Title: Classification accuracy of TMS for the diagnosis of neurodegenerative dementias

Running head: TMS for the diagnosis of dementia

Alberto Benussi[1*], Mario Grassi[2*], Fernando Palluzzi[2*], Giacomo Koch[3,4], Vincenzo Di Lazzaro[5], Raffaele Nardone[6,7], Valentina Cantoni[1], Valentina Dell'Era[1], Enrico Premi[1], Alessandro Martorana[3,8], Francesco di Lorenzo[3], Sonia Bonnì[3], Federico Ranieri[9], Fioravante Capone[5], Gabriella Musumeci[5], Maria Sofia Cotelli[10], Alessandro Padovani[1], Barbara Borroni[1]

*these authors equally contributed to the present work

[1]*Department of Medical and Experimental Sciences, Centre for Neurodegenerative Disorders, Neurology Unit, University of Brescia, Italy;*
[2]*Department of Brain and Behavioural Sciences, Medical and Genomic Statistics Unit, University of Pavia, Pavia, Italy*
[3]*Non Invasive Brain Stimulation Unit, IRCCS Santa Lucia Foundation, Rome, Italy*
[4]*Stroke Unit, Policlinico Tor Vergata, Rome, Italy*
[5]*Unit of Neurology, Neurophysiology, Neurobiology, Department of Medicine, Campus Bio-Medico University, Rome, Italy*
[6]*Department of Neurology, Franz Tappeiner Hospital, Merano, Italy;*
[7]*Department of Neurology, Christian Doppler Klinik, Paracelsus Medical University, Salzburg, Austria*
[8]*Neurology Unit, Department of System Medicine, University of Tor Vergata, Rome, Italy*
[9]*Department of Neuroscience, Biomedicine and Movement Sciences, University of Verona, Verona, Italy*
[10]*Neurology Unit, Vallecamonica Hospital, Esine, Brescia, Italy*

***Corresponding author**:
Barbara Borroni, MD

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ana.25677

Clinica Neurologica, Università degli Studi di Brescia
P.le Spedali Civili 1, 25123, Brescia, Italy
Phone: 0039 0303995632
Email: bborroni@inwind.it

Search terms: Alzheimer's disease; frontotemporal dementia; dementia with Lewy bodies; transcranial magnetic stimulation; diagnostic accuracy; decision tree; short interval intracortical inhibition; intracortical facilitation; short latency afferent inhibition.

Abstract

Objective: Transcranial Magnetic Stimulation (TMS) has been suggested as a reliable, non-invasive, and inexpensive tool for the diagnosis of neurodegenerative dementias. In this study we assessed the classification performance of TMS parameters in the differential diagnosis of common neurodegenerative disorders, including Alzheimer's disease (AD), dementia with Lewy bodies (DLB) and frontotemporal dementia (FTD).

Methods: We performed a multicenter study enrolling patients referred to four dementia centers in Italy, in accordance with the Standards for Reporting of Diagnostic Accuracy. All patients underwent TMS assessment at recruitment (index test), with application of reference clinical criteria, to predict different neurodegenerative disorders. The investigators who performed the index test were masked to the results of the reference test and all other investigations.

We trained and tested a Random Forests classifier using 5-fold cross validation. The primary outcome measures were the classification accuracy, precision, recall and F1-score of TMS in differentiating each neurodegenerative disorder.

Results: 694 participants were included, namely 273 patients diagnosed as AD, 67 as DLB, 207 as FTD, and 147 as healthy controls (HC). A series of 3 binary classifiers was employed, and the prediction model exhibited high classification accuracy (ranging from 0.89 to 0.92), high precision (0.86-0.92), high recall (0.93-0.98), and high F1 scores (0.89-0.95), in differentiating each neurodegenerative disorder.

Interpretation: TMS is a non-invasive procedure which reliably and selectively distinguishes AD, DLB, FTD and HC, representing a useful additional screening tool to be used in clinical practice.

Introduction

Alzheimer's disease (AD) and other dementias are a major and increasing global health challenge worldwide, with 40–50 million people currently living with dementia,[1] the majority of individuals coming from low- and middle-income countries.[2] Even though AD is the most common form,[3] recent epidemiological studies[4] and the refinement of new clinical criteria[5–7] have clearly shown that frontotemporal dementia (FTD) and dementia with Lewy bodies (DLB) are much more frequent than previously thought.

These premises claim for the urgent need of reliable diagnostic markers, able to identify the different forms of neurodegenerative dementias, since the early disease stages, to be easily introduced as a screening tool in memory clinics, even in primary or secondary referral centers. Currently, validated markers, divided into imaging modalities and fluid measures, are used on clinical grounds and have proven to be highly accurate in diagnosing dementia.[8] However, a number of drawbacks may limit the use of these markers, thus being considered only in selected cases. In particular, some are able to identify AD, but are unhelpful in other forms of dementia (i.e., amyloid PET imaging or cerebrospinal fluid $A\beta_{42}$ and Tau dosages), others are not useful in early disease stages at single subject level (i.e., brain MRI); moreover, the invasiveness of the procedure (i.e., cerebrospinal fluid analysis) or the expensiveness (i.e., PET amyloid) may further confine their availability. Notably, the ideal marker, besides having high accuracy and reliability, should be non-invasive, simple to perform and inexpensive.[9]

