Complexity analysis of the magnetoencephalogram background activity in Alzheimer's disease patients

Carlos Gómez[1], Roberto Hornero[1], Daniel Abásolo[1], Alberto Fernández[2], Miguel López[1]

[1]E.T.S. Ingenieros de Telecomunicación, Universidad de Valladolid, Spain

[2]Centro de Magnetoencefalografía Dr. Pérez-Modrego, Universidad Complutense de Madrid, Spain

AUTHOR'S ADDRESS: Carlos Gómez

E.T.S. Ingenieros de Telecomunicación

University of Valladolid

Camino del Cementerio s/n

47011 - Valladolid (Spain)

E-mail: carlos.gomez@tel.uva.es

Abstract

The aim of the present study was to analyse the magnetoencephalogram (MEG) background activity in patients with Alzheimer's disease (AD) using the Lempel-Ziv (LZ) complexity. This non-linear method measures the complexity of finite sequences and is related to the number of distinct substrings and the rate of their occurrence along the sequence. The MEGs were recorded with a 148-channel whole-head magnetometer (MAGNES 2500 WH, 4D Neuroimaging) in twenty one patients with AD and in twenty one age-matched control subjects. Artefact-free epochs were selected for complexity analysis. Results showed that MEG signals from AD patients had lower complexity than control subjects' MEGs and the differences were statistically significant ($p < 0.01$). In order to reduce the dimension of the LZ complexity results, a principal components analysis (PCA) was applied, and only the first principal component was retained. The first component score from PCA was graphically analysed using a box plot and a Receiver Operating Characteristic (ROC) curve. A specificity of 85.71%, a sensitivity of 80.95% and an area under the ROC curve of 0.9002 were obtained. These preliminary results suggest that cognitive dysfunction in AD is associated with a decreased LZ complexity in the MEG signals.

Keywords: Alzheimer's disease; magnetoencephalogram; Lempel-Ziv complexity; non-linear analysis; principal component analysis.

1. Introduction

Alzheimer's disease (AD) is considered to be the main cause of dementia in western countries [1]. This neurodegenerative disorder is characterized by neuronal loss and the appearance of neurofibrillary tangles and senile plaques [2, 3]. These plaques are spherical structures containing amyloid-β-peptide. As a definite diagnosis is only possible by brain autopsy following the death of the patient, the differential diagnosis with other types of dementia is becoming more important. The AD diagnosis includes Functional Assessment Staging (FAST) [4], computerized tomography and magnetic resonance imaging. Mini Mental State Examination (MMSE) of Folstein [5] is also used to assess the severity of cognitive deficit.

Magnetoencephalography is a non-invasive technique that allows recording the magnetic fields produced by brain activity. It provides high temporal and spatial resolution and it is independent of any reference point. In addition, magnetic fields are less distorted by the resistive properties of the skull [6].

There are several studies of the electroencephalogram (EEG) and the magnetoencephalogram (MEG) in AD patients. Spectral analysis seems to discriminate AD patients from control subjects through an increased EEG/MEG activity in delta and theta bands associated with AD [7-9]. On the other hand, some authors have studied this dementia using non-linear analysis. The most widely used non-linear methods to characterize the complexity of a system are the first positive Lyapunov exponent (L_1) and the correlation dimension (D_2). L_1 is a dynamic complexity measure that describes the divergence of trajectories starting at nearby initial states [10], while D_2 is a static measure of the system dimensional complexity [11, 12]. Jeong *et al.* [13] employed a method proposed by Kennel *et al.* [14] to calculate the minimum embedding dimension and demonstrated that AD patients exhibit significantly lower D_2 and L_1 values than controls in many EEG channels. Other

authors used the Grassberger and Procaccia algorithm with time-delay embedding [11, 12] and multichannel EEG to find that the dimensional complexity was significantly lower in AD patients than in age-matched control subjects [15]. Van Cappellen van Walsum *et al.* [16] used MEG data and computed the D_2 in different frequency bands. In the 0.5 – 40 Hz frequency band, the mean D_2 was lower in AD patients compared with control subjects. Woyshville and Calabrese [17] used another metric, the fractal dimension, to analyse the EEG background activity of control subjects, probable AD patients and autopsy-confirmed AD patients. These studies showed a decreased complexity of the EEG/MEG background activity in AD. Nevertheless, the classical measures for estimating the non-linear dynamic complexity have some drawbacks. Reliable estimation of L_1 and D_2 requires a large quantity of data [18] and stationary and noise free time series [13]. These assumptions cannot be achieved for physiological data. Since computational complexity of both methods is also high, specially if the amount of data is large, other measures are needed for EEG/MEG complexity analysis.

Several complexity measures have been proposed [19]. Firstly, dimensional complexity or just complexity, referring to D_2 [13, 15, 16, 20-23]. Secondly, the algorithmic or Kolmogorov complexity of a string [24], which is defined as the length of the string's shortest description in some fixed description language. Other measure is the neural complexity: statistical measures that capture regularities based on the deviation from independence among subsets of a system [25]. In our study we have used the Lempel-Ziv (*LZ*) complexity, an appropriate measure of complexity in Kolmogorov's sense. Moreover, it contains the notion of complexity in a statistical sense [26]. *LZ* complexity is a nonparametric measure of complexity for finite sequences related to the number of distinct substrings and the rate of their occurrence along the sequence [27] without the drawbacks of the classical non-linear complexity methods. *LZ* complexity was applied to EEGs in order to quantify the relationship between brain activity patterns and depth of anaesthesia [28, 29]. In [30],

complexity measures were extracted from mutual information time series of EEGs to predict response during isoflurane anaesthesia. Moreover, it has been used to predict movement during anaesthesia in dogs [31]. *LZ* complexity was also used to study the EEG signal of focal cerebral ischemia [32], the brain function [33] and the brain information transmission [345]. Other authors used this non-linear method to analyse neural discharges (spike trains) [35, 36], to detect ventricular tachycardia and ventricular fibrillation [37]. Finally, it has been used to quantify the regularity in uterine electromyography [38] and in epileptic seizure EEG time series data [39].

In this preliminary study we examined the MEG background activity in AD patients with the *LZ* complexity and principal components analysis (PCA), a statistical technique that allows to reduce the dimension of a data set. The purpose of this study was to test the hypothesis that the complexity measure applied to MEG data was lower in AD patients compared to control subjects.

