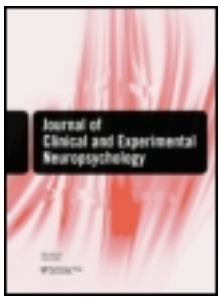

Journal of Clinical and Experimental Neuropsychology

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ncen20

Using artificial neural networks in clinical neuropsychology: High performance in mild cognitive impairment and Alzheimer's disease

María Quintana [a], Joan Guàrdia [b], Gonzalo Sánchez-Benavides [a], Miguel Aguilar [c], José Luis Molinuevo [d], Alfredo Robles [e], María Sagrario Barquero [f], Carmen Antúnez [g], Carlos Martínez-Parra [h], Anna Frank-García [i], Manuel Fernández [j], Rafael Blesa [k], Jordi Peña-Casanova [l] & for the Neuronorma Study Team

[a] Group of Behavioral Neurology, Neuropsychopharmacology Program, Institut Municipal d'Investigació Mèdica, Barcelona, Spain

[b] Deparment de Metodologia de les Ciències del Comportament, Facultat de Psicologia, Universitat de Barcelona, Institut de Recerca Cervell, Cognició i Conducta (IR3C) Grup de Recerca Tècniques Estadístiques Avançades Aplicades a la Psicologia, Barcelona, Spain

[c] Service of Neurology, Hospital Mútua de Terrassa, Terrassa, Spain

[d] Service of Neurology, Hospital Clínic, Barcelona, Spain

[e] Service of Neurology, Hospital Clínico Universitario, Santiago de Compostela, Spain

[f] Service of Neurology, Hospital Clínico San Carlos, Madrid, Spain

[g] Service of Neurology, Hospital Virgen Arrixaca, Murcia, Spain

[h] Service of Neurology, Hospital Virgen Macarena, Sevilla, Spain

[i] Department of Neurology, Hospital Universitario La Paz, Madrid, Spain

[j] Service of Neurology, Hospital de Cruces, Bilbao, Spain

[k] Service of Neurology, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

[l] Section of Behavioral Neurology and Dementias, Hospital del Mar, Parc de Salut Mar, Barcelona, Spain

Available online: 14 Dec 2011

JOURNAL OF CLINICAL AND EXPERIMENTAL NEUROPSYCHOLOGY
2011, i*First*, 1–14

Psychology Press
Taylor & Francis Group

Using artificial neural networks in clinical neuropsychology: High performance in mild cognitive impairment and Alzheimer's disease

María Quintana[1], Joan Guàrdia[2], Gonzalo Sánchez-Benavides[1], Miguel Aguilar[3], José Luis Molinuevo[4], Alfredo Robles[5], María Sagrario Barquero[6†], Carmen Antúnez[7], Carlos Martínez-Parra[8], Anna Frank-García[9], Manuel Fernández[10], Rafael Blesa[11], and Jordi Peña-Casanova[1,12] for the Neuronorma Study Team

[1]Group of Behavioral Neurology, Neuropsychopharmacology Program, Institut Municipal d'Investigació Mèdica, Barcelona, Spain
[2]Deparment de Metodologia de les Ciències del Comportament, Facultat de Psicologia, Universitat de Barcelona, Institut de Recerca Cervell, Cognició i Conducta (IR3C) Grup de Recerca Tècniques Estadístiques Avançades Aplicades a la Psicologia, Barcelona, Spain
[3]Service of Neurology, Hospital Mútua de Terrassa, Terrassa, Spain
[4]Service of Neurology, Hospital Clínic, Barcelona, Spain
[5]Service of Neurology, Hospital Clínico Universitario, Santiago de Compostela, Spain
[6]Service of Neurology, Hospital Clínico San Carlos, Madrid, Spain
[7]Service of Neurology, Hospital Virgen Arrixaca, Murcia, Spain
[8]Service of Neurology, Hospital Virgen Macarena, Sevilla, Spain
[9]Department of Neurology, Hospital Universitario La Paz, Madrid, Spain
[10]Service of Neurology, Hospital de Cruces, Bilbao, Spain
[11]Service of Neurology, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain
[12]Section of Behavioral Neurology and Dementias, Hospital del Mar, Parc de Salut Mar, Barcelona, Spain

†Recently deceased.

Members of the Neuronorma Study Team: *Steering committee*: J. P.-C., Hospital del Mar, Barcelona, Spain; R. B., Hospital de la Santa Creu i Sant Pau, Barcelona, Spain; M. A., Hospital Mútua de Terrassa, Terrassa, Spain. *Principal investigators*: J. P.-C., Hospital del Mar, Barcelona, Spain; R. B., Hospital de la Santa Creu i Sant Pau, Barcelona, Spain; M. A., Hospital Mútua de Terrassa, Terrassa, Spain; J. L. M., Hospital Clínic, Barcelona, Spain; A. R., Hospital Clínico Universitario, Santiago de Compostela, Spain; M.S.B. (deceased), Hospital Clínico San Carlos, Madrid, Spain; C. A., Hospital Virgen Arrixaca, Murcia, Spain; C. M.-P., Hospital Virgen Macarena, Sevilla, Spain; A. F.-G., Hospital Universitario La Paz, Madrid, Spain; M. F., Hospital de Cruces, Bilbao, Spain. *Genetics substudy*: Rafael Oliva, Service of Genetics, Hospital Clínic, Barcelona, Spain. *Neuroimaging substudy*: Beatriz Gómez-Ansón, Radiology Department and IDIBAPS, Hospital Clínic, Barcelona, Spain. *Research Fellows*: Gemma Monte, Elena Alayrach, Aitor Sainz, and Claudia Caprile, Fundació Clinic, Hospital Clinic, Barcelona, Spain; Gonzalo Sánchez, Behavioral Neurology Group, Institut Municipal d'Investigació Médica, Barcelona, Spain. *Clinicians, psychologists, and neuropsychologists*: Nina Gramunt (Coordinator), Peter Böhm, Sonia González, Yolanda Buriel, María Quintana, Sonia Quiñones, Gonzalo Sánchez, Rosa M. Manero, Gracia Cucurella, Institut Municipal d'Investigació Mèdica, Barcelona, Spain; Eva Ruiz, Mónica Serradell, Laura Torner, Hospital Clínic, Barcelona, Spain; Dolors Badenes, Laura Casas, Noemí Cerulla, Silvia Ramos, Loli Cabello, Hospital Mútua de Terrassa, Terrassa, Spain; Dolores Rodríguez, Clinical Psychology and Psychobiology Dept. University of Santiago de Compostela, Spain; María Payno, Clara Villanueva, Hospital Clínico San Carlos, Madrid, Spain; Rafael Carles, Judit Jiménez, Martirio Antequera, Hospital Virgen Arrixaca, Murcia, Spain; Jose Manuel Gata, Pablo Duque, Laura Jiménez, Hospital Virgen Macarena, Sevilla, Spain; Azucena Sanz, María Dolores Aguilar, Hospital Universitario La Paz, Madrid, Spain; Ana Molano, Maitena Lasa, Hospital de Cruces, Bilbao, Spain. *Data management and biometrics*: Josep Maria Sol, Francisco Hernández, Irune Quevedo, Anna Salvà, Verónica Alfonso, European Biometrics Institute, Barcelona, Spain. *Administrative management*: Carme Pla (deceased), Romina Ribas, Department of Psychiatry and Forensic Medicine, Universitat Autònoma de Barcelona, and Behavioral Neurology Group, Institut Municipal d'Investigació Mèdica, Barcelona, Spain. *English edition*: Stephanie Lonsdale, Program of Neuropsychopharmacology, Institut Municipal d'Investigació Mèdica, Barcelona, Spain. ***Conflict of interest:*** None declared.