In this context, transcranial magnetic stimulation (TMS) has shown to be a reliable tool to non-invasively assess a series of intracortical circuits which indirectly rely on several neurotransmitters, as GABA, glutamate and acetylcholine.[10,11]

A considerable body of literature has historically shown that AD and DLB are characterized by a deficit in short latency afferent inhibition (SAI),[12–21] a marker of sensorimotor integration which largely relies on cholinergic circuits,[11] while FTD and DLB show a striking alteration in short interval intracortical inhibition and facilitation (SICI-ICF),[22–25] which substantially depend on GABAergic and glutamatergic circuits,[11] respectively. These findings have prompted subsequent studies, which have suggested that a neurophysiological assessment might not be very far from being ready to be translated from the experimental to the clinical setting.[26–30] However, to further confirm TMS utility for the diagnosis of AD and other neurodegenerative dementias and to extend its use broadly, multicenter studies assessing the best combination of TMS measures, thus achieving the highest classification performance, are desirable.

As a results of these observations, we hypothesized that by applying state-of-the-art machine learning techniques to neurophysiological measures, obtained from multicenter studies, we could obtain a very high diagnostic accuracy in the differential diagnosis of the most common neurodegenerative dementing disorders.

Methods

Subjects

In this study, we collected patients retrospectively from four centers in Italy, namely from the Neurology Unit, Department of Clinical and Experimental Sciences, University of Brescia, Italy, from the Memory Clinic of the Tor Vergata University, Rome, from the Neurology Unit, Campus Bio-Medico University, Rome, Italy, and the Neurology Unit, Merano Hospital, Merano, Italy. Several patients were already enrolled in previous published studies by the authors. Each included patient fulfilled current clinical criteria for probable AD,[31] DLB,[7] or FTD.[5,6] Dementia was defined when cognitive or behavioral (neuropsychiatric) symptoms interfered with the ability to function at work or at usual activities, representing a decline from previous levels of functioning and performing, not explained by other disorders (i.e. delirium, major psychiatric disorders). Regarding AD, patients met criteria for dementia, having an insidious onset with a clear-cut history of worsening of cognition, and an amnestic presentation.[31] In cases of non-amnestic presentations but with a high suspect of AD pathophysiology, patients underwent amyloid PET imaging or cerebrospinal fluid (CSF) analysis (see below). For DLB, patients fulfilled the diagnosis of dementia, alternately associated with fluctuating cognition with pronounced variations in attention and alertness, recurrent visual hallucinations, REM sleep behavior disorder and at least one spontaneous cardinal feature of parkinsonism.[7] Regarding FTD, patients were classified as probable behavioral variant FTD (bvFTD)[6] or primary progressive aphasia (PPA)[5] based on the initial and most prominent clinical features, associated with fronto-insular/temporal atrophy at magnetic MRI or hypometabolism at ^{18}F-fluorodeoxyglucose PET (FDG-PET).

All patients considered in the present study underwent an extensive neuropsychological evaluation, according to standard procedures at each center and based on the expertise of each clinician. Patients were followed for at least two years, and clinical diagnoses were confirmed at follow-up. Brain MRI with a 1.5 or 3T scanner was performed in all patients. In order to exclude patients with vascular cognitive disorders, patients with vascular lesions, small vessel disease, strategic lacunar infarcts, or cerebral hemorrhages at MRI,[32] were excluded from the present study.

In selected cases, when the diagnostic confidence was not satisfactory, FDG-PET, single-photon emission computed tomography (SPECT)-DaTSCAN or ^{123}I-MIBG myocardial scintigraphy was performed (i.e. differential diagnosis between AD and DLB).

In a subgroup of patients, diagnosis was accomplished by amyloid markers, such as CSF $A\beta_{42}$ determinations or amyloid PET imaging, which supported or ruled out AD diagnosis. A CSF AD-like profile was defined as $A\beta_{1\text{-}42} \leq 650$ ng/L and tau ≥ 400 ng/L using a commercial ELISA assay,[33] while PET amyloid imaging was acquired using 370 MBq (10 mCi) of [^{18}F]-florbetapir or [^{18}F]-flutemetamol and visual readings were performed by nuclear medicine physicians who were blinded to the patients' diagnosis, following the procedures provided by the ligand manufacturer, as previously reported.[34]

Exclusion criteria were as follows: *i*) use of drugs that could affect TMS variables, *ii*) history of head trauma, alcohol abuse, stroke or transient ischemic attack, or epilepsy; *iii*) presence of pacemaker or other cardiac devices, cochlear implants, or previous brain surgery, such as clipping of a cerebral aneurysm.

A group of healthy controls (HC) was included, who underwent a brief standardized neuropsychological assessment (MMSE ≥ 27/30); psychiatric or other neurological illnesses were considered as exclusion criterion.

Full written informed consent was obtained from all participants according to the Declaration of Helsinki. The study protocol was approved by the local ethics committees of the participating centers.

Study design

The study was performed in accordance with the Standard for Reporting of Diagnostic Accuracy (STARD) criteria, applying the reference and index test at recruitment (see **Figure 1**). All subjects underwent an extensive clinical and instrumental work-up and the diagnosis was made by neurologists with expertise in neurodegenerative disorders (AB, GC, VDL, AA, MSC, AP, BB) (i.e., reference test).