2. Subjects and MEG recording

The MEG data were acquired from 42 subjects. Twenty one patients ranging in age from 56 to 83 years (73.62 ± 8.38 years, mean ± standard deviation SD) fulfilling the criteria of probable AD took part in this study. Patients were recruited from the Asociación de Familiares de Enfermos de Alzheimer (AFAL). The MMSE and FAST scores in AD patients were 18.00 ± 3.91 and 4.05 ± 0.22 (mean ± SD), respectively. One patient had a MMSE score of less than 12 points, indicating a severe degree of dementia. MEG measurements were performed at the moment of AD diagnoses and before initiation of treatment. None of the patients had any kind of medication that could influence on the MEG recording. The AD diagnosis for all patients was made using criteria of the National Institute of Communicative Disorders and Stroke - Alzheimer's Disease and Related Disorders Association (NINCDS-

ADRDA) [40].

MEGs were also obtained from twenty one elderly control subjects without past or present neurological disorders (age = 70.29 ± 7.07 years; MMSE score = 29.1 ± 1.0 points; FAST score = 1.71 ± 0.46; mean ± SD). Both groups were carefully matched for age (mean age of 73.62 for the patients and 70.29 for the controls). The difference in the mean age of both populations is not statistically significant ($p = 0.17 > 0.01$, Student's *t*-test). Table 1 summarizes the sociodemographic data of all subjects. The local ethics committee approved this study. All control subjects and all caregivers of the patients gave their informed consent for the participation in the current study.

INSERT TABLE 1 AROUND HERE

MEGs were recorded using a 148-channel whole-head magnetometer (MAGNES 2500 WH, 4D Neuroimaging) in a magnetically shielded room. The subjects lay comfortably on a patient bed with eyes closed. They were asked to stay awake and to avoid eye and head movements. For each subject, five minutes of recording were acquired at a sampling frequency of 678.17 Hz. These recordings were down-sampled to 169.549 Hz (50863 data points). Artefact-free epochs of 20 seconds (3392 samples) were selected from MEG signal. These epochs were digitally filtered using a band-pass filter with cut-off frequencies at 0.5 and 40 Hz. Finally, the pre-processed data were copied as ASCII files to a personal computer for off-line complexity analysis.

3. Methods

In Fig. 1 we present the four MEG analysis steps. The first step consisted in a pre-processing of the signals: segmentation, artefact-free epoch selection and band-pass filtering.

In the second step we estimated the *LZ* complexity of the signals. The third step was a PCA, used to reduce the dimension of the complexity results. Finally, a statistical analysis was carried out with Student's *t*-test, a box plot and a Receiver Operating Characteristic (ROC) curve.

INSERT FIGURE 1 AROUND HERE

3.1. Lempel-Ziv complexity

The *LZ* complexity algorithm for sequences of finite length was suggested by Lempel and Ziv in 1976 [27]. Later, Kaspar and Schuster [41] presented a computer program that determined the *LZ* complexity using only two simple operations: to copy and to insert. This program is an appropriate measure of complexity in Kolmogorov sense as well as in a statistical sense [26]. According to Kolmogorov [24], the complexity of a given string of zeros and ones is given by the number of bits of the shortest computer program which can generate it. The complexity in this sense seems very general and computer-dependent. Nevertheless, the *LZ* complexity avoids these disadvantages [26].

LZ complexity analysis is based on a coarse-graining of the measurements, so the MEG time series must be transformed into a finite symbol sequence. In this study we used the simplest way, a binary sequence conversion (zeros and ones). The median value is estimated as a threshold T_d, as partitioning about the median is robust to outliers [38]. By comparison with the threshold, the original data are converted into a 0-1 sequence $P = s(1), s(2),\ldots, s(n)$, with $s(i)$ defined by [28]:

$$s(i) = \begin{cases} 0 & if\ x(i) < T_d \\ 1 & if\ x(i) \geq T_d \end{cases} \quad (1)$$

The string P is scanned from left to right and a complexity counter $c(n)$ is increased by one unit every time a new subsequence of consecutive characters is encountered in the scanning process. The detailed algorithm for the measure of the *LZ* complexity is as follows [28, 29, 31]:

1. Let S and Q denote two subsequences of the original sequence P and SQ be the concatenation of S and Q, while $SQ\pi$ is a string derived from SQ after its last character is deleted (π means the operation to delete the last character).
2. Let $v(SQ\pi)$ denote the vocabulary of all different substrings of $SQ\pi$.
3. At the beginning, the complexity counter $c(n) = 1$, $S = s(1)$, $Q = s(2)$, $SQ = s(1)$, $s(2)$ and $SQ\pi = s(1)$.
4. In general, suppose that $S = s(1), s(2),\ldots, s(r)$, $Q = s(r+1)$ and, therefore, $SQ\pi = s(1), s(2),\ldots, s(r)$. If $Q \in v(SQ\pi)$, then Q is a subsequence of $SQ\pi$, not a new sequence.
5. S does not change and renew Q to be $s(r+1), s(r+2)$, then judge if Q belongs to $v(SQ\pi)$ or not.
6. The steps 4 and 5 are repeated until Q does not belong to $v(SQ\pi)$. Now $Q = s(r+1), s(r+2),\ldots, s(r+i)$ is not a subsequence of $SQ\pi = s(1), s(2),\ldots, s(r+i-1)$, so increase the counter by one.
7. Thereafter, S and Q are combined and renewed to be $s(1), s(2),\ldots, s(r+i)$, and $s(r+i+1)$, respectively.
8. Repeat the previous steps until Q is the last character. At this time, the number of different substrings is $c(n)$, the measure of complexity.

In order to obtain a complexity measure which is independent of the sequence length, $c(n)$ should be normalized. If the length of the sequence is n and the number of different symbols is α, it has been proved [27] that the upper bound of $c(n)$ is given by:

$$c(n) < \frac{n}{(1-\varepsilon_n)\log_\alpha(n)} \qquad (2)$$

where ε_n is a small quantity and $\varepsilon_n \to 0$ ($n \to \infty$). In general, $n/log_\alpha(n)$ is the upper limit of $c(n)$, where the base of the logarithm is α, i.e.,

$$\lim_{n \to \infty} c(n) = b(n) \equiv \frac{n}{\log_\alpha(n)} \qquad (3)$$

For a binary conversion $\alpha = 2$, and $b(n)$ is given by

$$b(n) \equiv \frac{n}{\log_2(n)} \qquad (4)$$

and $c(n)$ can be normalized via $b(n)$:

$$C(n) = \frac{c(n)}{b(n)} \qquad (5)$$

$C(n)$ is usually between zero and one. The normalized *LZ* complexity reflects the arising rate of new patterns along with the sequence [28]. To ensure that defect-related features will be included for the complexity calculation, a minimum data length needs to be considered [42]. Thus, the effect of the data length on the *LZ* complexity was analysed for the MEG epochs. These analyses showed that the complexity values decline quickly at the beginning and become stable from 3000 data points. Therefore, the normalized *LZ* complexity $C(n)$ can be viewed as independent of number of samples for $n > 3000$. As an example, Fig. 2 illustrates one curve of the complexity values against the data length for a MEG signal. Hence, an epoch length of 3392 (20 seconds of recording) was used in our study for the complexity measure.