Address correspondence to Jordi Peña-Casanova, Institut Municipal d'Investigació Mèdica, Biomedical Research Park Building. Carrer Dr. Aiguader, 88, 08003 Barcelona, Spain (E-mail: jpcasanova@parcdesalutmar.cat).

http://www.psypress.com/jcen http://dx.doi.org/10.1080/13803395.2011.630651

Mild cognitive impairment (MCI) is a transitional state between normal aging and Alzheimer disease (AD). Artificial neural networks (ANNs) are computational tools that can provide valuable support to clinical decision making, classification, and prediction of cognitive functioning. The aims of this study were to develop, train, and explore and develop the ability of ANNs to differentiate MCI and AD, and to study the relevant variables in MCI and AD diagnosis. The sample consisted of 346 controls and 79 MCI and 97 AD patients. A linear discriminant analysis (LDA) and ANNs with 12 input neurons (10 subtests of a neuropsychological test, the abbreviated Barcelona Test; age; and education), 4 hidden neurons, and output neuron (diagnosis) were used to classify the patients. The ANNs were superior to LDA in its ability to classify correctly patients (100–98.33% vs. 96.4–80%, respectively) and showed better predictive performance. Semantic fluency, working and episodic memory and education showed up as the most significant and sensitive variables for classification. Our results indicate that ANNs have an excellent capacity to discriminate MCI and AD patients from healthy controls. These findings provide evidence that ANNs can be a useful tool for the analysis of neuropsychological profiles related to clinical syndromes.

Keywords: Artificial neural networks; Mild cognitive impairment; Alzheimer disease; Assessment classification.

Improvements in health care over the past 50 years have extended average life expectancy, which has resulted in a substantial increase in the numbers of individuals over 65 years of age (Hebert, Beckett, Scherr, & Evans, 2001). This fact is associated with an increase of age-related diseases such as Alzheimer disease (AD) and other dementias.

Early diagnosis, in which the concept of mild cognitive impairment (MCI) plays a fundamental role, is a crucial aspect in the treatment of AD. MCI is considered to be a state between normal cognition and dementia, characterized by deficits not explainable by age, educational background, or medical illness (Petersen, 2003). Several studies (Busse, Hensel, Gühne, Angermeyer, & Riedel-Heller, 2006; Gauthier et al., 2006; Manly et al., 2008; Visser, Kester, Jolles, & Verhey, 2006) have demonstrated that MCI is associated with an increased risk of developing dementia, usually AD. Nevertheless, this concept is heterogeneous (DeCarli, 2003; Petersen et al., 2001) because many factors converge in a possible progression to AD diagnosis and other diseases.

Several markers of AD conversion have been studied in the last few years, such as the presence of the apolipoprotein E (APOE) $\varepsilon 4$ allele (Fleisher et al., 2007), phosphorylated tau levels in the cerebrospinal fluid (Ewers et al., 2007), reduction in the hippocampal and entorhinal cortex volumes (Jack et al., 2008; Risacher et al., 2009), and amyloid deposition (Okello et al., 2009). As neuropsychological markers, episodic memory and executive functions are particularly well-studied predictors of conversion (Landau et al., 2010; Rozzini et al., 2007; Tabert et al., 2006). Neuropsychological assessment is, therefore, essential, not only in the diagnosis but also in the monitoring of conversion to dementia. The study of the appropriate neuropsychological test and technique to support clinical decisions is, therefore, a relevant issue in the prediction of developing dementia.

Artificial neural networks (ANNs) have been proposed as viable computational tools in order to provide support to clinical decision making, classification, and the prediction of cognitive functioning (Lisboa & Taktak, 2006). ANNs are computing paradigms, inspired by neurosciences (Hebb, 1949), where the organization and storage of information are connected by computational units or nodes to allow signals to travel through the network. The models reflect the highly interactive processing functions of the human brain and are able to modify their internal structure in relation to a function objective. Like the brain, ANNs recognize patterns, manage data, and learn (Buscema et al., 2004; Di Luca et al., 2005). In statistical terms, ANNs are nonparametric models that carry out estimations of the so-called *free model* (Perez & Martin, 2003). They can provide several advantages with respect to conventional statistical models. In fact, a number of studies have reported their superiority to other statistical approaches such as logistic regression or linear discriminant analysis (LDA; French, Dawson, & Dobbs, 1997; Parsons, Rizzo, & Buckwalter, 2004).

A number of different areas of science, including medicine, biology, and psychology, have benefited from the applications of ANNs. Their ability to accurately classify and recognize patterns has encouraged researchers to employ them in solving numerous clinical problems. In medicine, the clinical applications include diagnosis and outcome prediction, clinical pharmacology, and neuroimaging amongst many others (Batx, 1995; Ramesh, Kambhampati, Monson, & Drew, 2004). In neurology, ANNs have been used for the

diagnosis and classification of neurodegenerative disorders with extrapyramidal features (Litvan et al., 1996). In dementia, AD is a good model for ANN application due to the fact that the course of the disease is nonlineal; progression is generally slower in the early stages and more rapid in the middle phase (Doody, Massman, & Dunn, 2001; Tandon, Adak, & Kaye, 2006). ANNs have been successfully implemented in the classification of pharmacological responders to donepezil in a group of AD patients (Mecocci et al., 2002), identifying the time-course of the disease in longitudinal data (Tandon et al., 2006), and studying the gender-related differences in clinical presentation (Grossi, Massini, Buscema, Savarè, & Maurelli, 2005). They have been employed as a practical analytic tool for electrophysiological data (Buscema, Rossini, Babiloni, & Grossi, 2007; Lehmann et al., 2007), neurofunctional imaging (DeFigueiredo et al., 1995; Horn et al., 2009; Page, Howard, O'Brien, Buxton-Thomas, & Pickering, 1996), biomarkers (Di Luca et al., 2005), and neuropathological findings in the prediction of AD (Buscema et al., 2004; Grossi, Buscema, Snowdon, & Antuono, 2007). In animal models, Leighty et al. (2008) applied ANNs in order to examine behavioral data in AD transgenic mice. Only one study where ANNs were used to identify the predictive values of risk factors on the conversion of amnestic MCI to AD has been described (Tabaton et al., 2010). This study supports the utility of ANN analysis in the interpretation of data from heterogeneous and distinct sources.

Although ANNs have been applied to various areas of neuroscience research there are, to our knowledge, no reports of MCI and AD with cognitive measures. In this study, our primary aim was to explore and develop the ability of ANNs to differentiate healthy controls and MCI and AD patients. The second aim was to study the more relevant variables in MCI and AD diagnosis.

METHOD

Participants

A whole sample of 522 subjects from the Neuronorma project took part in this study. They were divided into three groups: 346 healthy elderly participants, 79 patients diagnosed with MCI, and 97 patients diagnosed with AD.

The Spanish Multicenter Normative Studies (Neuronorma project) was designed to provide normative data for people aged over 49 years for commonly used neuropsychological tests (Peña-Casanova et al., 2009). This project was performed in nine services of neurology and units of neuropsychology in different Spanish regions.

Healthy elderly participants were recruited from a variety of sources such as: (a) spouses of patients evaluated at the participating centers, (b) various senior citizen activity centers, and (c) by word of mouth. MCI and AD patients were consecutively recruited at neurology services from the participating centers.

Three hundred and forty six healthy elderly controls (age range: 50–90 years) were studied. Entry criteria included consecutive individuals according to the following inclusion and exclusion criteria. *Inclusion criteria* were: (a) signed informed consent, (b) subjects of both genders aged over 49 years, (c) Spanish speakers with at least a minimal capacity in writing, (d) community dwelling and independent functioning individuals as measured by the Interview for Deterioration of Daily Living in Dementia (IDDD; Teunisse, Derix, & Cléber, 1991), and (e) absence of cognitive impairment, measured by the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975), adjusted for age and education.

Exclusion criteria were: (a) personal history of central nervous disease possibly causing neuropsychological deficits (e.g., stroke, epilepsy, movement disorder, multiple sclerosis, brain tumor, severe head trauma), (b) score of 4 or more on the Modified Ischemia Scale (Rosen, Terry, Fuld, Katzman, & Peck, 1980), (c) history of alcohol or other psychotropic substance abuse, (d) presence of active or uncontrolled systemic diseases associated with cognitive impairment (e.g., diabetes mellitus, hypothyroidism, B12 deficiency), (e) history of psychiatric diseases (e.g., major depression, bipolar mood disorder, psychosis), and (f) presence of severe sensorial deficits (loss of vision and/or hearing) that might have impeded the administration of cognitive tests.