All subjects underwent TMS study at recruitment, performed by examiners who had experience with neurophysiological techniques (RN, VC, VD, FR) and who were masked to the results of the reference test (i.e., index test). Data analysis was done by two separate statisticians (MG and FP).

Our primary research question was to determine the classification performance of AD and other dementias, considering the best combination of TMS indicators.

Transcranial magnetic stimulation parameters

The four centers applied comparable TMS protocols. A TMS figure-of-eight coil (each loop diameter 70 mm) connected to a monophasic Magstim Bistim or Bistim^2 system (Magstim

Company, Oxford, UK) was employed for all TMS paradigms. Electromyographic (EMG) recordings were performed from the first dorsal interosseous (FDI) muscles using 9 mm diameter, Ag-AgCl surface-cup electrodes. The active electrode was placed over the muscle belly and the reference electrode over the metacarpophalangeal joint of the index finger. Responses were amplified and filtered at 20 Hz and 2 kHz with a sampling rate of 5 kHz.

To locate the precise representation of the target muscle on the contralateral primary motor cortex, the TMS coil was positioned approximately 4 cm laterally and 2 cm anteriorly to Cz, tangentially on the scalp with the coil handle pointed 45° posteriorly and laterally to the sagittal plane. The "hot spot" was defined as the point in which magnetic stimulation resulted in the maximum motor evoked potential (MEP) amplitude with the minimum stimulator intensity. To obtain this, stimulator intensity was increased from 35% of the maximal stimulator output (MSO) in 5% steps until MEPs with an approximately 0.5-1 mV amplitude could be recorded. The coil was then moved in 0.5 cm steps medially, laterally, posteriorly and anteriorly while evoking 3 MEPs at each site.[35] This was performed until the site in which the largest MEPs could be located, which was marked with a felt tip pen on the scalp to ensure constant placement of the coil throughout the experiment.

RMT was defined as the minimal stimulus intensity needed to produce MEPs with an amplitude of at least 50 μV in 5 out of 10 consecutive trails during complete muscle relaxation, which was controlled by visually checking the absence of EMG activity at high-gain amplification. The active motor threshold (AMT) was determined during a slight tonic contraction of the target muscle at approximately 20% of the maximal muscle strength. MT was determined according to the relative frequency method, in which we started at a stimulus intensity of 35% MSO with the coil placed over the motor "hot spot", and stimulus intensity was gradually increased in steps of 5% MSO until

TMS consistently evoked MEPs with peak-to-peak amplitudes of >50 μV in each trial for RMT. Thereafter, stimulus intensity was gradually lowered in steps of 1% MSO until there were less than 5 positive responses out of 10 trials. For AMT, MEPs greater than 200 μV were judged to be positive.

SICI-ICF, LICI and SAI were studied using a paired-pulse protocol, employing a conditioning-test design. For all paradigms, the test stimulus (TS) was adjusted to evoke a MEP of approximately 1 mv peak-to-peak amplitude.

For SICI and ICF, the conditioning stimulus (CS) was adjusted at 70% of the RMT or 5% of below AMT (based on individual preferences at each center), employing multiple interstimulus intervals (ISIs), including 1, 2, 3, 5 ms for SICI and 7, 10, 15 ms for ICF.[36,37]

Long interval intracortical inhibition (LICI), which predominantly reflects $GABA_B$ergic transmission, was elicited by applying two suprathreshold stimuli at long ISIs (50, 100, 150 ms), with the CS set at 130% of the RMT preceding the TS.[38]

SAI was evaluated employing a CS consisting of a single pulse (200 μs) of electrical stimulation at the right median nerve at the wrist, using a bipolar electrode with the cathode positioned proximally, at an intensity sufficient to evoke a visible twitch of the thenar muscles.[39] Different ISIs were implemented ($^{-}$4, 0, $^{+}$4, $^{+}$8 ms), which were fixed relative to the latency of the N20 component of the somatosensory evoked potential of the median nerve.

For each ISI and for each protocol (SICI-ICF, LICI and SAI), from 5 to 10 (depending on each center) different paired CS-TS and control TS were delivered in all participants in a pseudo randomized sequence, with an inter-trial interval of 5 secs (±10%).

The conditioned MEP amplitude, evoked after delivering a paired CS-TS, was expressed as percentage of the average control MEP amplitude. Stimulation protocols were conducted in a randomized order. Audio-visual feedback was provided to ensure muscle relaxation during the entire experiment and trials were discarded if EMG activity exceeded 100 μV in the 250 ms prior to TMS stimulus delivery. Less that 5% of trials were discarded for each protocol. All of the participants were capable of following instructions and reaching complete muscle relaxation; if, however the data was corrupted by patient movement, the protocol was restarted and the initial recording was rejected.

The operators who administered TMS were blinded to the subjects' status; standardized TMS procedures were employed for all participants and stimuli were delivered in a randomized sequence, thus reducing possible biases in TMS recordings.