INSERT FIGURE 2 AROUND HERE

3.2. Principal Components Analysis

The magnetometer used in this study recorded 148 MEG signals for each subject. Thus, 148 complexity results were obtained after applying *LZ* complexity to the recordings. In order to reduce the dimensionality of these results we used PCA. This multivariate statistical technique decomposes a set of data into mutually independent contributions, ranked in decreasing importance. The algorithm to calculate the principal components is the following:

1. Let $\boldsymbol{X}$ denote the data matrix of order $m \times n$, where m is the number of observations and n is the number of features. In our case $m = 42$ is the number of subjects and $n = 148$ is the number of complexity results.
2. The data are centred and scaled. For a given column, its mean is subtracted from all the observations in that column. Then, each element of the matrix is divided by $m - 1$ in order to scale the data. The new matrix $\boldsymbol{Z}$ has the same dimension as the original data matrix $\boldsymbol{X}$.
3. A singular value decomposition (SVD) of $\boldsymbol{Z}$ is carried out to perform the PCA:

$$\boldsymbol{Z} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T \qquad (6)$$

   where $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary matrices and their columns are called the left and the right singular vectors, respectively. $\boldsymbol{S}$ is a diagonal matrix of the same dimension as $\boldsymbol{Z}$ and with nonnegative elements on the diagonal in decreasing order, called singular values.
4. The principal components are the columns of $\boldsymbol{V}$ and the principal components scores are the projections of the original data onto the principal components axes.

There are several methods to select the number of principal components to be retained. The two most widely used are the Kaiser's criterion [43] and the Cattell's scree test [44]. The Kaiser's rule says that only the principal components with eigenvalues greater than one should be retained. The scree test is a graphical method in which the number of an eigenvalue is plotted versus its actual value, from highest to lowest. A change of slope in the graph

indicates the number of components to select.

3.3. Statistical analysis

LZ complexity analyses were carried out separately for each MEG channel. Firstly, a Student's *t*-test was used to determine if there were any differences between the *LZ* complexity values in both groups (control subjects and AD patients). Normality and equality of variances are required for applying this parametric statistical test. Lilliefors and Levene tests were used to verify these hypotheses. Secondly, the first component score from PCA was analysed using Student's *t*-test, box plot and ROC curve.

A box plot provides a visual summary of the data. This statistical graph shows a box with three horizontal lines at the median, lower quartile (25^{th} percentile) and upper quartile (75^{th} percentile) values. The ends of the whiskers (vertical lines) indicate the minimum and maximum values. Finally, a plus sign is the symbol used for the outlier values.

A ROC curve [45] summarizes the performance of a two-class classifier across the range of possible thresholds. It is a graphical representation of the tradeoffs between sensitivity and specificity. Sensitivity is the true positive rate and specificity is equal to the true negative rate:

$$Sensitivity = \frac{TP}{TP + FN} \qquad (7)$$

$$Specificity = \frac{TN}{TN + FP} \qquad (8)$$

where false negatives (*FN*) are the control subjects classified as AD patients, and false positives (*FP*) are the patients classified as control subjects. True positives (*TP*) and true negatives (*TN*) are the patients and control subjects correctly recognized, respectively. The area under the ROC curve (AROC) is a single number summary of performance. The ideal value is 1 and the worst case value is 0.5. A rough guide to classify the AROC is the

traditional academic point system: excellent (A) if AROC values are between 0.9 and 1, good (B) when the values are ranged from 0.8 to 0.89, fair (C) for values between 0.7 and 0.79, poor (D) for results between 0.6 and 0.69, and bad (E) if the AROC is in the range 0.5 – 0.59.

4. Results

We have used the *LZ* complexity to quantify the complexity in MEG time-series of 3392 samples. Fig. 3 summarizes the average *LZ* complexity values estimated for the patients with AD and the control subjects, for all the MEG channels (A1-A148). These mean values were higher in the control group's MEGs for all channels. We have averaged the *LZ* complexity values for all channels in both groups. The values obtained were 0.68 ± 0.06 (Mean ± SD) for the controls and 0.57 ± 0.09 (Mean ± SD) for the AD patients. These results suggest that the complexity, in the sense of number of new subsequences in the data, in AD patients' MEGs is lower than in control subjects' ones. Moreover, the differences were statistically significant for all channels ($p < 0.01$).

INSERT FIGURE 3 AROUND HERE

Afterwards, PCA was used to reduce the dimension of the results. In order to select the number of principal components to retain we used the Cattell's scree test. Fig. 4 shows an abrupt change in the slope of the curve between the first and the second principal components. Using this criterion, only the first principal component should be retained. Moreover, only the first eigenvalue was near 1, while the others were far from this value. The Kaiser's rule says that only the eigenvalues greater than 1 should be retained. In addition, the variance explained by the first factor is 83.7%, as can be noticed in Fig. 5. Therefore, only the first principal component was retained and the first component score calculated.

INSERT FIGURE 4 AROUND HERE

INSERT FIGURE 5 AROUND HERE

Fig. 6 shows the box plot for the first component score values estimated for AD patients and control subjects. A box plot provides a graphical summary of the data. From the visual inspection of the plots it becomes evident that there are clear differences between the first component scores for AD patients and controls subjects. We can detect this difference graphically by observing that the boxes of both groups do not overlap. In order to obtain a numerical result, a *t*-test was applied to the first principal score. A *p*-value of $3.77 \cdot 10^{-6}$ ($p << 0.01$) was obtained with this statistical method. Moreover, we can see that the variability is greater in the AD patients' data than in the control subjects' data.

INSERT FIGURE 6 AROUND HERE

Finally, the first principal score was analysed by means of a ROC curve. ROC curves are used to summarize the accuracy of diagnostic tests. Accuracy is a parameter that quantifies the total number of subjects (AD patients and control subjects) precisely classified. In a good test, the ROC curve is towards the upper left corner. This means that *Sensitivity* is high and *1-Specificity* is low. The AROC is a summary of performance: the larger the area, the better the diagnostic test. Our results show an AROC of 0.9002. This AROC value indicates the probability that a randomly chosen AD patient has a first principal score value lower than a randomly selected control subject. Moreover, the ROC curve can be used to

select an optimum selection threshold. This threshold is the cut-off point in which the highest accuracy is obtained. In our case, we obtained an optimum threshold of 0.0951 for the first principal score. For this value, the specificity was 85.71%, the sensitivity 80.95% and the accuracy 83.33%.

