NeuroImage: Clinical 9 (2015) 103–109

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

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

Fernando Maestú PhD[a,*], Jose-Maria Peña PhD[a], Pilar Garcés MS[a], Santiago González PhD[a], Ricardo Bajo PhD[a], Anto Bagic MD, PhD[b], Pablo Cuesta MS[a], Michael Funke MD[c], Jyrki P. Mäkelä MD[d], Ernestina Menasalvas PhD[a], Akinori Nakamura MD[e], Lauri Parkkonen PhD[f,g], Maria E. López PhD[a], Francisco del Pozo PhD[a], Gustavo Sudre PhD[h], Edward Zamrini MD[i], Eero Pekkonen MD[j], Richard N. Henson PhD[k], James T. Becker PhD[b,l,m], Magnetoencephalography International Consortium of Alzheimer's Disease[1]

[a] Laboratory of Cognitive and Computational Neuroscience, Center for Biomedical Technology, Complutense University of Madrid and Technical University of Madrid, Madrid, Spain
[b] Department of Neurology, University of Pittsburgh, Pittsburgh, USA
[c] Department of Pediatrics, University of Texas Health Science Center, Houston, USA
[d] BioMag Laboratory, HUS Medical Imaging Center, Helsinki University Central Hospital, Hensinki, Finland
[e] Department of Clinical and Experimental Neuroimaging, National Center for Geriatrics and Gerontology, Obu, Japan
[f] Department of Biomedical Engineering and Computational Science, Aalto University School of Science, Aalto, Espoo, Finland
[g] Elekta Oy, Helsinki, Finland
[h] Human Genome Research Institute, National Institutes of Health, Bethesda, USA
[i] Department of Neurology, University of Utah, Salt Lake City, USA
[j] Department of Neurology, University of Helsinki, Finland
[k] Medical Research Council Cognition and Brain Sciences Unit, Cambridge, UK
[l] Department of Psychiatry, University of Pittsburgh, Pittsburgh, USA
[m] Department of Psychology, University of Pittsburgh, Pittsburgh, USA

ARTICLE INFO

Available online 1 August 2015

Keywords:
Magnetoencephalography
Mild Cognitive Impairment
Functional connectivity
Data mining
Machine learning
Synaptic dysfunction
Multicenter study

ABSTRACT

Synaptic disruption is an early pathological sign of the neurodegeneration of Dementia of the Alzheimer's type (DAT). The changes in network synchronization are evident in patients with Mild Cognitive Impairment (MCI) at the group level, but there are very few Magnetoencephalography (MEG) studies regarding discrimination at the individual level. In an international multicenter study, we used MEG and functional connectivity metrics to discriminate MCI from normal aging at the individual person level. A labeled sample of features (links) that distinguished MCI patients from controls in a training dataset was used to classify MCI subjects in two testing datasets from four other MEG centers. We identified a pattern of neuronal hypersynchronization in MCI, in which the features that best discriminated MCI were fronto-parietal and interhemispheric links. The hypersynchronization pattern found in the MCI patients was stable across the five different centers, and may be considered an early sign of synaptic disruption and a possible preclinical biomarker for MCI/DAT.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

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 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;

* Corresponding author at: Laboratory of Cognitive and Computational Neuroscience, Center for Biomedical Technology, Campus Montegancedo, Pozuelo de Alarcón, Madrid 28223, Spain. Tel.: +34 91 336 4641/+34 606561695.

[1] The Magnetoencephalography International Consortium of Alzheimer's Disease investigators and collaborators include: University of Utah — Edward Zamrini; Center for Biomedical Technology (UCM-UPM) — Fernando Maestu, Ricardo Bajo, Francisco del Pozo, Pablo Cuesta, Pilar Garcés, Maria Eugenia López, Santiago Gonzalez, Jose María Peña and Ernestina Menansalvas; Helsinki University Central Hospital and University of Helsinki — Jyrki Mäkelä and Eero Pekkonen; University of Pittsburgh — Anto Bagic, James T. Becker and Gustavo Sudre; University of Texas Health Science Center at Houston — Michael Funke; Aalto University — Lauri Parkonnen; National Center for Geriatrics and Gerontology, Obu, Japan — Akinori Nakamura and University of Cambridge — Richard Henson.

http://dx.doi.org/10.1016/j.nicl.2015.07.011
2213-1582/© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

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 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 data from five different 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.