For the purpose of the present study we considered as potential indicators each of the following parameters: SICI and ICF (at 1, 2, 3, 5, 7, 10, 15 ms ISIs), LICI (at 50, 100, 150 ms ISIs), and SAI (at –4, 0, +4, +8 ms ISIs). For every patient, considering that protocols were performed with slightly different parameters between each center (i.e. not every ISI was performed in all centers), machine learning algorithms were applied (see next section) to infer average values and trends of each measure for each paired-pulse protocol.

Statistical analysis

Descriptive analysis. TMS raw measures were compared using two-way mixed ANCOVA (for SICI-ICF, LICI and SAI) with GROUP as between-subjects factor and ISI as within-subjects factor, including age at TMS and center as covariates. If a significant main effect was observed, group

differences were evaluated with *post hoc* tests (Bonferroni correction for multiple comparisons). Mauchly's test was used to check for sphericity violation, applying Greenhouse-Geisser epsilon determinations.

Machine Learning (ML) model. The step-by-step ML design is depicted in **Figure 3**, and detailed below.

Step 1 - TMS parameters selection and missing data imputation. TMS intracortical connectivity measures, namely SICI-ICF, SAI and LICI, were considered. Missing values were imputed using the k-nearest neighbors (KNN) algorithm as previously reported.[40] To prevent the loss of predictive accuracy, imputation was restricted to subjects with at least 5 out of 7 data points for SICI-ICF and 2 out of 4 data points for SAI. No restrictions were applied for LICI, to avoid extensive data loss and having only a slight impact on prediction accuracy.

Step 2 - Regression analysis. Three distinct regression analyses, on the basis of previous published data on SICI-ICF, SAI and LICI curve patterns, were carried out (see **Figure 3, panel A)**. For each TMS protocol, we performed regression analysis to capture both baseline average values (i.e., intercept or zero-order parameters) and trends (i.e., regressor coefficients) through time (ISIs). Regressions take the general form: y ~ poly(t); i.e., the TMS parameter y for each subject is predicted as a polynomial function of time (ISIs). Each TMS parameter has its own polynomial function (see **Figure 3, panel B**).

LICI was analyzed with a simple linear regression (LR), in the form: y(LICI) = a0 + a1*t. The Broken Line Regression (BLR) was applied for SICI and ICF: y(SICI) = bs0 + bs*t and y(ICF) = bi0 + bi*t, where bs0 and bs are the intercept and slope parameters for SICI (at 1, 2, 3, 5 ms), and

bi0 and bi are the intercept and slope parameters for ICF (at 7, 10, 15 ms), respectively. The BLR allowed us to account for the dual behavior of the SICI-ICF signal at different ISIs. Finally, Quadratic Regression (QR) was used for SAI, considering its parabolic shape: y(SAI) = b0 + b1*t + b2*t^2.

This led to the production of a set of 9 regression parameters (i.e., those estimating baseline mean ISIs: a0, bs0, bi0, b0, and those estimating curve trends at different ISIs: a1, bs, bi, b1, b2); these regression parameters were the features input for the subsequent classification step, using either unadjusted values or adjusted values by demographic variables, including age at TMS, sex and center (as for ANCOVA procedures, we did not include age at onset because it was autocorrelated with age at TMS, r=0.962, p<0.001, and results did not differ adding this covariate). Covariate adjustment was performed by computing the regression parameter residuals per subject after the regression parameter fitting on demographic features.

Step 3 – Random Forest models. Random Forests (RFs) classifier was carried out[41] and regression parameters were used as predictors for binary classification through decision tree (see **Figure 3, panel C**). The RF is a classifier that includes a large number of decision tree classifiers. Each tree is trained with randomly selected with replacement (i.e., bootstrapped) learning samples and at each node a subset of features is randomly selected to generate the best split, and then the best classification is conducted based on a majority vote of the trees in the forest. We chose hierarchical two-group classifications approach, as this method usually yields higher predictive accuracy as compared to multi-group classification. Mean Decrease Accuracy (MDA), namely the importance of a feature in determining classification accuracy, and Mean Decrease Gini Impurity (MDG), namely the importance of a feature to discriminate between classes, were used for features ranking

and plotting. In addition, Multidimensional Scaling (MDS) plots were obtained for illustrative visualization of RF outputs.

Step 4 – Validation. To avoid classifier over-performance on a specific dataset, and consequent loss of classification generality and reproducibility, we performed a K-fold cross-validation analysis. At each iteration, K-1 partitions were merged into one and used for the learning process (training step), while the K-th left out partition (i.e., the validation set) was used to predict the outcome (i.e., the diagnostic class). Once every iteration cycle was completed, we moved to the next partitioning configuration, until every K-th partition has been used as validation set. We performed a multiple 10, 5, and 4-fold-cross validation analysis.

Performance of the classifier was computed considering the following indices:[42] *i)* classification accuracy, i.e. the ratio of correctly predicted (positive or negative) observations to the total observations; *ii)* precision, i.e. the ratio of correctly predicted positive observations to the total predicted positive observations (precision estimates classifier's ability to predict really positive observations when the test is positive); *iii)* recall, i.e. the ratio of correctly predicted positive observations to the total true positive observations (recall estimates the amount of true positive observations that were correctly classified as positive); and *iv)* F1-score, i.e. the harmonic average of precision and recall.