INSERT FIGURE 7 AROUND HERE

---

5. Discussion

We studied the MEG background activity in twenty one patients with probable AD and twenty one control subjects by means of the *LZ* complexity, a non-parametric measure that quantifies the complexity of the signals. Our purpose was to test the hypothesis that the complexity values were lower in the AD patients' group than in the control subjects' group.

The *LZ* complexity values were significantly lower in the AD patients indicating globally decreased complexity of brain activity in AD. The differences between groups were statistically significant in all channels ($p < 0.01$, Student's *t*-test). PCA was then applied to the *LZ* complexity results in order to reduce their dimension and only the first component was retained. Finally, the first component score was analysed graphically using a box plot and a ROC curve. A specificity of 85.71%, a sensitivity of 80.95% and an accuracy of 83.33% were obtained. This loss of complexity on the MEG background activity in AD might arise from neuronal death, deficiency of neurotransmitters like acetylcholine, and/or loss of connectivity of local neuronal networks [10].

Our results agree with previous research works of the MEG/EEG background activity in AD. Most of these studies were carried out using the correlation dimension (D_2) [13, 15, 16, 20-23] and the first positive Lyapunov exponent (L_1) [13, 23]. These studies showed that

patients with AD had significantly lower complexity in the brain electrical and magnetic activity than age-matched non-demented controls. Nevertheless, both methods have some drawbacks. Firstly, a large number of data points is necessary to estimate them. Secondly, the Grassberger and Procaccia algorithm [11, 12], usually used to estimate the D_2, assumes the signals to be stationary and noise free. Unfortunately, these problems can not be solved for physiological signals. Moreover, the complexity values depend on parameters used to compute them, such as the embedding dimension, the time delay or the number of data points. On the other hand, *LZ* complexity does not need any parameters and it is model-independent: only those differences between activity patterns that make a difference to the underlying system itself are considered, no matter whether the system is dominated by deterministic chaos or stochastic processes [28]. Thus, *LZ* complexity may be a better non-linear method for estimate the complexity of biomedical signals. In fact, this non-linear method has been applied to measure the complexity of EEG background activity in AD patients and in control subjects [46].

Another measure of neural complexity (C_N) was proposed by Tononi *et al.* [25] in 1994. Their hypothesis that the C_N decreased in neurological disorders as AD is in agreement with our study. However, this hypothesis was rejected by van Cappellen van Walsum *et al.* [16]. In their study, C_N and D_2 were applied to MEG data in different frequency bands. Mean neural complexity is higher in AD group than in control subjects group for the low frequency bands (2-4 Hz, 4-8 Hz and 8-12 Hz) and lower in the others (from 14 to 40 Hz). Other authors have studied spectral analysis in AD patients' EEGs [7] and MEGs [8, 9]. They found an increased EEG/MEG activity in AD patients compared with control subjects in delta and theta bands. These studies show a neuronal dysfunction associated with AD.

LZ complexity is an easy and fast method to measure the time series complexity. Only two simple mathematical operations are needed for its calculation: sequence comparison and

number accumulation. Moreover, the median value used as threshold is robust to outliers. These advantages support the proposal that the *LZ* complexity is a good measure of the complexity of biological signals as the MEG.

Nevertheless, our method has some limitations that must be considered:

- *LZ* complexity is based on a coarse-graining measure of the measurements [28]. The MEG data were transformed into a pattern of a few symbols, only two (0-1) in our study. Thus, it is possible that other conversions with more symbols could keep more information from the signal. Previous studies have suggested that a binary conversion is enough to study the complexity of a system [28], although a recent paper does seem to contradict this asseveration [46].
- In our study AD patients group and control group were carefully matched for age (AD patients, mean age ± SD = 73.62 ± 8.38; control subjects, mean age ± SD = 70.29 ± 7.07). Therefore, the complexity loss might represent the cognitive dysfunction in AD. Nevertheless, the decreased complexity is not specific of AD. This complexity loss appears also on other types of dementia as schizophrenia [47], epilepsy [48] and vascular dementia [23]. Thus, we need a larger database including recordings from patients with these dementias to confirm the performance of our method.

In summary, this paper presents the *LZ* complexity as a method to measure the complexity of the MEG background activity. We have found significant differences between the two groups we have studied, AD patients and control subjects. This pilot study is only a first step for the use of non-linear analysis in the diagnosis of AD and further investigation is needed to confirm our results.

Acknowledgements

This research has been supported by a project from the Consejería de Fomento de la Junta de Castilla y León and by the Fundación Caja Madrid. The authors would like to thank the Asociación de Enfermos de Alzheimer (AFAL) for supplying the patients who have participated in our study.

References

[1] Bird TD. Alzheimer's disease and other primary dementias. In: Braunwald E, Fauci AS, Kasper DL, Hauser SL, Longo DL, Jameson JL, editors. 'Harrison's Principles of Internal Medicine'. New York: The McGraw-Hill Companies Inc, 2001:2391–9.

[2] Cummings BJ, Pike CJ, Shankle R, Cotman CW. Beta-amyloid deposition and other measures of neuropathology predict cognitive status in Alzheimer's disease. Neurobiol Aging 1996;17:921–33.

[3] Trojanowski JQ. The cellular and molecular correlates of cognitive impairments in the Alzheimer's disease brain. Neurobiol Aging 1996;17:941–3.

[4] Reisberg B. Functional assessment staging (FAST). Psychopharmacol Bull 1988;24:653–9.

[5] Folstein MF, Folstein SE, McHugh PR. Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–98.

[6] Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV. Magnetoencephalography – theory, instrumentation, and applications to noninvasive studies of the working human brain. Rev Mod Phys 1993;65:413-97.

[7] Signorino M, Pucci E, Belardinelli N, Nolle G, Angeleri F. EEG spectral analysis in vascular and Alzheimer dementia. Electroen clin Neuro 1995;94:313–25.

[8] Osipova D, Ahveninen J, Jensen O, Ylikoski A, Pekkonen E. Altered generation of spontaneous oscillations in Alzheimer's disease. Neuroimage 2005;27:835–41.

[9] Fernández A, Maestú F, Amo C, Gil P, Fehr T, Wienbruch C, Rockstroh B, Elbert T, Ortiz T. Focal tempoparietal slow activity in Alzheimer's disease revealed by magnetoencephalography. Biol Psychiatry 2002;52:764–70.

[10] Jeong J. EEG dynamics in patients with Alzheimer's disease. Clin Neurophysiol 2004;115:1490–505.

[11] Grassberger P, Procaccia I. Characterization of strange attractors. Phys Rev Lett 1983;50 (5):346–9.

[12] Grassberger P, Procaccia I. Measuring the strangeness of strange attractors. Physica D 1983;9:189–208.

[13] Jeong J, Kim SJ, Han SH. Non-linear dynamical analysis of the EEG in Alzheimer's disease with optimal embedding dimension. Electroenceph clin Neurophysiol 1998;106:220–228.