Sample characteristics, recruitment procedures, and general methods of the Neuronorma project have been reported in a previous paper (Peña-Casanova et al., 2009). Seventy-nine patients were diagnosed with MCI following IPA–WHO (International Psychogeriatric Association–World Health Organization; Levy, 1994) criteria. The above-mentioned criteria are the following: (a) no age restriction (in our project the subjects had to be at least 50 years old); (b) decline of cognitive capacity affirmed by the patient and/or informant; (c) gradual decline and of minimal duration

of 6 months; (d) any of the following cognitive functions could be affected: memory and learning, attention and concentration, thinking, language or visuospatial functioning; (e) neuropsychological performance more than 1 standard deviation below the age and education norms in well-standardized neuropsychological tests; (f) the disorder does not have sufficient intensity to establish the diagnosis of dementia, nor does delirium exist; (g) there do not exist any cerebral, systemic, or psychiatric processes that could explain the symptoms. In our study, the test selected as neuropsychological test was the abbreviated Barcelona Test, being a test that assesses each of the cognitive functions that can be affected in the MCI, such as memory and learning, attention and concentration, thinking, language, or visuospatial functioning.

In our study we considered that MCI equated to Stage 3 of the GDS (Global Deterioration Scale; Reisberg, Ferris, de Leon, & Crook, 1972). It is crucial to indicate that diagnosis of MCI was done using clinical criteria alone; thus it is a non-psychometric diagnosis, and therefore the score in the Barcelona Test must be taken as another measure, but not the only measure. The diagnosis of MCI does not focus only on the score of a neuropsychological test.

Ninety-seven patients with a diagnosis of probable AD fulfilled *DSM–IV* (*Diagnostic and Statistical Manual for Mental Disorders–Fourth Edition*) criteria (American Psychiatric Association, 1994) and NINCDS-ADRDA (National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association) criteria (McKhann et al., 1984); they were at Stages 4 and 5 in the GDS. Subjects rated GDS 6 and 7 were excluded. The GDS scale used in this study includes stages of cognitive impairment from "cognitive normality" up to the most advanced phases of dementia. In no case was it used to distinguish between MCI and AD.

The majority (95%) of study participants were right-handed, as confirmed by the Edinburgh Handedness Inventory (Oldfield, 1971). The ethnic background of all participants was Caucasian, and all were living in Spain.

Approval for the study was obtained from the Research Ethics Committee of the Municipal Institute of Medical Care of Barcelona, Spain and the participating centers. The study was conducted in accordance with the Declaration of Helsinki and its subsequent amendments and the European Union regulations concerning medical research. All participants signed an informed consent before being tested; they received no financial reimbursement or any other compensation.

Neuropsychological evaluation

The MMSE, a global cognition measure (Folstein et al., 1975) in a validated Spanish version (Blesa et al., 2001) was used to select study participants. The study cutoff, adjusted for age and education, was 24. Functional changes were evaluated by the IDDD (Teunisse et al., 1991) in its validated Spanish version (Böhm et al., 1998).

The abbreviated Barcelona Test (a-BT) was administered as part of a larger battery of neuropsychological measures in the Neuronorma project (Peña-Casanova et al., 2009). Testing and scoring were performed by neuropsychologists specifically trained for this project. Standard administration and scoring procedures were followed as outlined in the original a-BT manual (Peña-Casanova, 1990). This test has been recently described in detail (Quintana et al., 2011). Briefly, it consists of 41 subtests that generate 55 variables, which encompass a basic spectrum of the neuropsychological functions: language, attention, mental tracking, working memory, repetition, confrontation naming, semantic fluency, verbal comprehension, reading, writing, praxis, visual perceptual functions, verbal memory (story), visual memory (figures), numerical reasoning, concept formation, sustained attention, speed, and visuospatial and motor skills. It takes only 30–45 minutes to administer.

As to psychometric characteristics, this test has shown higher convergent validity with the Alzheimer Disease Assessment Scale–cognitive part (ADAS-Cog; Peña-Casanova et al., 1997) and excellent test–retest and inter-rater reliability (Serra-Mayoral & Peña-Casanova, 2006).

Data analysis

The Statistical Package for Social Sciences (SPSS v.18) was used to investigate group differences in demographic and neuropsychological scores. Analysis of variance (ANOVA) with post hoc analysis (Bonferroni) was used to compare sociodemographic and neuropsychological data amongst groups. Gender differences were assessed by means of chi-square tests. The significance level was set at .05.

LDA with a stepwise method was performed to compare the performance of the ANNs. Software Easy NN-Plus (Easy Neural Network Plus, v.10; Neural Planner Software, 2010) was used to simulate ANNs.

ANNs

ANNs are adaptive models, inspired by the functioning processes in the human brain, for the analysis of data (Tabaton et al., 2010). In addition, a nonlinear pattern recognition system and massively parallel distributed processor make up simple processing units (Haykin, 2008).

We applied supervised ANNs whose processing result (the desired output) had already been defined. Multilayer perceptron (MLP), as described by Rumelhart and McClelland (1986), was used for the analysis. It is defined by different, interconnected layers of nodes characterized by a nonlinear function, generally of the sigmoidal type, trained by backpropagation error. It includes an additional layer between the input and output, which permits the classification of nonseparable linear patterns. In addition, it involves the modification of the learning rule to enable the teaching of hidden neurons (Pérez & Martin, 2003). This model is widely used in the nonseparable, linear classification of patterns and function prediction.

The nodes, also known as processing elements, and the connections are fundamental elements of ANNs. Each node is connected to other nodes to link communications with a numerical value, also known as weights (w). The knowledge acquired is represented in the connection weights. If w is more than zero, the connection is excitatory; if it is less than zero, it is inhibitory (Basheer & Hajmeer, 2000). In addition, it requires a rule, the so-called activation function, to combine the inputs. The sigmoid function logistic type is the most widely used activation function.

The nodes are distributed in layers: input, hidden, and output. The input layer receives the external inputs—that is, the predictor variable values. In this layer, the received signal is not usually processed but is sent instead to the following layer (Pérez & Martin, 2003). The hidden layer learns to recode the inputs. The activity of each hidden unit is determined by the activities of the input units and the weights between the hidden and output units. More than one hidden layer can be used. The last layer is the output layer; the activation levels of the neurons in this layer are considered to be the output of the neural network (Dayhoff & DeLeo, 2001).

With the objective of minimizing the function error and adjusting the weights from the input to the hidden units, the backpropagation error algorithm, known as a generalization of the delta rule, was used. Although this algorithm was initially developed by Werbos (1974), Rumelhart and McClelland (1986) claimed its popularity. The backpropagation error algorithm has two phases.

In the forward phase, the synaptic weights of the network are fixed, and the input signal is propagated through the network, layer by layer, until it reaches the output (Haykin, 2008). The error at the output unit is therefore calculated. In the backward phase, the errors between the desired output and the network are propagated layer by layer in the backward direction, and the error at the hidden nodes is calculated. This process is iterative.

It is advisable to select the input variables in order to eliminate background noise and achieve an optimum learning function. Indiscriminate use of neurons can, moreover, cause the network to memorize the training data and generalize—that is to say, it cannot give a suitable answer when new data are presented (Pérez & Martín, 2003). Taking into account these aspects, we selected input variables with the best discrimination amongst the three diagnosis groups studied. Ten a-BT subtests were included: temporal orientation, backward series, semantic fluency, immediate and delayed story memory, free and cued recall, visual memory, similarities, and digit symbol. The sociodemographic variables with the greatest influence on cognitive performance, age, and years of education were also incorporated.

Our ANN had three neuron layers: input, hidden, and output (see Figure 1). The input layer had 12 neurons: age and years of education, and a selection from a previously presented a-BT subtest. The entire distribution of the input variables was employed. The single hidden layer had four nodes. The output layer had one neuron: the diagnosis. This neural network architecture was selected by choosing an optimal number of neurons in the hidden layer. All inputs and outputs were scaled and offset to the range 0.0 to 1.0. In an ANN analysis, quantitative variables are used to predict a categorical variable—in this case, the diagnosis.