The 2×2 frequency table (the so-called confusion matrix) was obtained at each iteration of the K-fold cross-validation, and the performance indices were computed both by averaging the indices of K 2×2 tables and by using the overall 2×2 table over the K iterations.

Software

Descriptive analyses were carried out using SPSS software (SPSS 21.0. Armonk, NY). Random forest classifier and evaluation of classification performance were carried out in R-3.6.0, using *RandomForest* package, with ntree=1000=number of trees to grow and mtry=sqrt(9)=number of variables randomly sampled as candidates at each split;[43] *reprtree* package for selection of the most representative trees;[44] *CMA* package for performance evaluation with K-fold cross-validation and custom R visualization functions (MDS_plot).[45]

Results

Subjects

Seven hundred and eight participants were assessed for eligibility, and 14 were excluded because they could not undergo TMS testing (2%): they were carrying electronic implants (n=4), they had a positive history of seizures (n=7), or because 1 mV MEPs could not be obtained by using stimulator intensities <85% of the maximum stimulator output (n=3) (see **Figure 1**).

Thus, 694 subjects were considered in the present analysis, 504 of whom were recruited at the University of Brescia, 103 at the University of Tor Vergata in Rome, 64 at Campus Bio-Medico University in Rome and 23 at Merano Hospital, Italy. Of these, 54.2% have been included in previous studies published by the authors.

Demographic characteristics of the diagnostic groups are reported in **Table 1**. We included 273 patients with AD, 207 patients with FTD (60 with Primary Progressive Aphasia and 147 with behavioral variant FTD), 67 patients with DLB, and 147 subjects as HC.

Three-hundred thirty-seven (48.5%) patients had at least one amyloid marker (PET amyloid or CSF $A\beta_{1\text{-}42,}$ and tau dosage), which further supported or excluded an AD diagnosis.

TMS connectivity measures and regression parameters estimation

TMS connectivity measures, i.e. SICI-ICF, SAI and LICI in the different diagnostic groups are reported in **Figure 2.** We observed a significant interaction at the two-way mixed ANCOVA for SICI-ICF [$F(12.0,2411.2)=81.1$, $p<0.001$, partial $\eta^2=0.29$, $\varepsilon=0.67$], LICI [$F(5.6,867.1)=2.4$, $p=0.031$, partial $\eta^2=0.02$, $\varepsilon=0.93$] and SAI [$F(7.2,1382.1)=13.8$, $p<0.001$, partial $\eta^2=0.07$, $\varepsilon=0.80$]. *Post-hoc* differences, corrected for multiple comparisons, between groups and for each ISI, are reported in **Figure 2**. Briefly, in comparison to healthy controls, SICI-ICF resulted significantly

impaired in both FTD and DLB, SAI was significantly impaired in AD and DLB, while LICI was significantly impaired in FTD.

We did not observe significant differences in MT intensities (expressed as % of MSO) between centers, $F(6,680)=1.82$, $p=0.092$, $\eta^2=0.02$, or in SICI-ICF measures obtained with different conditioning stimulus intensities (70% RMT or AMT-5%), which were distinctly adopted between centers, $F(1,682)=0.12$, $p=0.726$, $\eta^2<0.01$.

For each subject, the set of 9 regression parameters on the basis of ISI values and curve shapes of SICI-ICF, SAI and LICI, was calculated (see method section for details) and used in the Random Forests classifier. The mean values of each regression parameter according to diagnostic group, and significant differences between groups are reported in **Supplementary Table 1**.

Random Forests (RFs) models and classification performances

As reported in **Figure 1 (lower part)**, a series of subsequent 3binary (two-groups) and independent classifiers were employed: 1) cases (AD, FTD or DLB) vs HC; 2) FTD vs non-FTD (AD or DLB), 3) AD vs DLB.

The specific order of classification resulted in the greatest accuracy, i.e. fewer classification errors. The first two-groups classification allowed us to classify each subject as "case" (i.e., patient with dementia) or "control"; if the subject fell into the "case" category, the next order of classification was considered, and the FTD vs non-FTD classifier was carried out; once again, if the patient fell into the "non-FTD" category, the third classifier allowed us to classify the patient into AD vs DLB.

For each of the three classifiers, Multi-dimensional scaling (MDS) plots are reported in **Figure 4**. MDS plots tend to a separate shape: the more pronounced is the separation the better the

classification performance is obtained (especially visible in the first and second RF classifier). In the first classifier, cases vs HC, the top-ranking variables of importance were SAI b0, SICI bs and ICF bi; in the second classifier, FTD vs non-FTD, were SAI b0, SICI bs and ICF bi0, while in the third classifier, AD vs DLB, were ICF bi, SICI bs0 and ICF bi0.

Classification indices with 5-fold-cross validation (10 and 4-fold cross validation run similar results), unadjusted and adjusted for age, sex and center, are reported in **Table 2** (A-B), after removing outliers with Brier score >1.[46] The prediction model exhibited overall high accuracy (ranging from 0.89 to 0.92), high precision (0.86-0.92), high recall (0.93-0.98), and high F1-scores (0.89-0.95).

To evaluate if results were consistent in patients with a higher diagnostic confidence, classification performance was also evaluated by 5-fold-cross validation on control subjects and the patient subgroups with a biomarker-supported diagnosis, obtaining very similar results [see **Table 2** (C)].