[14] Kennel MB, Brown R, Abarbanel HDI. Determining embedding dimension for phase-space reconstruction using a geometrical construction. Phys Rev A 1992;45:3403–11.

[15] Jelles B, van Birgelen JH, Slaets JPJ, Hekster REM, Jonkman EJ, Stam CJ. Decrease of non-linear structure in the EEG of Alzheimer patients compared to healthy controls. Clin Neurophysiol 1999;110:1159–1167.

[16] van Cappellen van Walsum AM, Pijnenburg YAL, Berendse HW, van Dijk BW, Knol DL, Scheltens Ph, Stam CJ. A neural complexity measure applied to MEG data in Alzheimer's disease. Clin Neurophysiol 2003;114:1034–1040.

[17] Woyshville MJ, Calabrese JR. Quantification of occipital EEG changes in Alzheimer's disease utilizing a new metric: the fractal dimension. Biol Psychiatry. 1994;35:381–7.

[18] Eckmann JP, Ruelle D. Fundamental limitations for estimating dimensions and Lyapunov exponents in dynamical systems. Physica D 1992;56:185–7.

[19] Tononi G, Edelman GM, Sporns O. Complexity and coherency: integrating information in the brain. Trends Cogn Sci 1998;2(12):474–84.

[20] Pritchard WS, Duke DW, Coburn KL, Moore NC, Tucker KA, Jann MW, Hostetler RM. EEG-based neural-net predictive classification of Alzheimer's disease versus control subjects is augmented by non-linear EEG measures. Electroenceph clin Neurophysiol 1994;91:118–130.

[21] Stam CJ, Jelles B, Achtereekte HAM, Rombouts SARB, Slaets JPJ, Keunen RWM. Investigation of EEG nonlinearity in dementia and Parkinson's disease. Electroenceph clin Neurophysiol 1995;95:309–317.

[22] Besthorn C, Zerfass R, Geiger-Kabisch C, Sattel H, Daniel S, Schreiter-Gasser U, Förstl H. Discrimination of Alzheimer's disease and normal aging by EEG data. Electroenceph clin Neurophysiol 1997;103:241–248.

[23] Jeong J, Chae JH, Kim SY, Han SH. Nonlinear dynamic analysis of the EEG in patients with Alzheimer's disease and vascular dementia. J Clin Neurophysiol 2001;18:58–67.

[24] Kolmogorov AN. Three approaches to the quantitative definition of information. Infor Trans 1965;1:3–11.

[25] Tononi G, Sporns O, Edelman GM. A measure for brain complexity: Relating functional segregation and integration in the nervous system. Proc Natl Acad Sci USA 1994;91:5033–5037.

[26] Zozor S, Ravier P, Buttelli O. On Lempel-Ziv complexity for multidimensional data analysis. Physica A 2005;345:285–302.

[27] Lempel A, Ziv J. On the complexity of finite sequences. IEEE Trans Inform Theory 1976;22:75–81.

[28] Zhang XS, Roy RJ, Jensen EW. EEG complexity as a measure of depth of anesthesia for patients. IEEE Trans Biomed Eng 2001;48:1424–33.

[29] Zhang XS, Roy RJ. Derived fuzzy knowledge model for estimating the depth of anesthesia. IEEE Trans Biomed Eng 2001;48:312–23.

[30] Huang L, Yu P, Ju F, Cheng J. Prediction of response to incision using the mutual information of electroencephalogram during anesthesia. Med Eng Phys 2003;25:321–7.

[31] Zhang XS, Roy RJ. Predicting movement during anaesthesia by complexity analysis of electroencephalograms. Med Biol Eng Comput 1999;37(3):327–34.

[32] Zhang JW, Zheng CX, Jiang DZ. EEG complexity measurement of focal ischemic cerebral injury. Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society 1998;20(4):2027–9.

[33] Wu X, Xu J. Complexity and brain functions. Acta Biophysica Sinica 1991;7:103–6.

[34] Xu J, Liu ZR, Liu R, Yang QF. Information transformation in human cerebral cortex. Physica D 1997;106:363–74.

[35] Amigó JM, Szczepánski J, Wajnryb E, Sanchez-Vives, MV. Estimating the entropy rate of spike trains via Lempel-Ziv complexity. Neural Comput 2004;16:717–36.

[36] Szczepánski J, Amigó JM, Wajnryb E, Sanchez-Vives, MV Application of Lempel-Ziv complexity to the analysis of neural discharges. Network-Comp Neural 2003;14:335–50.

[37] Zhang XS, Zhu YS, Thakor NV, Wang ZZ. Detecting ventricular tachycardia and fibrillation by complexity measure. IEEE Trans Biomed Eng 1999;46:548–555.

[38] Nagarajan R. Quantifying physiological data with Lempel-Ziv complexity – Certain issues. IEEE Trans Biomed Eng 2002;49:1371–1373.

[39] Radhakrishnan N, Gangadhar BN. Estimating regularity in epileptic seizure time-series data. A complexity-measure approach. IEEE Eng Med Biol 1998;17:89–94.

[40] McKhann G, Drachman D, Folstein M. Clinical diagnosis of Alzheimer's disease: report of NINCDS-ADRDA work group under the auspices of department of health and human services task force on Alzheimer's disease. Neurology 1984;34:939–44.

[41] Kaspar F, Schuster HG. Easily calculable measure for the complexity of spatiotemporal patterns. Phys Rev A 1987;36(2):842–8.

[42] Yan R, Gao RX. Complexity as a measure for machine health evaluation. IEEE Trans Instrum Meas 2004;53(4):1327–34.

[43] Kaiser HF. The application of electronic computers to factor analysis. Educ Phychol Meas 1960;20:141–51.

[44] Cattell RB. The scree test for the number of factors. Multivar Behav Res 1966;1(2):140–61.

[45] Zweig MH, Campbell G. Receiver-Operating Characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 1993;39:561–577.

[46] Abásolo D, Hornero R, Gómez C, García M, López M. Analysis of EEG background activity in Alzheimer's disease patients with Lempel-Ziv complexity and Central Tendency Measure. Med Eng Phys 2005. Article in press. doi:10.1016/j.medengphy. 2005.07.004.

[47] Na SH, Jin SH, Kim SY, Ham BJ. EEG is schizophrenic patients: mutual information analysis. Clin Neurophysiol 2002;113: 1954–1960.

[48] Jing H, Takigawa M. Comparison of human ictal, interictal and normal non-linear component analyses. Clin Neurophysiol 2000;111: 1282–1292.

Tables

Table 1. Sociodemographic data of Alzheimer's disease patients and control subjects.