The neuron learning consists of the modification of the weight's vector, w, so that the output will be appropriate for the task. These weights can be both positive (excitatory) and negative (inhibitory). This algorithm is repetitive. The equation of the weight's modification is the following:

$$\Delta_p w_{ij} = \gamma \delta_i O_j$$

where: w_{ij} = weight between neuron i and neuron j; γ = increase in learning in every step; δ_i = neuron i error; and O_j = pattern output.

Once the patterns are presented, the weights are updated so that a cycle of learning is completed. This process attempts to minimize the following function of error:

$$E = {}_p 1/2_k (d_{pk} - x_{pk})^2$$

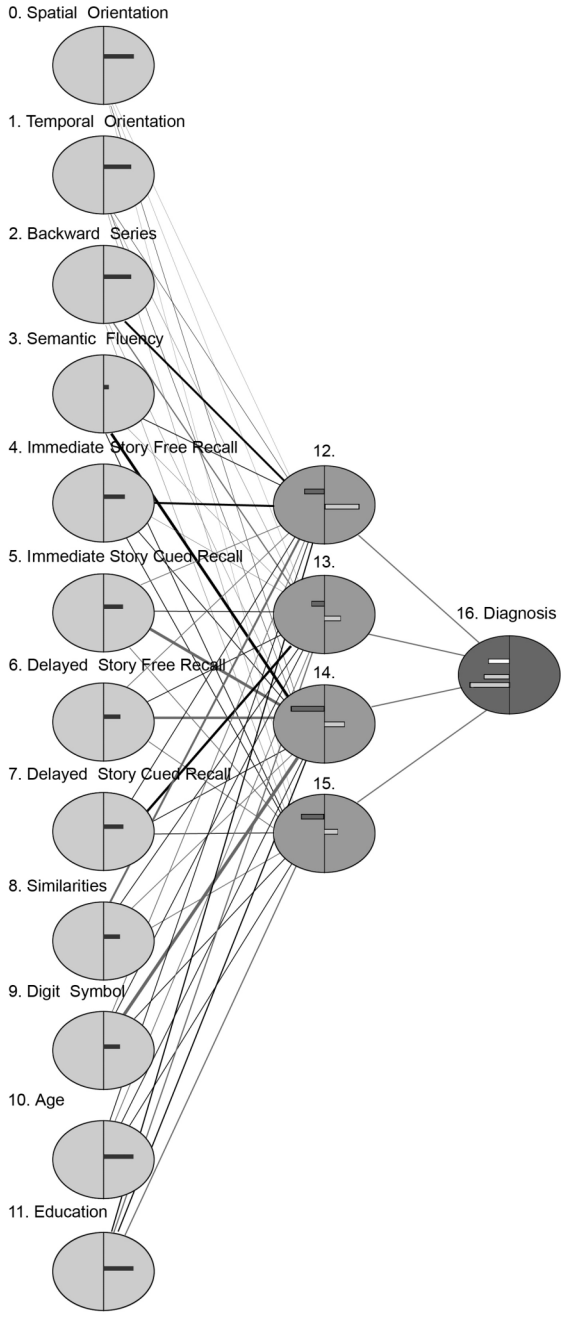

Figure 1. Graphical representation of a neural network, showing input layer (cognitive measures, age, and education), hidden layer, and output layer (the diagnosis).

where: $p1/2k$ = minimization of the error in every nodule p in the value k; d_{pk} = desired output neuron of pattern p in output neuron; and x_{pk} = real output of pattern p in output neuron.

Each neuron has a nonlinear activation function, usually a logistic one. The most popular activation function is the sigmoid:

$$f(\mathrm{s}) = \frac{1}{1 + \exp(-\mathrm{bx})}$$

The learning rate and momentum parameters were set to 0.6 and 0.8, respectively, for all analyses. Prior to training, networks were initialized by randomly setting connection weights in the range from –.5 to .5. A priori, we did not set the number of learning cycles or interaction in the training phase.

As for the hidden layer, we decided to include only one layer. Due to the problem of overadjustment, the minimum number of hidden neurons should be employed in order to obtain the network structure showing the optimum generalized performance. One hidden layer with four neurons was, therefore, selected.

The importance of each input variable was assessed during the training phrase. The input importance is a parameter expressing the magnitude of the activation of a given node during this phase. The magnitude of the activation is arbitrarily expressed by a number that ranges from zero to infinity (Grossi et al., 2007). The equation of the input importance is the following:

$$R_i = 1/K * \sum_{c}^{K} \sum_{j}^{N} w_{c,j,i}$$

where: $R_{\mathrm{i}=}$ the mean importance of the ith input variable of the dataset; K = the number of classifiers used in the training phase; N = the number of hidden units of the trained K classifiers; $w_{c,j,i}$ = the trained weight of the cth classifier, connecting the ith input to the jth hidden unit.

Another evaluated parameter was the sensitivity of each input variable. Sensitivity (S) is understood as the variation of the output variable before changes in input values. It is detailed below:

$$S_X^Q = \frac{X}{Q} \frac{\partial Q}{\partial X}$$

where: X = the matrix of input variables; Q = the vector of loading coefficients of each input variable in relation to the output variable; ∂Q = the first derivate to zero of Q; and ∂X = the first derivative to zero of X.

This general procedure was applied in three ANN substudies:

- Substudy 1: The whole sample (three diagnosis groups)
- Substudy 2: Healthy controls and MCI patients
- Substudy 3: Healthy controls and AD patients

We decided to apply ANN analysis to a more complex classification issue: the first substudy, which

included the three diagnosis groups. Furthermore, we included two common diagnostic dichotomies: healthy controls versus MCI, and healthy controls versus AD.

Validation protocol

In order to increase the probability of generalization, and to avoid the overfitting of the predictive performance, we used a standard training set for the network and a test set to validate its predictive performance. This procedure is indispensable in order to verify the model's ability to generalize the results reached in the testing phase of each model (Grossi et al., 2007).

In the training phase, the ANNs learn to associate the input variables with those that are indicated as targets (Buscema et al., 2004). The testing phase is used to assess the predictive ability of the network for generations of future practical applications.

In Substudy 1, the total sample of 522 subjects was randomly subdivided into two subsamples: the training phase ($n = 432$) and the prediction phase (testing; $n = 90$ subjects, 30 from each group: controls, MCI, and AD, respectively). In Substudy 2, we excluded the AD patients. This sample included 425 subjects in the training phase and 60 subjects (30 healthy controls and 30 MCI patients) in the testing phase. Finally, in Substudy 3, the training subsample had 443 subjects and the testing subsample 60 subjects (30 controls and 30 AD patients).

RESULTS

Sample characteristics

Demographic data, MMSE, IDDD, and a-BT scores are given in Table 1. Age and years of education were unequally distributed across groups. Participants in the control group were significantly younger and more highly educated than MCI and AD patients. However, in post hoc analysis (Bonferroni) these patients did not differ in age and education. The slight difference between groups in the gender balance was nonsignificant. Patients' MMSE scores were, as assumed from diagnosis of MCI and AD, lower than those for the control group. The scores in the functional scale IDDD were distributed according to an expected pattern—that is to say, with functional normality in the control group, whereas there were higher scores in MCI patients, and AD patients presented functional impairment. All a-BT measures varied significantly ($p < .0001$) in the ANOVA analysis.