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	0	0
Total	54	78	15	13	13	11

MCI: Mild Cognitive Impairment; NC: elderly control subjects.

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; Nenonen 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

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 = 0.13 ± 0.22) mm/s and MCIs (M ± SD = 0.17 ± 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 $N = 15$ bins) were computed from the corresponding times series $x(t)$, and the probability $p(x \in i)$, $i = 1, 2, \ldots, N$ of each bin i was evaluated. Additionally, for each pair of sensors x and y, the joint probability $p(x \in i, y \in j)$, $i = 1, 2, \ldots N, j = 1, 2, \ldots N$ was computed, which evaluates the probability that $x(t)$ belongs to bin i and $y(t)$ belongs to bin j. The MI between $x(t)$ and $y(t)$ was then 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 $H_{yy} = -\sum_{i=1}^{N} p(y \in i)\log(p(y \in i))$ 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., $(102 \times 101)/2$), 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 (for instance, gene expression). In this particular case the clinical information used was demographic data from the subjects (age and sex) and the massive data records are the aforementioned links (MEG functional pair-channel synchronization). The algorithm evaluated different data partitioning criteria in an iterative greedy manner using the demographic information. To evaluate each tentative partition criterion, the algorithm split the dataset according to a selected demographic attribute in two or more partitions. Then each partition was used as separate training datasets using the corresponding links from the instances in each partition.

CliDaPa applied a combined process made of a specific feature selection and a set of machine-learning classifiers (selecting the classifier with the best performance for the data in that partition). Then the accuracy achieved for a particular partition criterion was obtained as the weight average of the best classifier for each partition. This partitioning process continued whereas there was any improvement in the averaged accuracy of the whole partitioning scheme and there were more demographic attributes to consider for the data partitioning. The partition process might also stop if one of the partitions included less than a minimum number of instances (20 instances in this study).

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 (χ^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.

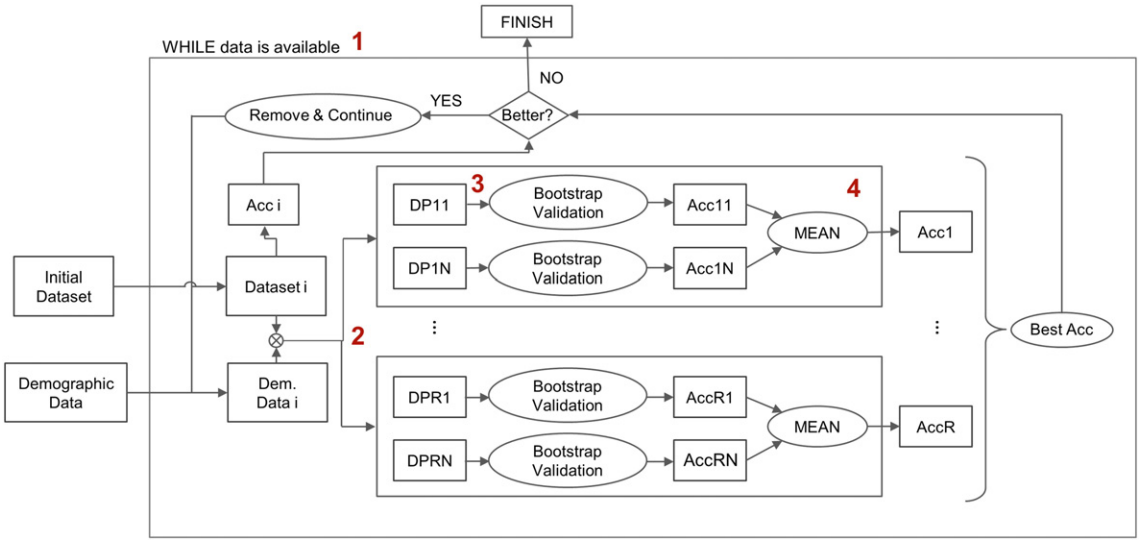

Fig. 1. General structure of the CliDaPa algorithm: (1) the iterative greedy partitioning loop, (2) where the R partition criteria are generated, and (3) each partition criterion includes N partitions of the massive data records which is classified and (4) validated by means of a bootstrap method.