AD patients				Control subjects			
	Age	MMSE	FAST		Age	MMSE	FAST
Alz-1	71	15	4	Con-1	68	30	2
Alz-2	67	12	4	Con-2	61	29	2
Alz-3	56	14	4	Con-3	70	30	2
Alz-4	64	15	4	Con-4	64	30	2
Alz-5	59	20	4	Con-5	60	30	1
Alz-6	60	16	4	Con-6	63	30	1
Alz-7	72	15	4	Con-7	73	29	1
Alz-8	71	15	4	Con-8	69	29	1
Alz-9	75	22	4	Con-9	56	27	2
Alz-10	82	21	4	Con-10	79	29	2
Alz-11	72	17	4	Con-11	79	30	2
Alz-12	80	24	4	Con-12	75	29	2
Alz-13	83	10	5	Con-13	67	29	2
Alz-14	77	21	4	Con-14	68	29	2
Alz-15	82	19	4	Con-15	84	29	2
Alz-16	83	20	4	Con-16	68	27	2
Alz-17	80	20	4	Con-17	73	30	2
Alz-18	73	23	4	Con-18	71	29	1
Alz-19	77	24	4	Con-19	74	30	2
Alz-20	79	19	4	Con-20	78	27	2
Alz-21	83	16	4	Con-21	76	29	2
Mean ± SD	73.62 ± 8.38	18.00 ± 3.91	4.05 ± 0.22	Mean ± SD	70.29 ± 7.07	29.10 ± 1.00	1.71 ± 0.46

Figure legends

Fig. 1. Block diagram of the steps followed in the MEG analysis: signal pre-processing, complexity analysis, principal components analysis and statistical analysis.

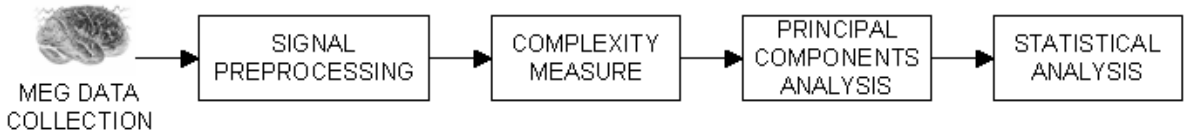


Fig. 2. Effect of the epoch length on the $C(n)$ value.

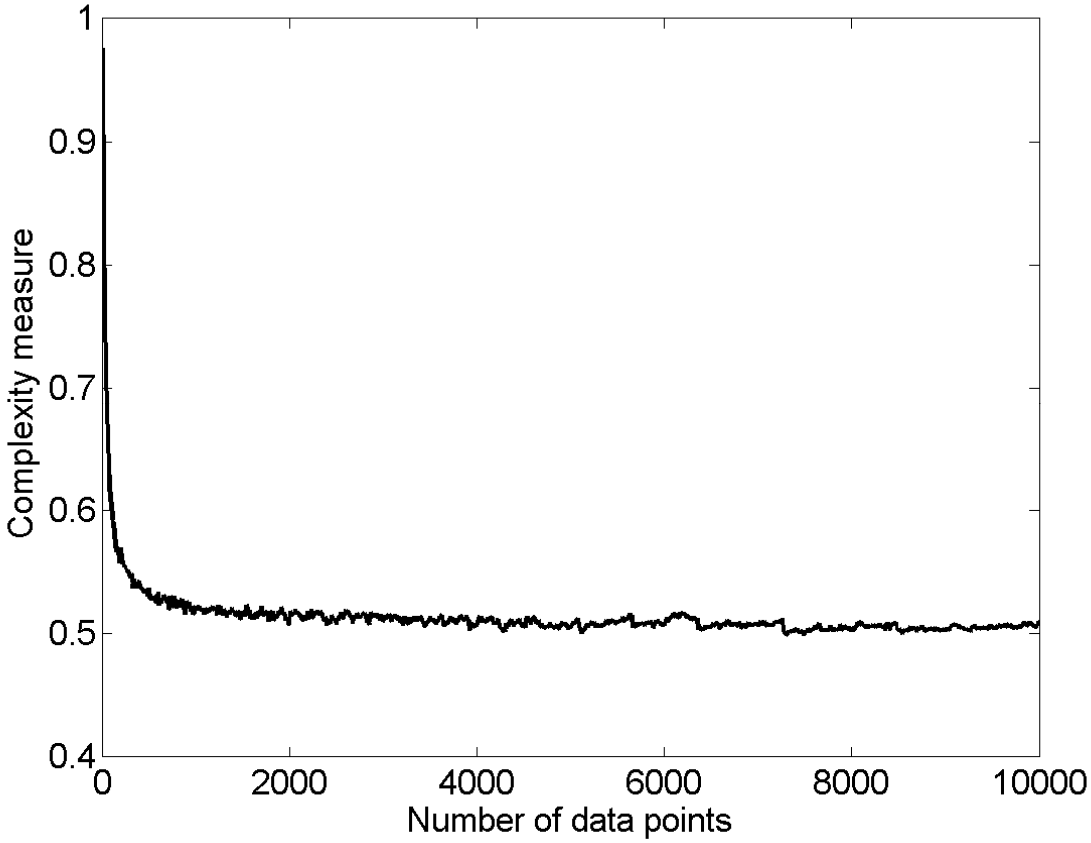


Fig. 3. Average LZ complexity values of the MEGs in AD patients and control subjects for all channels, from A1 to A148.