ANN analysis

In this study, three substudies of ANNs were carried out in order to distinguish healthy controls, MCI, and AD patients. In the first model, which included the three diagnosis groups, ANNs correctly classified 66.67% of subjects (average error 0.02). In addition to evaluating the overall performance rate in discriminating between patients, the ANN models permitted the definitions of the

TABLE 1
Demographic characteristics and functional and neuropsychological scores for the three groups studied

	Control (n = 346)	*MCI (n = 79)*	*AD (n = 97)*	*Sign.*
Age, years	65.04 ± 9.38	72.82 ± 6.53	74.69 ± 7.49	<.0001
Female gender, %	60	57	65	*ns*
Education, years	10.56 ± 5.46	8.03 ± 4.76	7.49 ± 4.55	<.0001
MMSE	28.80 ± 1.50	25.77 ± 2.22	20.37 ± 3.98	<.0001
IDDD	33.19 ± 0.61	36.01 ± 2.56	48.87 ± 10.12	<.0001
Temporal orientation	22.93 ± 0.40	21.30 ± 3.61	13.29 ± 7.14	<.0001
Backward series	2.92 ± 0.27	2.47 ± 0.68	1.67 ± 1.10	<.0001
Semantic fluency	19.73 ± 5.76	13.58 ± 4.83	9.73 ± 4.07	<.0001
Immediate story memory: free recall	12.98 ± 3.72	7.34 ± 3.31	4.68 ± 2.95	<.0001
Immediate story memory: cued recall	16.62 ± 3.33	11.61 ± 3.62	8.14 ± 3.85	<.0001
Delayed story memory: free recall	13.01 ± 4.55	5.43 ± 4.34	2.00 ± 3.16	<.0001
Delayed story memory: cued recall	16.41 ± 3.71	9.72 ± 4.58	5.15 ± 4.62	<.0001
Visual memory	10.09 ± 3.64	4.48 ± 2.55	2.03 ± 2.49	<.0001
Similarities	8.17 ± 2.35	6.05 ± 2.07	4.65 ± 2.79	<.0001
Digit symbol	23.20 ± 10.36	12.57 ± 6.10	7.18 ± 5.49	<.0001

Note. Data are presented in the form: mean ± standard deviation. MCI = mild cognitive impairment; AD = Alzheimer's disease; MMSE = Mini-Mental State Examination; IDDD = Interview For Deterioration In Daily Living Scale; Sign. = significance. Group differences were performed by analysis of variance (ANOVA) with post hoc analysis (Bonferroni). Gender percentage comparison was performed by means chi-square test with continuity correction.

TABLE 2
The importance value for each input variable

	Substudy		
	1 (Whole sample)	*2 (Controls & MCI)*	*3 (Controls & AD)*
Temporal orientation	16.6842	59.2188	17.4212
Backward series	39.6352	89.9924	15.8443
Semantic fluency	58.3462[a]	54.8496	17.9992
Immediate story memory: free recall	24.2160	81.1996	16.7558
Immediate story memory: cued recall	34.4266	82.7592	7.3857
Delayed story memory: free recall	13.7516	72.0599	14.5154
Delayed story memory: cued recall	35.7109	149.2006[a]	2.5344
Visual memory	21.3789	128.1294	19.8522[a]
Similarities	25.5158	75.9835	11.5521
Digit symbol	46.7522	87.3062	8.5909
Age	35.0616	51.3258	7.08
Education	30.3801	91.1110	16.8025

Note. MCI = mild cognitive impairment; AD = Alzheimer's disease.
[a]The most important input in ANN (artificial neural networks) model.

net importance of each input in the occurring model. Table 2 summarizes the importance for each input variable represented by the addition of the weight's absolute values of each input variable toward each neuron in the intermediate layer. As can be observed, semantic fluency, followed by digit symbol, accounted for the highest input importance.

The ANN model also contributes to the sensitivity value of each independent variable, which is represented as the variation of the output variable before changes in the input values (see Table 3). Immediate story memory: cued recall showed the most sensitive input variable within the ANN model.

In the second substudy, which included healthy controls and MCI patients, the ANN model showed a generalized mean of 98.33%, and only 1 subject was misclassified. Delayed story memory: cued recall was the most important input variable. In sensitivity, three input variables presented maximum values: backward series, delayed story memory cued recall, and visual memory.

Finally, in the third substudy, the ANN model correctly classified all subjects ($n = 60$, 100%). The input variables in this model showed different degrees of importance and sensitivity (see Tables 2 and 3). The highest input of importance and sensitivity was reported in visual memory.

Linear discriminant analysis

The results obtained by ANN analysis were compared with those from LDA using the three substudies. Variables within the linear discriminant function in each case are given in Table 4. The results from the LDA classifier comparing the ANN results are shown in Table 5. ANNs showed better predictive performance than LDA in Substudy 2 (controls and MCI) and Substudy 3 (controls and AD) but not in Substudy 1 (controls, MCI, and AD).

In relation to subjects wrongly classified under one or the other technique, the results derived from the individual classifications were analyzed in order to evaluate the presence of a specific pattern of coincidences. Only 12% of the subjects were wrongly classified with both techniques. So, this result suggests the almost absolute independence between both criteria of classification.

DISCUSSION

The purpose of this report was to develop, train, and explore the ability of artificial neural networks to differentiate healthy controls, mild cognitive impairment subjects, and Alzheimer disease patients. In addition, the ANNs were employed as a viable model to study the relevant variables in MCI and AD diagnosis.

Our results support the utility of ANNs as a new tool in the interpretation of data from heterogeneous and distinct sources, with an overall predictive accuracy ranging from 66.67% to 100%. Our results showed that ANN analysis was more efficient than conventional statistical analysis such as LDA, but not in all the studied situations, since when three groups were present it was the LDA that obtained a better result. The first ANN tackled the most difficult classification problem due

TABLE 3
The sensitivity value for each input variable

Substudy	*Input variable*	*Value*
1: Whole sample	Immediate story memory: cued recall	.90227
	Digit-symbol	.89658
	Semantic fluency	.89589
	Backward series	.89419
	Similarities	.89386
	Education	.89061
	Temporal orientation	.87803
	Immediate story memory: free recall	.83096
	Delayed story memory: free recall	.80967
	Delayed story memory: cued recall	.61686
	Age	.61134
	Visual memory	.40397
2: Controls & MCI	Visual memory	1.00000
	Backward series	1.00000
	Delayed story memory: cued recall	1.00000
	Immediate story memory: free recall	.99999
	Delayed story memory: free recall	.99999
	Age	.99896
	Digit-symbol	.21181
	Immediate story memory: cued recall	.02487
	Education	.00138
	Similarities	.00001
	Semantic fluency	.00001
	Temporal orientation	.00001
3: Controls & AD	Visual memory	.87715
	Education	.83429
	Semantic fluency	.81618
	Immediate story memory: free recall	.81072
	Temporal orientation	.79918
	Delayed story memory: free recall	.76317
	Backward series	.66772
	Digit symbol	.31101
	Immediate story memory: cued recall	.23355
	Age	.02053
	Delayed story memory: cued recall	.01758
	Similarities	.01556

Note. MCI = mild cognitive impairment; AD = Alzheimer's disease.

TABLE 4
Variables into the linear discriminant function

	Substudy 1		*Substudy 2*		*Substudy 3*	
Step	*Input variable*	*Wilks's lambda*	*Input variable*	*Wilks's lambda*	*Input variable*	*Wilks's lambda*
1	Delayed story memory: cued recall	.44*	Delayed story memory: cued recall	.69*	Temporal orientation	.41*
2	Temporal orientation	.31*	Backward series	.62*	Delayed story memory: cued recall	.31*
3	Backward series	.28*	Visual memory	.59*	Backward series	.28*
4	Visual memory	.26*	Education	.57*	Education	.26*
5	Education	.24*	Temporal orientation	.56*	Visual memory	.25*
6	Semantic fluency	.24*			Semantic fluency	.24*
7	Delayed story memory: free recall	.23*				

*$p < .0001$.