Considering the large sample size, we reported classification results using a single random independent validation set, corresponding to roughly half of the sample size (for each RF classifier), obtaining comparable results [see **Table 2** (D)].

Discussion

In this work, we aimed at maximizing the potential diagnostic performances of TMS measures by applying Random Forest classifiers to the diagnosis of AD and other neurodegenerative dementias. We have previously demonstrated that SICI-ICF, LICI and SAI parameters, as well as SICI-ICF/SAI ratio, are helpful to differentiate AD, FTD and DLB[27] with high accuracy. However, the previously proposed cut-offs (not cross-validated) cannot theoretically reproduce comparable results when different centers and new subjects are considered (data not shown). Even though the same findings were independently obtained by other groups in the past,[24,47,48] a systematic evaluation of TMS performances was essential to prove its diagnostic utility.

In this multicenter study which included a very large sample size, we propose a diagnostic approach which is able to differentiate the most common forms of dementias, whose performance metric was validated using a machine-learning based computer-aided approach.

We observed very high levels of classification accuracy, precision and recall, applying an intuitive and straightforward step-by-step approach: in the first step healthy controls are identified and excluded, in the second step FTD is recognized from the dementias of the central cholinergic system (i.e., AD and DLB), while in the third step AD is differentiated from DLB. Indeed, using the classification parameters obtained in the Random Forest analysis, an automated R script was coded to allow the simple and straightforward entry of raw TMS measures, which are computed and elaborated, resulting in a diagnostic class for each diagnosis at the single subject level (an user-friendly R package and related documentation are available online at the GitHub Repository https://github.com/fernandoPalluzzi/tmsClassifier). The decision process is achieved evaluating 1000 rules (i.e. trees) per subject, and then the best classification is based on a majority vote of the

trees in the forest. As compared to other biological markers, TMS parameters are possibly able to selectively identify most of the spectrum of neurodegenerative dementias, such as AD, FTD, DLB and distinguish them from healthy ageing.

The high accuracy obtained using TMS measures possibly relies on the biological bases of dementia. If neuropathological criteria for AD, FTD and DLB classify each disease by a specific proteinopathy, parallelly, biological studies have detected a specific, now well-established, neurotransmitter impairment.

Indeed, as SICI and LICI are considered to reflect short-lasting postsynaptic inhibition mediated through the $GABA_A$ and $GABA_B$ receptors at the level of local interneurons,[36,37] and ICF is thought to represent a net facilitation most likely mediated by glutamatergic NMDA receptors,[11,37] the impairment observed in FTD suggests a deficit of GABAergic and glutamatergic interneurons.[49–51] On the other hand, Aβ peptides have been shown to impair acetylcholine synthesis and release, and to induce cholinergic cell toxicity,[52] which could reflect the impairment of SAI observed in AD.[12–21] In the same view, the well-recognized cholinergic deficit and the documented impairment of GABAergic and glutamatergic neurotransmission in DLB might explain the impairment of both SAI and SICI-ICF in these patients.[27,53,54] Summarizing, from what we have observed in this study, SICI-ICF was significantly impaired in FTD and DLB, LICI mostly in FTD, while SAI was impaired in both DLB and AD patients. The combination of these measures has been shown to be accurate in differentiating these neurodegenerative diseases from one another.

Diagnosis of neurodegenerative dementias is still challenging, with substantial diagnostic delays, employing different methods which may be invasive or expensive, as cerebrospinal fluid analysis or amyloid PET imaging. TMS has the advantage that it can be carried out concomitantly in the

outpatient visit, being non-invasive, relatively non-time consuming and rather inexpensive. For these reasons, beyond resulting as a marker with diagnostic accuracy values comparable to those of well-established biomarkers, TMS might represent a reliable and accessible screening tool,[55] considering the exponential increase in the prevalence of people living with dementia.[1]

The strength of the present work relies on the very large sample size with a multicenter enrolment and the machine-learning approach to data analysis to obtain highest possible values of diagnostic accuracy. Indeed, as compared to a previous work,[28] the large sample size allowed us to apply a 5-fold cross validation, which prevents results from overfitting and guarantees replicability to other datasets. Moreover, the Random Forests classifier is superior to the most available learning algorithms since it is easy to parameterize, robust against overfitting, not sensitive to noise in the data set (i.e. good at dealing with outliers in training data) and able to avoid biases due to unrelated centers.

However, the main limitation is that TMS remains a specialized technique requiring specific technology (special hardware and software) and training. Consequently, although the tolerability of the index test was very good, the practicability in primary or secondary referral dementia centers is still limited. This issue needs to be further assessed considering that worldwide the majority of patients with dementia come from low- and middle-income countries,[2] and inexpensive screening tools will be mandatory.

Moreover, a number of questions needs to be further addressed. Above all, the performance of TMS in real-world situations as a screening tool should be tested as well as the accuracy in subjects in the preclinical or prodromal phases of dementia, even though most of the patients included in the present work had a mild disease stage. Another relevant aspect relies on the possible overlap

between pathologies, considering that approximately 40% of dementias are of mixed type,[56] and which were not accounted for in this study. Nevertheless, excluding patients with a vascular cognitive decline based also on MRI assessments,[32] which account for a significant overlap in pathological studies, could partially mitigate this issue. We cannot also exclude that healthy controls included in the present study are biomarker positive in a preclinical phase of disease, since the majority did not undergo any biological marker assessment.