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 anterio-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

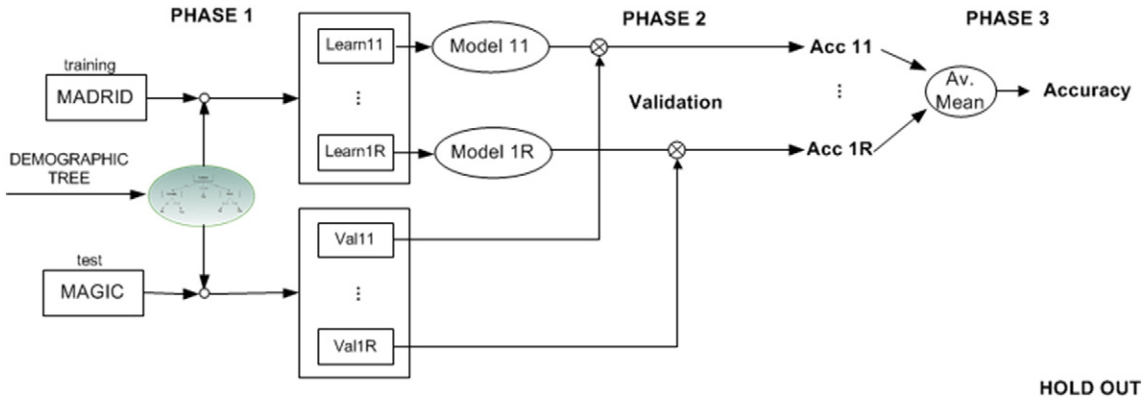

Fig. 2. The external validation is performed for each of the partitions defined by the obtained tree (using demographic attributes), from the validation data it is divided into different groups according to the application of the same partitioning criteria, and for each of the partitions the classification model is applied. The results are the average of the classification results.

Table 2
Results of the validation of the first model. Dataset 1 is used to build a model that is validated 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).

Results of first validation				
Data to construct model	Dataset 1 (CliDaPa)			
Data to test model	Dataset 1 (bootstrap)		Dataset 2 (hold-out)	
	Real class		Real class	
Predicted class	MCI	Normal	MCI	Normal
MCI	65	15	12	4
Normal	13	39	1	11
	Sensitivity	Specificity	Sensitivity	Specificity
	.83	.72	.92	.73
Accuracy	.79		.82	

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 imaging 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 controls, with approximately one third of individuals classified as "MCI" actually 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 becomes 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 efficient in information-processing terms and may arise from synaptic loss, particularly if that loss occurs at hubs in the normal (non-random) network. 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 syndrome progresses until at some point the loss of organization in the network, 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 required to test this hypothesis, since other alternative hypothesis such as compensation could be plausible to explain this phenomenon. Nevertheless, our data add to the evidence that hypersynchronization is a reliable feature of MCI. The breakdown of the normal neuronal networks in an individual patient can be used to identify the presence of neuropathological 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

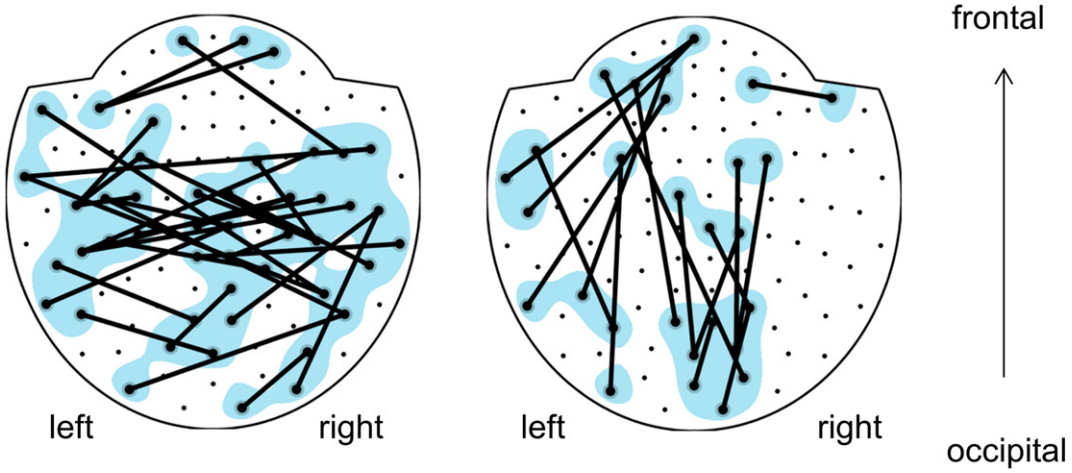

Fig. 3. Graphical representation of the synchronization links selected as classifier features in model 1. Interhemispheric and antero-posterior links are shown in green and yellow, respectively.