TABLE 5
Comparison of predictive performance of linear discriminant analysis and artificial neural networks

Substudy	*Sample included*	*LDA (%)*	*ANN (%)*
1	Controls, MCI, and AD	79.9	66.67
2	Controls and MCI	80	98.33
3	Controls and AD	96.4	100

Note. Percentage of subjects correctly classified. LDA = linear discriminant analysis; ANN = artificial neural networks; MCI = mild cognitive impairment; AD = Alzheimer's disease.

to the fact that all three diagnosis groups were included. Nevertheless, the percentage of correct classification was high (66.67%). This result could be explained by the model's difficulty in separating the intermediate group, the MCI patients. In fact, the same effect is obtained applying LDA. The common characteristics of MCI patients and healthy controls (such as the conservation of fixed cognitive functions), and MCI patients and AD patients who had impaired episodic memory, might have been the cause. The result is, however, consistent with the heterogenic concept of MCI (DeCarli, 2003; Petersen et al., 2001), which overlaps in a similar manner with other examples of clinical diagnosis, such as normal aging and AD. In fact, when we excluded the AD group (Substudy 2), the percentage of correctly classified subjects increased considerably (98.33%). And, finally, when we only included healthy controls and AD patients, the ANN model classified all subjects with an error average of 2%. Consequently, through the results of the input variables included in ANN, we could predict diagnosis with an accuracy that varied according to the complexity of decision. Only in the first substudy was the minimal classification of LDA better than that of ANN (79.9% vs. 66.67%, respectively). In this case, LDA showed similar difficulties in separating the intermediate group, the MCI patients.

With respect to importance and sensitivity within the ANN model, in Substudy 1 semantic fluency was the most significant factor for patient classification. Impaired semantic fluency has been reported as a sign in MCI and AD disease (Murphy, Rich, & Troyer, 2006; Stokholm, Vogel, Gade, & Waldemar, 2006). Multiple cognitive processes are involved in the execution of a semantic fluency task—these include semantic memory, phonological process, executive function, and general speed of processing (Adlam, Bozeat, Arnold, Watson, & Hodges, 2006; Moreno-Martínez, Laws, & Schulz, 2008; Radanovic et al., 2009; Raoux et al., 2008). On the other hand, the delayed story memory (free recall) was the most important input in the second ANN model for both healthy controls and MCI patients. Both free recall and cued recall have been shown to be sensitive measures for MCI diagnosis (Dubois & Albert, 2004; Ritchie, Artero, & Touchon, 2001) and predictors of conversion from MCI to AD (Artero, Tierney, Touchon, & Ritchie, 2003; Dierckx et al., 2009). Finally, visual memory was the most important variable for the classification of healthy controls and AD patients.

In addition, the sociodemographic variable, years of education, showed high importance when we included two groups (Substudies 1 and 2). A large quantity of evidence indicates that education, as a variable of cognitive reserve, protects against the appearance of cognitive symptoms of aging and AD (Rentz et al., 2010; see Stern, 2002, for review). In the case of age, the effect is the contrary.

With respect to sensitivity, immediate story memory (free recall) was the most sensitive measure within the ANN in the first substudy. In the second model, backward series as a measure of working memory, delayed story memory (free and cued recall), and visual memory showed the highest values. Episodic memory assessment, through list learning test and/or story memory task (Rabin et al., 2009) is essential due to the fact that impairment in episodic memory is a typical sign of prodromal dementia—for example, amnestic MCI. In the latter case, the sensitivity values showed that the majority of these variables fit impaired cognitive functions in AD.

In this study, no direct comparison was made between the results obtained by ANN and LDA because both logical and technical models belong to different analysis, and therefore it would be more complex to develop here a model of contrast between the two results. The LDA is a linear model structure based on the concept of the general linear model, whereas ANN, which is based, as quoted in the paper, on models of transition, is not necessarily linear and is more like the logic of procedures such as Markov chains, with no single pattern from which to fit.

In situations involving clinical decisions with a large number of predictive variables and, as a consequence, with a high degree of complexity in reaching an adequate diagnostic decision, the application of neural networks may represent a potential solution. ANN models have flexibility, in a wide variety of areas, to perform with significant diagnostic accuracy. They resemble the brain in several aspects (Haykin, 2008). On the one hand, they acquire knowledge through a learning process, where the intensity of the interneuron connections is used to store knowledge (Sánchez & Analis, 2006). On the other hand, ANNs can recognize complex patterns, manage data, learn the hidden relationship among different variables, and resolve classification problems or prediction of results (Buscema et al., 2004).

Neural networks present a series of advantages. First, they are nonlinear systems. As mentioned above, AD is a good model to explore the usefulness of ANNs as it is a slow-paced progressive disease in which cognitive deterioration follows a nonlinear course and, as with aging, its course is curvilinear (Doody et al., 2001; Embretson & Reise, 2000; Mungas & Reed, 2000; Tandon et al., 2006). A priori, ANNs do not require initial principles or data restrictions, such as distribution assumptions (e.g. normal distribution; Sargent, 2001). They are parallel-distributed systems that are flexible and failure tolerant (Lisboa & Taktak, 2006). Furthermore, an ANN is capable of "learning" and, as a result, the network is able to transform entry and exit data. ANNs can also work with imprecise data, given their capacity to learn and generalize. A further advantage is their adaptability (i.e., their capacity to modify their parameters, even in real time), and, as adaptable systems, they are able to solve complex problems through the combination of multiple factors or simultaneous variables (Di Luca et al., 2005). Moreover, as distributed systems, ANNs allow for the failure of some neurons without significantly altering the overall response of the system and can, therefore, be considered as failure-tolerant systems. They can work with incomplete, noisy, and inconsistent information and provide uniformity in analysis and design through joint theories that describe the different algorithms and applications. ANNs are also analogous to biological networks (Haykin, 2008; Ramesh et al., 2004).

In our study, ANNs were more efficient in their ability to correctly classify MCI and AD patients and separating them from control subjects than was LDA. Several previous studies in ANNs have compared the efficiency of these models with other statistical approaches such as logistic regression or LDA, and ANNs have demonstrated superior performance (Buscema et al., 2004; Di Luca et al., 2005; French et al., 1997; Grossi et al., 2007; Leighty et al., 2008; Page et al., 1996). Despite these arguments, in the field of health sciences, classical statistical methods are still preferred, for reasons that include their relative simplicity and the wide availability of statistical programs (Sargent, 2001).

With respect to the application of our results, it is particularly important to bear in mind that the neural-network models used do not facilitate diagnosis. In other words, the diagnosis is already known, as we are using a multilayer perceptron based on a supervised-learning paradigm, and, as a result, the network learns entry and exit patterns. Therefore, ANNs should be considered instruments for pattern classification and function prediction that can help in the taking of clinical decisions; they are not diagnostic tools.

Our study has some limitations. First, ANNs are considered as "black boxes" to identify relationships, and it is difficult to understand the nature of the internal representations generated by an extremely complex network that transforms inputs into predetermined outputs (Montaño & Palmer, 2003). Moreover, in spite of their wide application in different fields, their methodology is little known in comparison to other statistical techniques (Sargent, 2001). As regards ANN implementation, the empirical processes present problems at a methodological level that are not yet fully solved. Second, whilst we used the MCI IPA-WHO criteria, we did not classify them into MCI subtypes. In fact, we assumed that all patients could be amnestic or multidomain amnestic.

In conclusion, the ANNs were applied to groups of healthy controls and MCI and AD patients to predict diagnosis. We obtained excellent classificatory power and more efficient classifier sensitivity than LDA. These results support previous findings that reported ANNs to be a versatile tool capable of separating nonlinear associations—for example, cognition and aging. Furthermore, the use of ANNs is a novel methodology in neuropsychology. Finally, this approach, as an innovative and powerful modeling tool, could be increasingly applied to develop predictive models in neuroscience and should prove useful for predicting cognitive impairment.

Original manuscript received 16 June 2011
Revised manuscript accepted 21 September 2011
First published online 13 December 2011

REFERENCES

Adlam, A. L., Bozeat, S., Arnold, R., Watson, P., & Hodges, J. R. (2006). Semantic knowledge in mild cognitive impairment and mild Alzheimer's disease. *Cortex*, *42*, 675–684.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Artero, S, Tierney, M. C., Touchon, J., & Ritchie, K. (2003). Prediction of transition from cognitive impairment to senile dementia: A prospective, longitudinal study. *Acta Psychiatrica Scandinavica*, *107*, 390–393.

Basheer, I. A., & Hajmeer, M. (2000). Artificial neural networks: Fundamentals, computing, design, and application. *Journal of Microbiological Methods*, *43*, 3–31.