Despite these limitations, the addition of TMS measures to the routine diagnostic assessment could allow for an earlier diagnosis, increasing the enrolment of patients with dementia into therapeutic trials when combined with clinical and conventional methods of diagnosis. These findings support for the use of TMS intracortical connectivity measures to be translated from the experimental setting to the clinical practice.

Acknowledgements

We thank the patients and their families for participating in this research. We also thank the many physicians, genetic counselors, and families for referring patients.

Author contributions

Conception and design of the study: AB, MG and BB. Acquisition and analysis of data: AB, MG, FP, GC, VDL, RN, VC, VD, EP, AM, FDL, SB, FR, FC, GB, MSC, AP and BB. Drafting the manuscript and figures: AB, MG, FP and BB.

Potential Conflicts of Interest

Nothing to report.

References

1. GBD 2016 Dementia Collaborators. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. [Internet]. Lancet. Neurol. 2019;18(1):88–106.Available from: https://linkinghub.elsevier.com/retrieve/pii/S1474442218304034
2. World Alzheimer Report. The Global Impact of Dementia: an analysIs of prevalence, incidence, cost and trends [Internet]. 2015;Available from: https://www.alz.co.uk/research/WorldAlzheimerReport2015.pdf
3. Prince M, Bryce R, Albanese E, et al. The global prevalence of dementia: A systematic review and metaanalysis [Internet]. Alzheimer's Dement. 2013;9(1):63-75.e2.[cited 2019 Apr 30] Available from: https://www.sciencedirect.com/science/article/pii/S1552526012025319
4. Logroscino G, Piccininni M, Binetti G, et al. Incidence of frontotemporal lobar degeneration in Italy: The Salento-Brescia Registry study [Internet]. Neurology 2019;92(20):e2355–e2363.[cited 2019 Apr 30] Available from: http://www.ncbi.nlm.nih.gov/pubmed/30979859
5. Gorno-Tempini ML, Hillis AE, Weintraub S, et al. Classification of primary progressive aphasia and its variants. [Internet]. Neurology 2011;76(11):1006–1014.Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=21325651&retmode=ref&cmd=prlinks
6. Rascovsky K, Hodges JR, Knopman D, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. [Internet]. Brain 2011;134(Pt 9):2456–77.Available from: http://www.ncbi.nlm.nih.gov.bibliosan.cilea.it/pmc/articles/PMC3170532/pdf/awr179.pdf
7. McKeith IG, Boeve BF, Dickson DW, et al. Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium. [Internet]. Neurology 2017;89(1):88–100.Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=28592453&retmode=ref&cmd=prlinks
8. Frisoni GB, Boccardi M, Barkhof F, et al. Strategic roadmap for an early diagnosis of Alzheimer's disease based on biomarkers. [Internet]. Lancet. Neurol. 2017;16(8):661–676.Available from: http://linkinghub.elsevier.com/retrieve/pii/S147444221730159X
9. The Ronald and Nancy Reagan Research Institute of the Alzheimer's Association and National Institute on Aging Working Group. Consensus Report of the Working Group on: "Molecular and Biochemical Markers of Alzheimer's Disease" [Internet]. Neurobiol. Aging 1998;19(2):109–116.[cited 2019 Apr 30] Available from: https://www.sciencedirect.com/science/article/pii/S0197458098000220?via%3Dihub
10. Rossini PM, Burke D, Chen R, et al. Non-invasive electrical and magnetic stimulation of the brain, spinal cord, roots and peripheral nerves: Basic principles and procedures for routine clinical and research application. An updated report from an I.F.C.N. Committee. [Internet].

Clin. Neurophysiol. 2015;126(6):1071–1107.Available from: http://linkinghub.elsevier.com/retrieve/pii/S1388245715000711

11. Ziemann U, Reis J, Schwenkreis P, et al. TMS and drugs revisited 2014. [Internet]. Clin. Neurophysiol. 2015;126(10):1847–68.Available from: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L601025833%0Ahttp://dx.doi.org/10.1016/j.clinph.2014.08.028

12. Di Lazzaro V, Oliviero A, Tonali PA, et al. Noninvasive in vivo assessment of cholinergic cortical circuits in AD using transcranial magnetic stimulation [Internet]. Neurology 2002;59(3):392–397.Available from: http://www.neurology.org/cgi/doi/10.1212/WNL.59.3.392

13. Di Lazzaro V, Oliviero A, Pilato F, Saturno E. Motor cortex hyperexcitability to transcranial magnetic stimulation in Alzheimer's disease [Internet]. J. Neurol. 2004;75(4):555–559.Available from: http://jnnp.bmj.com/cgi/doi/10.1136/jnnp.2003.018127

14. Nardone R, Bratti A, Tezzon F. Motor cortex inhibitory circuits in dementia with Lewy bodies and in Alzheimer's disease [Internet]. J. Neural Transm. 2006;113(11):1679–1684.Available from: papers3://publication/doi/10.1007/s00702-006-0551-1