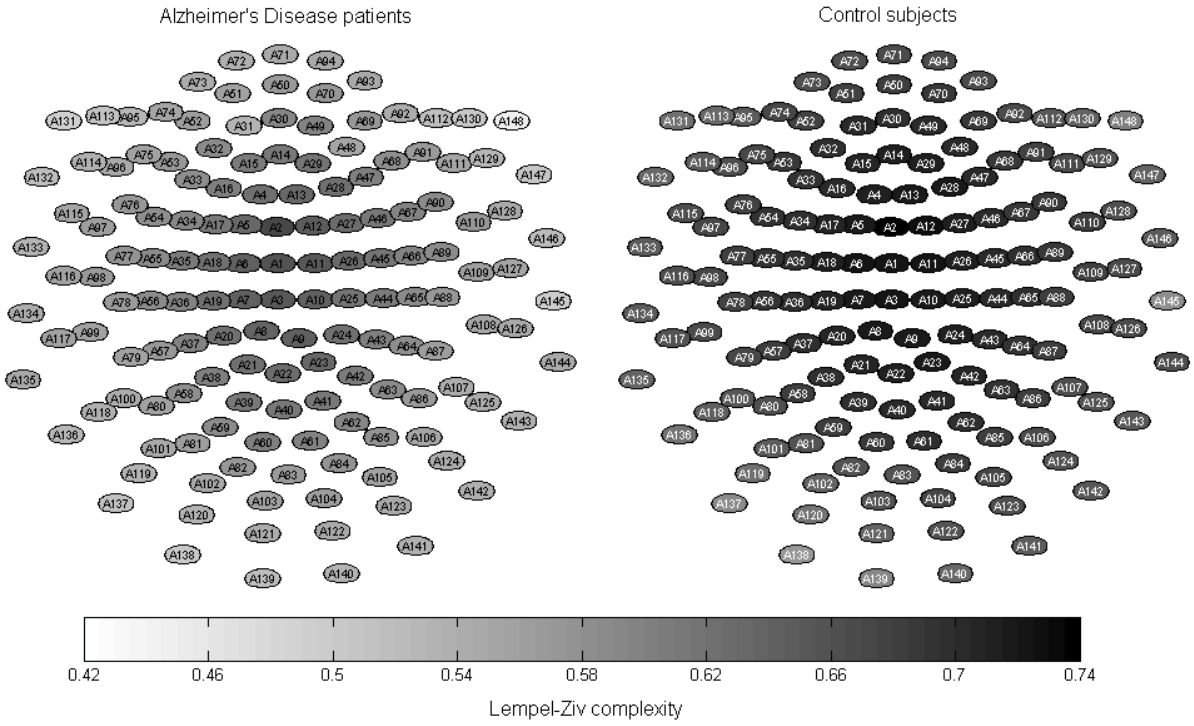


Fig. 4. Scree plot of the ten first eigenvalues for the PCA in AD patients and control subjects.

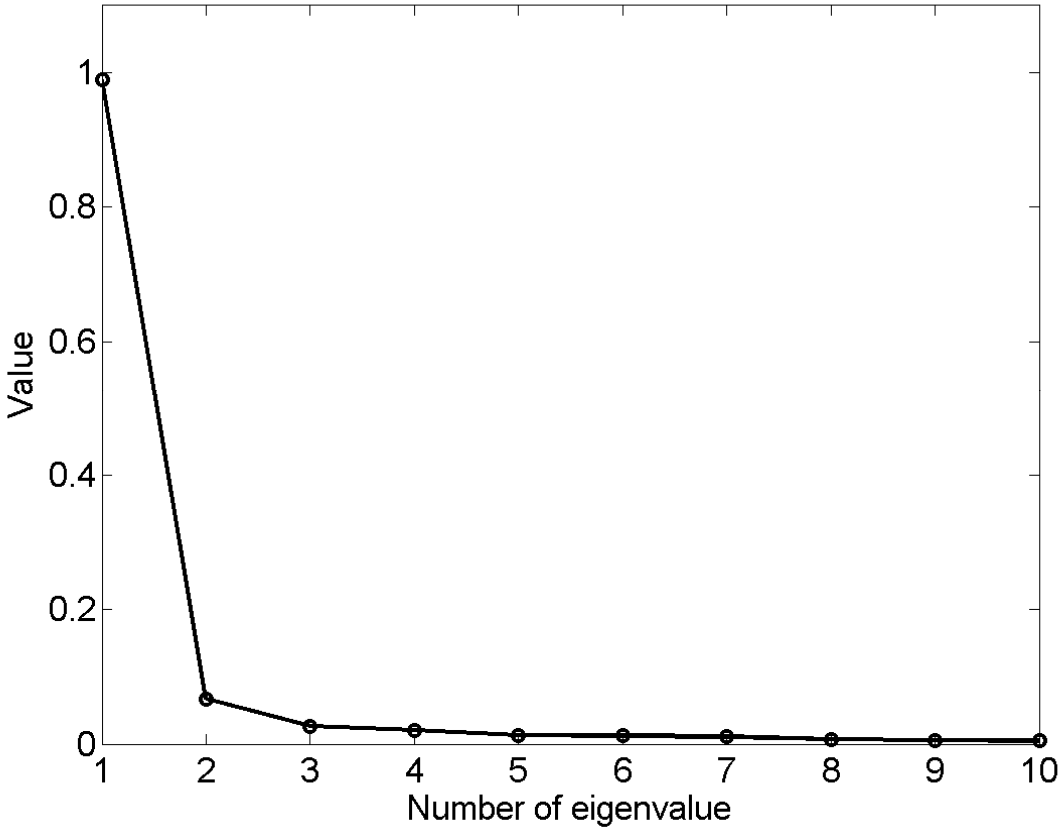


Fig. 5. Percent variance explained by each principal component. The variance explained by the first principal component was 83.7%.

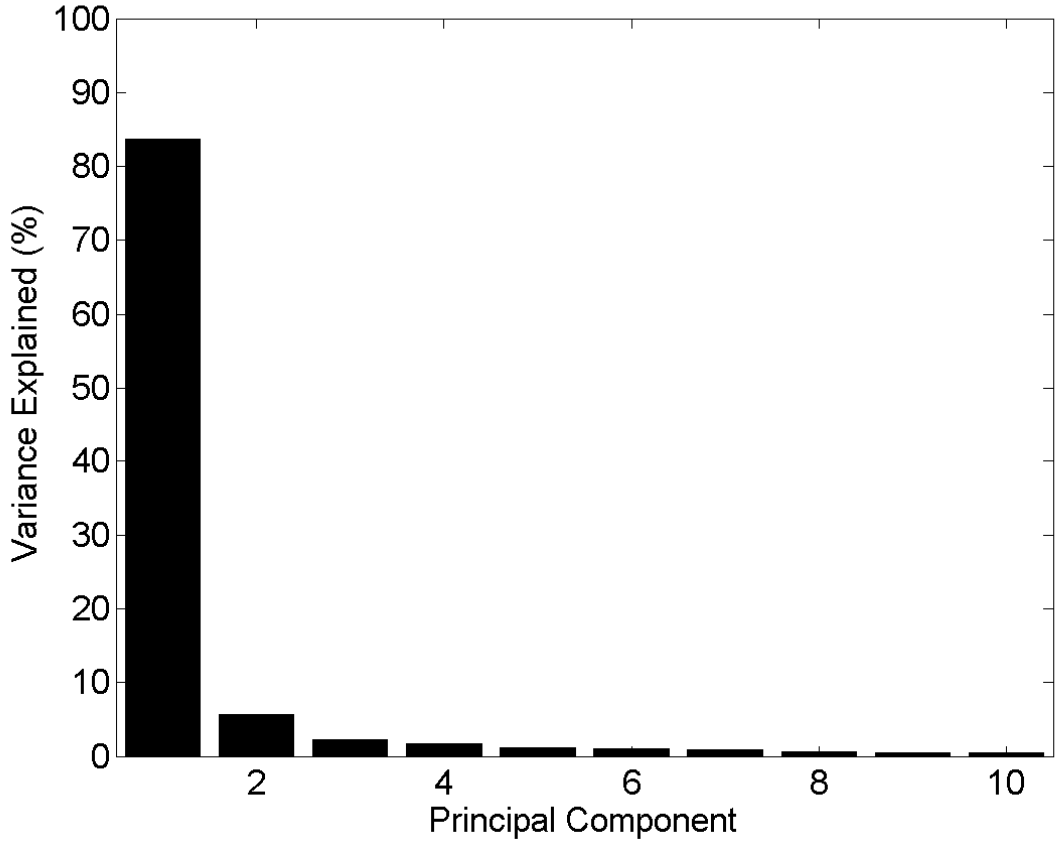


Fig. 6. Box plot for the first component score for the AD patients group and the control group. A potential outlier is marked with a plus sign in the plot.