Batx, W. (1995). Application of artificial neural networks to clinical medicine. *Lancet*, *346*, 1135–1138.

Blesa, R., Pujol, M., Aguilar, M., Santacruz, P., Bertran-Serra, I., Hernandez, G., et al. (2001). NORMAlisation of cognitive and functional instruments for DEMentia. Clinical validity of the "mini-mental state" for Spanish speaking communities. *Neuropsychologia*, *39*, 1150–1157.

Böhm, P., Peña-Casanova, J., Aguilar, M., Hernández, G., Sol, J. M., & Blesa, R. (1998). Clinical validity and utility of the Interview for Deterioration of Daily Living in Dementia for Spanish-speaking communities. *International Psychogeriatrics*, *10*, 261–270.

Buscema, M., Grossa, E., Snowdon, D., Antuono, P., Intraligi, M., Maurelli, G., et al. (2004) Artificial neural networks and artificial organisms can predict Alzheimer pathology in individual patients only on the basis of cognitive and functional status. *Neuroinformatics*, *2*, 399–416.

Buscema, M., Rossini, P., Babiloni, C., & Grossi, E. (2007). The IFAST model, a novel parallel nonlinear EEG analysis technique, distinguishes mild cognitive impairment and Alzheimer's disease patients with high degree of accuracy. *Artificial Intelligence in Medicine*, *40*, 127–141.

Busse, A., Hensel, A., Gühne, U., Angermeyer, M. C., & Riedel-Heller, S. G. (2006). Mild cognitive impairment: Long-term course for four clinical subtypes. *Neurology*, *67*, 2176–2185.

Dayhoff, J. E., & DeLeo, J. M. (2001). Artificial neural networks. Opening the black box. *Cancer*, *91*, 1615–1635.

DeCarli, C. (2003). Mild cognitive impairment: Prevalence, prognosis, aetiology, and treatment. *Lancet Neurology*, *2*, 15–21.

DeFigueiredo, R. J., Shankle, W. R., Maccato, A., Dick, M. B., Mundkur, P., Mena, I., et al. (1995). Neural-network-based classification of cognitively normal, demented, Alzheimer disease and vascular dementia from single photon emission with computed tomography image data from brain. *Proceedings of the National Academy of Sciences of the United States of America*, *92*, 5530–5534.

Dierckx, E., Engelborghs, S., De Raedt, R., Van Buggenhout, M., De Deyn, P. P., Verté, D., et al. (2009). Verbal cued recall as a predictor of conversion to Alzheimer's disease in mild cognitive impairment. *International Journal of Geriatric Psychiatry*, *24*, 1094–1100.

Di Luca, M., Grossi, E., Borroni, B., Zimmermann, M., Marcello, E., Colciaghi, F., et al. (2005). Artificial neural networks allow the use of simultaneous measurements of Alzheimer disease markers for early detection of the disease. *Journal of Translational Medicine*, *27*, 3–30.

Doody, R. S., Massman, P., & Dunn, J. K. (2001). A method for estimating progression rates in Alzheimer disease. *Archives of Neurology*, *58*, 449–454.

Dubois, B., & Albert, M. L. (2004). Amnestic MCI or prodromal Alzheimer's disease? *Lancet Neurology*, *3*, 246–248.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates.

Ewers, M., Buerger, K., Teipel, S. J., Scheltens, P., Schröder, J., Zinkowski, R. P., et al. (2007). Multicenter assessment of CSF-phosphorylated tau for the prediction of conversion of MCI. *Neurology*, *69*, 2205–2212.

Fleisher, A. S., Sowell, B. B., Taylor, C., Gamst, A. C., Petersen, R. C., Thal, L. J., et al. (2007). Clinical predictors of progression to Alzheimer disease in amnestic mild cognitive impairment. *Neurology*, *68*, 1588–1595.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198.

French, B. M., Dawson, M. R., & Dobbs, A. R. (1997). Classification and staging of dementia of the Alzheimer type: A comparison between neural networks and linear discriminant analysis. *Archives of Neurology*, *54*, 1001–1009.

Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R. C., Ritchie, K., Broich, K., et al. (2006), Mild cognitive impairment. *Lancet*, *367*, 1262–1270.

Grossi, E., Buscema, M., Snowdon, D., & Antuono, P. (2007). Neuropathological findings processed by artificial neural networks (ANNs) can perfectly distinguish Alzheimer's patients from controls in the Nun Study. *BioMed Central Neurology*, *21*, 7–15.

Grossi, E., Massini, G., Buscema, M., Savarè, R., & Maurelli, G. (2005). Two different Alzheimer diseases in men and women: Clues from advanced neural networks and artificial intelligence. *Gender Medicine*, *2*, 106–117.

Haykin, S. (2008). *Neural networks and learning machine* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Hebb, D. O. (1949). *The organization of behaviour: A neuropsychological theory*. New York, NY: Wiley.

Hebert, L. E., Beckett, L. A., Scherr, P. A., & Evans, D. A. (2001). Annual incidence of Alzheimer disease in the United States projected to the years 2000 through 2050. *Alzheimer Disease and Associated Disorders*, *15*, 169–173.

Horn, J. F., Habert, M. O., Kas, A., Malek, Z., Maksud, P., Lacomblez, L., et al. (2009). Differential automatic diagnosis between Alzheimer's disease and frontotemporal dementia based on perfusion SPECT images. *Artificial Intelligence in Medicine*, *47*, 147–158.

Jack, C. R., Weigand, S. D., Shiung, M. M., Przybelski, S. A., O'Brien, P. C., Gunter, J. L., et al. (2008). Atrophy rates accelerate in amnestic mild cognitive impairment. *Neurology*, *70*, 1740–1742.

Landau, S. M., Harvey, D., Madison, C. M., Reiman, E. M., Foster, N. L., Aisen, P. S., et al. (2010). Comparing predictors of conversion and decline in mild cognitive impairment. *Neurology*, *75*, 230–238.

Lehmann, C., Koening, T., Jelic, V., Prichep, L., John, R. E., Wahlund, L., et al. (2007). Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG). *Journal of Neuroscience Methods*, *161*, 342–350.

Leighty, R. E., Runfeldt, M. J., Berndt, D. J., Schleif, W. S., Cracchiolo, J. R., Potter, H., et al. (2008). Use of artificial neural networks to determine cognitive impairment and therapeutic effectiveness in Alzheimer's transgenic mice. *Journal of Neuroscience Methods*, *167*, 358–366.

Levy, R. (1994). Aging-associated cognitive decline. Working Party of the International Psychogeriatric Association in collaboration with the World Health Organization. *International Psychogeriatrics*, *6*, 63–68.

Lisboa, P. J., & Taktak, A. F. G. (2006). The use of artificial neural networks in decision support in cancer: A systematic review. *Neural Networks*, *19*, 408–415.

Litvan, I., DeLeo, J. M., Hauw, J. J., Daniel, S. E., Jellinger, K., McKee, A., et al. (1996). What can artificial neural networks teach us about neurodegenerative disorders with extrapyramidal features? *Brain*, *119*, 831–839.

Manly, J. J., Tang, M. X., Schupf, N., Stern, Y., Vonsattel, J. P., & Mayeux, R. (2008). Frequency and course of mild cognitive impairment in a multiethnic community. *Annals of Neurology*, *63*, 494–506.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA work group under the auspices of the department of health and human services task force on Alzheimer's disease. *Neurology*, *34*, 939–945.

Mecocci, P., Grossi, E., Buscema, M., Intraligi, M., Savarè, R., Rinaldi, P., et al. (2002). Use of artificial networks in clinical trials: A pilot study to predict responsiveness to donepezil in Alzheimer's disease. *Journal of American Geriatrics Society*, *50*, 1857–1860.

Montaño, J. J., & Palmer, A. (2003). Numeric sensitivity analysis applied to feedforward neural networks. *Neural Computing & Applications*, *12*, 119–125.

Moreno-Martínez, F. J., Laws, K. R., & Schulz, J. (2008). The impact of dementia, age and sex on category fluency: Greater deficits in women with Alzheimer's disease. *Cortex*, *44*, 1256–1264.