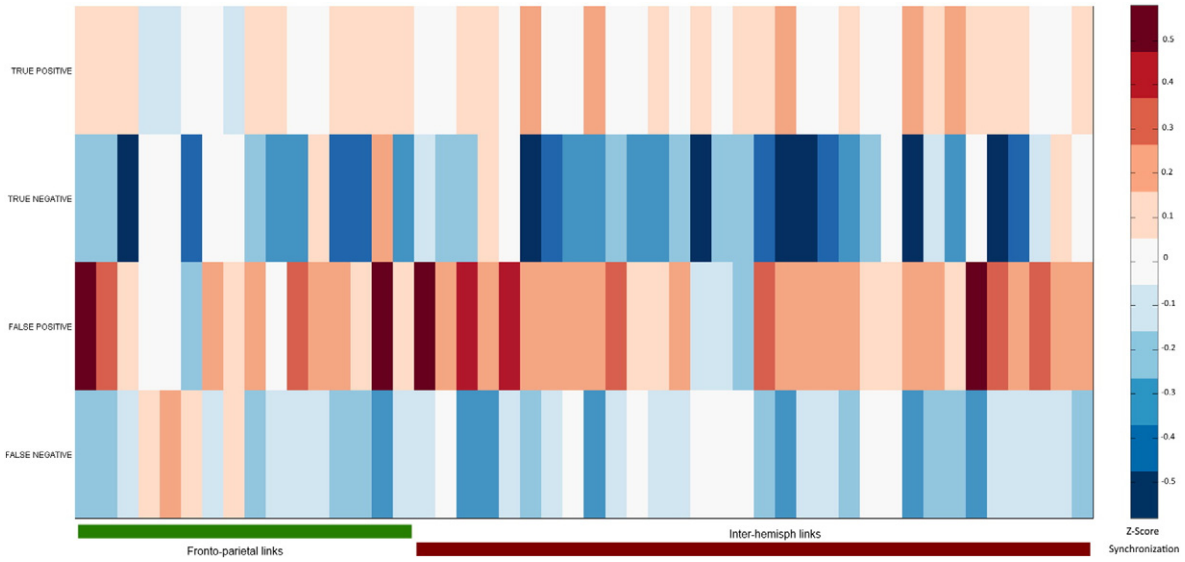

Fig. 4. HeatMap representation showing the z-score value of the average synchronization links, divided according to True Positive, True Negative, False Positive and False Negative cases from the classification using the subjects from both Datasets 1 and 2.

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, including episodic memory decline. Besides, 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 (Villemagne et al., 2013) 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 homogenous 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
Results of the external validation of the first and the second model. Dataset 1 was used to build the first model (two columns on the left-side). Datasets 1 and 2 were used to build the second model (two columns on the right-hand side). Both models were validated with Dataset 3.

Results of the second validation				
Data to construct model	Dataset 1 (CliDaPa)		Datasets 1 and 2 (CliDaPa)	
Data to test model	Dataset 3 (hold-out)			
	Real class		Real class	
Predicted class	MCI	Normal	MCI	Normal
MCI	10	4	11	4
Normal	1	9	0	9
	Sensitivity	Specificity	Sensitivity	Specificity
	.91	.69	1.00	.69
Accuracy	.79		.83	

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 neuropathological 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), SalWe 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.nicl.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.2011.06.00721813210.

Albert, M.S., DeKosky, S.T., Dickson, D., Dubois, B., Feldman, H.H., Fox, N.C., Gamst, A., Holtzman, D.M., 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.2011.03.00821514249.

Bajo, R., Castellanos, N.P., Cuesta, P., Aurtenetxe, 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.2011.006922458376.

Bajo, R., Maestú, F., Nevado, A., Sancho, M., Gutiérrez, R., Campo, P., Castellanos, N.P., Gil, P., Moratti, 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.