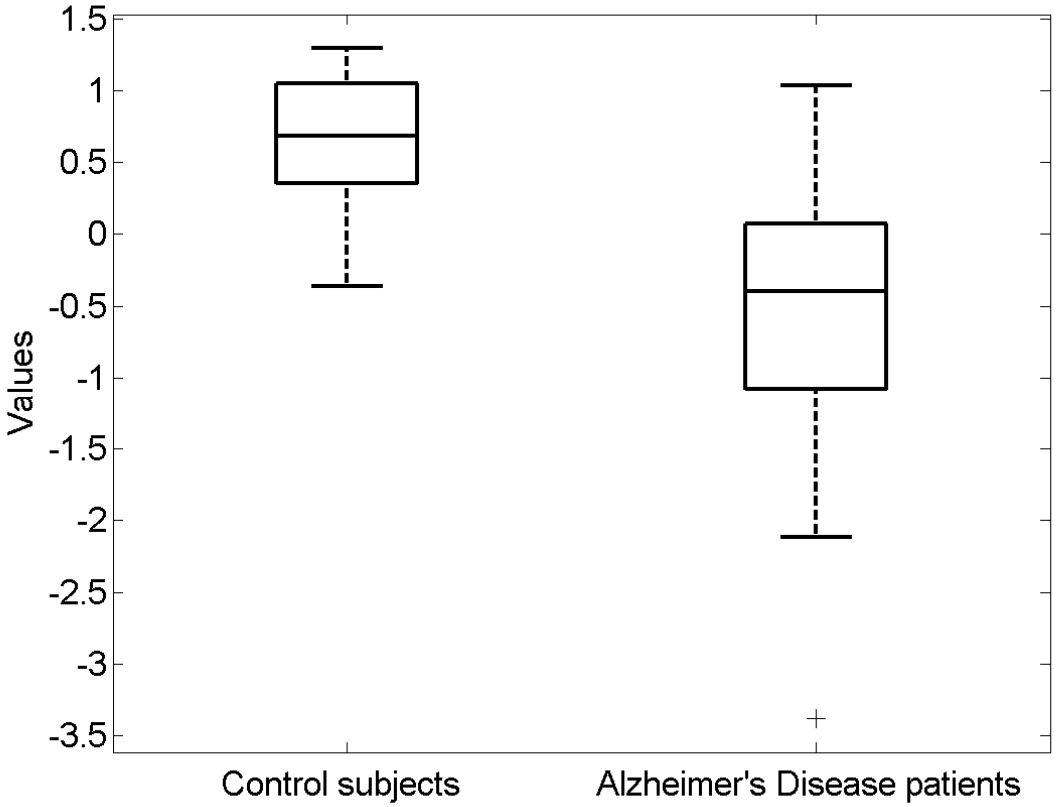
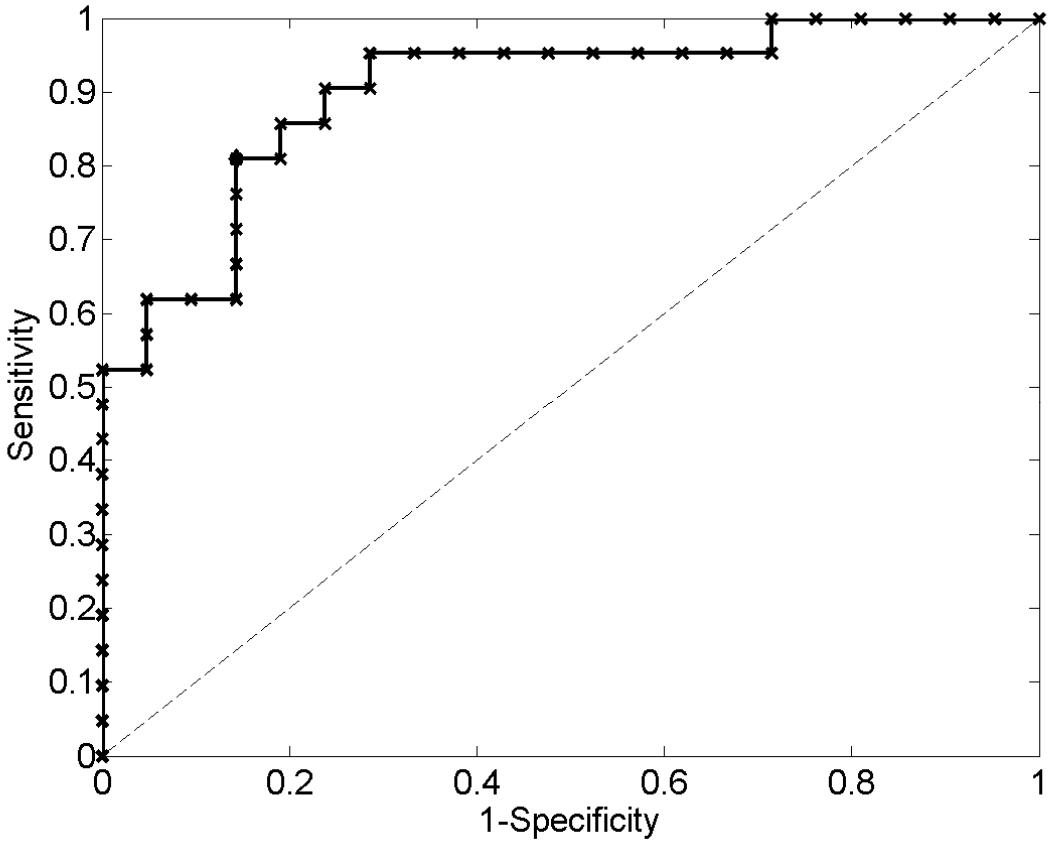


Fig. 7. ROC curve for discriminating AD patients and control subjects for the first component score, obtained from the LZ complexity values at all MEG channels. The ROC curve values are marked with a cross. The symbol ♦ indicates the optimum cut-off point.



Conflict of interest statement

There are no conflicts of interest that could inappropriately influence this research work.