Mungas, D., & Reed, B. R. (2000). Application of item response theory for development of a global functioning measure of dementia with linear measurement properties. *Statistics in Medicine*, *19*, 1631–1644.

Murphy, K. J., Rich, J. B., & Troyer, A. K. (2006). Verbal fluency patterns in amnestic mild cognitive impairment are characteristic of Alzheimer's type dementia. *Journal of the International Neuropsychological Society*, *12*, 570–574.

Neural Planner Software. (2010). *Easy Neural Network Plus* (Version 10). Retrieved from http://www.easynn.com

Okello, A., Koivunen, J., Edison, P., Archer, H. A., Turkheimer, F. E., Nagren, K., et al. (2009). Conversion of amyloid positive and negative MCI to AD over 3 years. *Neurology*, *73*, 754–760.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, *9*, 97–113.

Page, M. P., Howard, R. J., O'Brien, J. T., Buxton-Thomas, M. S., & Pickering, A. D. (1996). Use of neural networks in brain SPECT to diagnose Alzheimer's disease. *Journal of Nuclear Medicine*, *37*, 195–200.

Parsons, T., Rizzo, A. A., & Buckwalter, G. (2004). Backpropagation and regression: Comparative utility for neuropsychologist. *Journal of Clinical and Experimental Neuropsychology*, *26*, 95–104.

Peña-Casanova, J. (1990). *Programa integrado en la exploración neuropsicológica—Test Barcelona. Manual* [Integrated program of neuropsychological assessment—Revised Barcelona Test. Manual]. Barcelona, Spain: Masson.

Peña-Casanova, J., Blesa, R., Aguilar, M., Gramunt-Fombuena, N., Gómez-Ansón, B., Oliva, R., et al. (2009). Spanish multicenter normative studies (NEURONORMA project): Methods and sample characteristics. *Archives of Clinical Neuropsychology*, *24*, 307–319.

Peña-Casanova, J., Meza, M., Bertran-Serra, I., Manero, R. M., Espel, G., Martí, A., et al. (1997). Versión abreviada del test Barcelona (III): Validez de criterio con el ADAS-Cog [Shortened version of the Barcelona Test (III): Criterion validity with the ADAS-Cog]. *Neurología*, *12*, 117–119.

Pérez, M. L., & Martín, Q. (2003). *Aplicaciones de las redes neuronales artificiales a la estadística* [Applications of the artificial neural networks to statistics]. Madrid, Spain: Editorial La Murralla.

Petersen, R. C. (2003). Conceptual overview. In R. C. Petersen (Ed.), *Mild cognitive impairment: Aging to Alzheimer's disease* (pp. 1–14). New York, NY: Oxford University Press.

Petersen, R. C., Doody, R., Kurz, A., Mohs, R. C., Morris, J. C., Rabins, P. V., et al. (2001). Current concepts in mild cognitive impairment. *Archives of Neurology*, *58*, 1985–1992.

Quintana, M., Peña-Casanova, J., Sánchez-Benavides, G., Langohr, K., Manero, R. M., Aguilar, M., et al. (2011). Norms for the abbreviated Barcelona Test. *Archives of Clinical Neuropsychology, 26*, 144–157.

Rabin, L. A., Paré, N., Saykin, A. J., Brown, M. J., Wishart, H. A., Flashman, L. A., et al. (2009). Differential memory test sensitivity for diagnosing amnestic mild cognitive impairment and predicting conversion to Alzheimer's disease. *Aging, Neuropsychology, and Cognition*, *16*, 357–376.

Radanovic, M., Diniz, B. S., Mirandez, R. M., Novaretti, T. M., Flacks, M. K., Yassuda, M. S., et al. (2009). Verbal fluency in the detection of mild cognitive impairment and Alzheimer's disease among Brazilian Portuguese speakers: The influence of education. *International Psychogeriatrics*, *21*, 1081–1087.

Ramesh, A. N., Kambhampati, C., Monson, J. R. T., & Drew, P. J. (2004). Artificial intelligence in medicine. *Annals of The Royal College of Surgeons of England*, *86*, 334–338.

Raoux, N., Amieva, H., Le Goff, M., Auriacombe, S., Carcaillon, L., Letenneur, L., et al. (2008). Clustering and switching processes in semantic verbal fluency in the course of Alzheimer's disease subjects: Results from the PAQUID longitudinal study. *Cortex*, *44*, 1188–1196.

Reisberg, B., Ferris, S. H., de Leon, M. J., & Crook, T. (1972). The global deterioration scale for assessment of primary degenerative dementia. *The American Journal of Psychiatry*, *139*, 1136–1139.

Rentz, D. M., Locascio, J. J., Becker, J. A., Moran, E. K., Eng, E., Buckner, R. L., et al. (2010). Cognition, reserve, and amyloid deposition in normal aging. *Annals of Neurology*, *67*, 353–364.

Risacher, S. L., Saykin, A. J., West, J. D., Shen, L., Firpi, H. A., McDonald, B. C., et al. (2009). Baseline MRI predictors of conversion from MCI to probable AD in the ADNI cohort. *Current Alzheimer Research*, *6*, 347–361.

Ritchie, K., Artero, S., & Touchon, J. (2001). Classification criteria for mild cognitive impairment: A population-based validation study. *Neurology*, *56*, 37–42.

Rosen, W. G., Terry, R. D., Fuld, P., Katzman, R., & Peck, A. (1980). Pathological verification of ischemia score in differentiation of dementias. *Annals of Neurology*, *7*, 486–488.

Rozzini, L., Chilovi, B. V., Conti, M., Bertoletti, E., Delrio, I., Trabucchi, M., et al. (2007). Conversion of amnestic mild cognitive impairment to dementia of Alzheimer type is independent to memory deterioration. *International Journal of Psychiatry*, *22*, 1217–1222.

Rumelhart, D. E., & McClelland, J. (1986). *Parallel distributed processing* (Vol. 1). Cambridge, MA: Massachusetts Institute of Technology Press.

Sánchez, E., & Analís, A. (2006). *Redes neuronales, automática y robótica* [Neural networks, automatic and robotics]. Madrid, Spain: Pearson-Prentice.

Sargent, D. J. (2001). Comparison of artificial neural networks with other statistical approaches. *Cancer*, *91*, 1636–1642.

Serra-Mayoral, A., & Peña-Casanova, J. (2006). Fiabilidad test–retest e interevaluador del Test Barcelona [Test–retest and interrater reliability of Barcelona Test]. *Neurología*, *21*, 277–281.

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*, *8*, 448–460.

Stokholm, J., Vogel, A., Gade, A., & Waldemar, G. (2006). Heterogeneity in executive impairment in patients with very mild Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, *22*, 54–59.

Tabaton, M., Odetti, P., Cammarata, S., Borghi, R., Monacelli, F., Caltagirone, C., et al. (2010). Artificial neural networks identify the predictive values of risk factors on the conversion of amnestic mild cognitive impairment. *Journal of Alzheimer's Disease*, *19*, 1035–1040.

Tabert, M. H., Manly, J. J., Liu, X., Pelton, G. H., Rosenblum, S., Jacobs, M., et al. (2006). Neuropsychological prediction of conversion to Alzheimer disease in patients with mild cognitive impairment. *Archives of General Psychiatry*, *63*, 916–924.

Tandon, R., Adak, S., & Kaye, J. A. (2006). Neural networks for longitudinal studies in Alzheimer's disease. *Artificial Intelligence in Medicine*, *36*, 245–255.

Teunisse, S., Derix, M. M., & Cléber, H. (1991). Assessing the severity of dementia. Patient and caregiver. *Archives of Neurology*, *48*, 274–277.

Visser, P. J., Kester, A., Jolles, J., & Verhey, F. (2006). Ten-year risk of dementia in subjects with mild cognitive impairment. *Neurology*, *10*, 1201–1207.

Werbos, P. J. (1974). *Beyond regression: New tools for prediction and analysis in the behavioral sciences* (Unpublished PhD thesis). Harvard University, Cambridge, MA.