Binnewijzend, M.A., Schoonheim, M.M., Sanz-Arigita, E., Wink, A.M., van der Flier, W.M., Tolboom, N., Adriaanse, S.M., Damoiseaux, J.S., Scheltens, P., van Berckel, B.N., Barkhof, F., 2012. Resting-state fMRI changes in Alzheimer's disease and mild cognitive impairment. Neurobiol. Aging 33 (9), 2018–2028. http://dx.doi.org/10.1016/j.neurobiolaging.2011.07.00321862179.

Braga-Neto, U.M., Dougherty, E.R., 2004. Is cross-validation valid for small-sample microarray classification? Bioinformatics. 20 (3), 374–380.

Breiman, L., 2001. Random Forests. Mach. Learn. 45 (1), 5–32.

Brier, M.R., Thomas, J.B., Snyder, A.Z., Benzinger, T.L., Zhang, D., Raichle, M.E., Holtzman, D.M., Morris, J.C., Ances, B.M., 2012. Loss of intranetwork and internetwork resting state functional connections with Alzheimer's disease progression. J. Neurosci. 32 (26), 8890–8899. http://dx.doi.org/10.1523/JNEUROSCI.5698-11.201222745490.

Buldú, J.M., Bajo, R., Maestú, F., Castellanos, N., Leyva, I., Gil, P., Sendiña-Nadal, I., Almendral, J.A., Nevado, A., del-Pozo, F., et al., 2011. Reorganization of functional networks in mild cognitive impairment. PLOS One 6 (5), e19584. http://dx.doi.org/10.1371/journal.pone.001958421625430.

Buntine, W., 1991. Theory refinement on Bayesian networks. In Proceedings of the Seventh Conference on Uncertainty in Artificial Intelligence. Morgan Kaufmann Publishers Inc, pp. 52–60.

Cirrito, J.R., Kang, J.-E., Lee, J., Stewart, F.R., Verges, D.K., Silverio, L.M., Bu, G., Mennerick, 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.2008.02.00318400162.

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/10.1023/A:1023832305702128870 40.

Efron, B., Tibshirani, R., 1993. An introduction to the bootstrap. Champman and Hall/CRC press.

Garcia-Marin, V., Blazquez-Llorca, L., Rodriguez, J.-R., Boluda, S., Muntane, G., Ferrer, I., Defelipe, J., 2009. Diminished perisomatic GABAergic terminals on cortical neurons adjacent to amyloid plaques. Front. Neuroanat. 3, 28. http://dx.doi.org/10.3389/neuro.05.028.200919949482.

Gonzalez-Moreno, A., Aurtenetxe, S., Lopez-Garcia, M.E., del Pozo, F., Maestu, F., Nevado, A., 2014. Signal-to-noise ratio of the MEG signal after preprocessing. J. Neurosci. Methods 222, 56–61. http://dx.doi.org/10.1016/j.jneumeth.2013.10.01924200506.

González, S., Guerra, L., Robles, V., Peña, J.M., Famili, F., 2010. CliDaPa: A new approach to combining clinical data with DNA microarrays. Intell. Data Anal. 14 (2), 207–231.

Hämäläinen, M., 1993. Magnetoencephalography — theory, instrumentation, and applications to noninvasive studies of the working human brain. Rev. Mod. Phys. 65, 413–497.

Hlavackovaschindler, K., Palus, M., Vejmelka, M., Bhattacharya, J., 2007. Causality detection based on information-theoretic approaches in time series analysis. Phys. Rep. 441 (1), 1–46. http://dx.doi.org/10.1016/j.physrep.2006.12.004.

Jack Jr, C.R., Wiste, H.J., Weigand, S.D., Rocca, W.A., Knopman, D.S., Mielke, M.M., Lowe, V.J., Senjem, M.L., Gunter, J.L., Preboske, G.M., Pankratz, V.S., Vemuri, P., Petersen, R.C., 2014. Age-specific population frequencies of cerebral β-amyloidosis and neurodegeneration among people with normal cognitive function aged 50–89 years: a cross-sectional study. Lancet Neurol. 13 (10), 997–1005.

