features could be compared with resting state fMRI networks that have been found to be altered in MCI or DAT (Agosta et al., 2012; Binnewijzend et al., 2012; Brier et al., 2012) – we do not think source reconstruction would add extra information that would improve classification performance, which was the main aim of this study. Our sensor level approach is simpler, without making the extra assumptions required for source reconstruction, and therefore easier to transfer to clinical environments.

There are several aspects of our work that need to be addressed in future studies. Neuropsychological tests used for the diagnosis of MCI were not always the same in every site, due to cultural and regional reasons. This could be a source of noise in our sample selection and a limitation of this multisite study. However, in order to overcome this limitation, all laboratories used the same clinical criteria for MCI diagnosis. In addition, we had a relatively small set of test samples, and we did not have in vivo amyloid and/or Tau imaging. Of these limitations, the latter is perhaps the most important because we cannot confirm that the diagnosed MCI patients had amyloid and/or Tau deposition, and therefore that “MCI was due to AD” (with intermediate or high likelihood) (Albert et al., 2011). However, given the increasing evidence that many older people have amyloid accumulation but no MCI (Jack et al., 2014) and the fact that progression to DAT is not always preceded by measurable amyloid, this only reinforces the need for a larger, multimodal imaging study to fully evaluate the relationships among synaptic disruption, amyloid and Tau deposition, and incident DAT (Pievani et al., 2011). According to this, it should be noticed that the intrinsic heterogeneity that characterizes the MCI population may affect the results. Dividing MCI patients into subgroups with more homogeneous clinical information could generate subgroup-specific models and classifiers with greater accuracy. Furthermore, we did not take into account potential differences due to ethnicity, although the present classifiers were equally accurate across MEG centers. In addition, significant “site” effects remain in our data, despite our attempts to minimize between-site differences in biomagnetic noise or recording conditions. This is why we developed a training model that included data from all of the five sites, and this resulted in slightly better classification accuracy (i.e., to 83% from 79%).

Table 3

| Data to construct model | Dataset 1 (CliDaPa) | Datasets 1 and 2 (CliDaPa) |
|------------------------|--------------------|--------------------------|
| Data to test model     | Dataset 3 (hold-out) |                          |
| Predicted class        | Real class | Real class |
| MCI                    | MCI | Normal | MCI | Normal |
| Normal                 | Accuracy | Sensitivity | Specificity | Sensitivity | Specificity |
|                         | .79 | .91 | .69 | .90 | .69 |

5. Conclusion

Here, we report for the first time, a blind study using machine learning discrimination methods applied to data acquired from multiple sites that can accurately identify patients with MCI based on the pattern of functional connectivity measured with MEG. Because MEG relies on measures of neuronal function through the magnetic fields produced by postsynaptic currents, this noninvasive technique may provide a “window” into CNS dysfunction early in the course of the neurological process. To the extent that DAT may, at least in its earliest stages, reflect synaptic loss and dysfunction, then it would appear that MEG functional connectivity may be an ideal candidate biomarker for early, presymptomatic detection of the neuropathology of DAT, and for identifying MCI-patients at high risk of having DAT.

Conflicts of interest

The authors have no conflicts of interest to declare. Travels for the MAGIC-AD meetings were funded by Elekta company. The funding sources had no role in the study design, data collection, analysis, or the interpretation of the data.
Acknowledgments

This research was supported in part by funds from the National Institutes of Health (P01AG05133, R21MH098745), the Spanish Ministry of Innovation and Science (PSI2009-14415-C03-01 and PSI2012-38375-C03-01), SaWe Research Program for Mind and Body (grant number: 1104/10), the UK Medical Research Council (MC_A060_5PR10), and JSPS KAKENHI grant number 24590908. Thanks to Lisa Brindley and Elisa Cooper for helping to collect the Cambridge MEG data, and Peter Nestor for referring the Cambridge patients. We are also grateful to Melissa Fabrizio and Wiltrud Fassbinder for collecting the Pittsburgh MEG data.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at http://dx.doi.org/10.1016/j.clinph.2015.07.011.

References

Agosta, F., Pievani, M., Geroldi, C., Copetti, M., Frisoni, G.B., Filippi, M., 2012. Resting state fMRI in Alzheimer’s disease: beyond the default mode network. Neurobiol. Aging 33 (8), 1564–1578. http://dx.doi.org/10.1016/j.neurobiolaging.2011.06.021.18210.1002/jnr.21005.
Albert, M.S., DeKosky, S.T., Dickson, D., Dubois, B., Feldman, H.H., Fox, N.C., Gamst, A., Holzman, DM., Jagust, W.J., Petersen, R.C., et al., 2011. The diagnosis of mild cognitive impairment 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. 7 (3), 270–279. http://dx.doi.org/10.1016/j.jalz.2011.03.014.215139.
Bajo, R., Castellanos, N.P., Cuesta, P., Aurrtenetxe, S., Garcia-Prieto, J., Gil-Gregorio, P., del Pozo, F., Maestu, F., 2012. Differential patterns of connectivity in progressive mild cognitive impairment. Brain Connect. 2 (1), 21–24. http://dx.doi.org/10.1089/brain.2011.0292.185498.
Bajo, R., Maestu, F., Nieves, A., Sancho, M., Gutierrez, R., Campo, P., Castellanos, N.P., Gil, P., Morait, S., Pereda, E., et al., 2010. Functional connectivity in mild cognitive impairment during a memory task: implications for the disconnection hypothesis. J. Alzheimers Dis. 22 (1), 183–193. http://dx.doi.org/10.3233/JAD-2010-10017720847450.
Blinn-Wierend, M.A., Schoonheim, M.M., Sanz-Arigita, E., Wink, A.M., van der Flier, W.M., Delbeuck, X., Van der Linden, M., Collette, F., 2003. Alzheimer’s disease. Alzheimers Dement. 7 (3), 270–279. http://dx.doi.org/10.1016/j.jalz.2011.03.014.215139.
Braga-Neto, U.M., Dougherty, E.R., 2004. Is cross-validation valid for small-sample microarray classification? Bioinformatics. 20 (3), 374–380. http://dx.doi.org/10.1093/bioinformatics/bth040.
Cirrito, J.R., Kang, J.-E., Lee, J., Stewart, F.R., Verge, D.K., Silverio, L.M., Bu, G., Memerick, S., Holtzman, D.M., 2008. Endocytosis is required for synaptic activity-dependent release of amyloid-beta in vivo. Neuron 58 (1), 42–51. http://dx.doi.org/10.1016/j.neuron.2008.02.001.18400162.
Cover, T.M., Hart, P.E., 1967. Nearest neighbor pattern classification. IEEE Trans. Inform. Theory. 13 (1), 21–27.
Delbeuck, X., Van der Linden, M., Collette, F., 2003. Alzheimer’s disease as a disconnection syndrome? Neuropsychol. Rev. 13 (2), 79–92. http://dx.doi.org/10.1007/A10238321/057012887040.
Efron, B., Tibshirani, R., 1993. An introduction to the bootstrap. Chapman and Hall/CRC press.