López, M.E., Bruña, R., Aurtenetxe, S., Pineda-Pardo, J.Á, Marcos, A., Arrazola, J., Reinoso, A.I., Montejo, P., Bajo, R., Maestú, F., 2014. Alpha-band hypersynchronization in progressive mild cognitive impairment: a magnetoencephalography study. J. Neurosci. 34 (44), 14551–14559. http://dx.doi.org/10.1523/JNEUROSCI.0964-14.201425355209.

Morrison, J.H., Hof, P.R., Bouras, C., 1991. An anatomic substrate for visual disconnection in Alzheimer's disease. Ann. N. Y. Acad. Sci. 640, 36–431776757.

Nenonen, J., Nurminen, J., Kičić, D., Bikmullina, R., Lioumis, P., Jousmäki, V., Taulu, S., Parkkonen, L., Putaala, M., Kähkönen, S., 2012. Validation of head movement correction and spatiotemporal signal space separation in magnetoencephalography. Clin. Neurophysiol. 123 (11), 2180–2191. http://dx.doi.org/10.1016/j.clinph.2012.03.08022633918.

Ng, A.Y., Jordan, M.I., 2002. On Discriminative vs. Generative Classifiers: A comparison of Logistic Regression and Naive Bayes. In: Mitchell, T.M. (Ed.), Neural Information Processing Systems. In Machine Learning. McGraw Hill.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011, 156869. http://dx.doi.org/10.1155/2011/15686921253357.

Petersen, R.C., 2004. Mild cognitive impairment as a diagnostic entity. J. Intern. Med. 256 (3), 183–194. http://dx.doi.org/10.1111/j.1365-2796.2004.01388.x15324362.

Pievani, M., de Haan, W., Wu, T., Seeley, W.W., Frisoni, G.B., 2011. Functional network disruption in the degenerative dementias. Lancet Neurol. 10 (9), 829–843. http://dx.doi.org/10.1016/S1474-4422(11)70158-221778116.

Platt, J.C., 1999. Fast training of support vector machines using sequential minimal. Advances in Kernel Methods - Support Vector Learning. MIT Press, Cambridge, pp. 185–208.

Qiu, C., Kivipelto, M., von Strauss, E., 2009. Epidemiology of Alzheimer's disease: occurrence, determinants, and strategies toward intervention. Dial. Clin. Neurosci. 11 (2), 111–12819585947.

Quinlan, J.R., 1993. Programs for Machine Learning. Morgan Kaufmann Publishers.

Selkoe, D.J., 2002. Alzheimer's disease is a synaptic failure. Science 298 (5594), 789–791. http://dx.doi.org/10.1126/science.107406912399581.

Stam, C.J., de Haan, W., Daffertshofer, A., Jones, B.F., Manshanden, I., van Cappellen van Walsum, A.M., Montez, T., Verbunt, J.P.a., de Munck, J.C., van Dijk, B.W., et al., 2009. Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease. Brain 132 (1), 213–224. http://dx.doi.org/10.1093/brain/awn26218952674.

Taniguchi, T., Kawamata, T., Mukai, H., Hasegawa, H., Isagawa, T., Yasuda, M., Hashimoto, T., Terashima, A., Nakai, M., Mori, H., et al., 2001. Phosphorylation of tau is regulated by PKN. J. Biol. Chem. 276 (13), 10025–10031. http://dx.doi.org/10.1074/jbc.M00742720011104762.

Taulu, S., Simola, J., 2006. Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. Phys. Med. Biol. 51 (7), 1759–1768. http://dx.doi.org/10.1088/0031-9155/51/7/00816552102.

Villemagne, V.L., Burnham, S., Bourgeat, P., Brown, B., Ellis, K.A., Salvado, O., Szoeke, C., Macaulay, S.L., Martins, R., Maruff, P., et al., 2013. Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. Lancet Neurol. 12 (4), 357–367. http://dx.doi.org/10.1016/S1474-4422(13)70044-923477989.

Zamrini, E., Maestu, F., Pekkonen, E., Funke, M., Makela, J., Riley, M., Bajo, R., Sudre, G., Fernandez, A., Castellanos, N., et al., 2011. Magnetoencephalography as a putative biomarker for Alzheimer's disease. Int. J. Alzheimers Dis. 2011, 280289. http://dx.doi.org/10.4061/2011/28028921547221.