15. Di Lazzaro V, Pilato F, Dileone M, et al. In vivo functional evaluation of central cholinergic circuits in vascular dementia [Internet]. Clin. Neurophysiol. 2008;119(11):2494–2500.Available from: http://linkinghub.elsevier.com/retrieve/pii/S1388245708009279

16. Nardone R, Bergmann J, Kronbichler M, et al. Abnormal short latency afferent inhibition in early Alzheimer's disease: a transcranial magnetic demonstration [Internet]. J. Neural Transm. 2008;115(11):1557–1562.Available from: http://link.springer.com/10.1007/s00702-008-0129-1

17. Martorana A, Esposito Z, Di Lorenzo F, et al. Cerebrospinal fluid levels of Aβ42 relationship with cholinergic cortical activity in Alzheimer's disease patients [Internet]. J. Neural Transm. 2012;119(7):771–778.Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=22402892&retmode=ref&cmd=prlinks

18. Di Lorenzo F, Martorana A, Ponzo V, et al. Cerebellar theta burst stimulation modulates short latency afferent inhibition in Alzheimer's disease patients. [Internet]. Front. Aging Neurosci. 2013;5:2.Available from: http://journal.frontiersin.org/article/10.3389/fnagi.2013.00002/abstract

19. Nardone R, Tezzon F, Höller Y, et al. Transcranial magnetic stimulation (TMS)/repetitive TMS in mild cognitive impairment and Alzheimer's disease [Internet]. Acta Neurol. Scand. 2014;129(6):351–366.Available from: http://doi.wiley.com/10.1111/ane.12223

20. Koch G, Di Lorenzo F, Del Olmo MF, et al. Reversal of LTP-Like Cortical Plasticity in Alzheimer's Disease Patients with Tau-Related Faster Clinical Progression. [Internet]. J. Alzheimers. Dis. 2016;50(2):605–616.Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=26757193&retmode=ref&cmd=prlinks

21. Di Lazzaro V, Oliviero A, Pilato F, et al. Neurophysiological predictors of long term response to AChE inhibitors in AD patients. [Internet]. J. Neurol. Neurosurg. Psychiatry 2005;76(8):1064–9.Available from: http://jnnp.bmj.com/cgi/doi/10.1136/jnnp.2004.051334

22. Benussi A, Gazzina S, Premi E, et al. Clinical and biomarker changes in presymptomatic genetic frontotemporal dementia. [Internet]. Neurobiol. Aging 2019;76(2019):133–140.Available from: https://doi.org/10.1016/j.neurobiolaging.2018.12.018

23. Gazzina S, Benussi A, Premi E, et al. Neuroanatomical Correlates of Transcranial Magnetic Stimulation in Presymptomatic Granulin Mutation Carriers [Internet]. Brain Topogr. 2018;31(3):488–497.Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=29256051&retmode=ref&cmd=prlinks

24. Burrell JR, Kiernan MC, Vucic S, Hodges JR. Motor Neuron dysfunction in frontotemporal dementia [Internet]. Brain 2011;134(9):2582–2594.Available from: http://www.brain.oxfordjournals.org/cgi/doi/10.1093/brain/awr195

25. Benussi A, Cosseddu M, Filareto I, et al. Impaired long-term potentiation-like cortical plasticity in presymptomatic genetic frontotemporal dementia. [Internet]. Ann. Neurol. 2016;80(3):472–6.Available from: http://doi.wiley.com/10.1002/ana.24731

26. Pottier C, Zhou X, Perkerson III RB, et al. Potential genetic modifiers of disease risk and age at onset in patients with frontotemporal lobar degeneration and GRN mutations: a genome-wide association study [Internet]. Lancet Neurol. 2018;17(6):548–558.Available from: https://linkinghub.elsevier.com/retrieve/pii/S1474442218301261

27. Benussi A, Dell'Era V, Cantoni V, et al. Discrimination of atypical parkinsonisms with transcranial magnetic stimulation [Internet]. Brain Stimul. 2018;11(2):366–373.Available from: http://linkinghub.elsevier.com/retrieve/pii/S1935861X17309695

28. Benussi A, Di Lorenzo F, Dell'Era V, et al. Transcranial magnetic stimulation distinguishes Alzheimer disease from frontotemporal dementia. Neurology 2017;89(7):665–672.

29. Motta C, Di Lorenzo F, Ponzo V, et al. Transcranial magnetic stimulation predicts cognitive decline in patients with Alzheimer's disease. [Internet]. J. Neurol. Neurosurg. Psychiatry 2018;89(12):1237–1242.Available from: http://jnnp.bmj.com/lookup/doi/10.1136/jnnp-2017-317879

30. Padovani A, Benussi A, Cantoni V, et al. Diagnosis of mild cognitive impairment due to Alzheimer's disease with transcranial magnetic stimulation [Internet]. J. Alzheimer's Dis. 2018;65(1):221–230.Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=30010131&retmode=ref&cmd=prlinks

31. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease [Internet]. JALZ 2011;7(3):263–269.Available from: http://dx.doi.org/10.1016/j.jalz.2011.03.005

32. Sachdev P, Kalaria R, O'Brien J, et al. Diagnostic criteria for vascular cognitive disorders: A VASCOG statement. Alzheimer Dis. Assoc. Disord. 2014;28(3):206–218.

33. Borroni B, Benussi A, Archetti S, et al. Csf p-tau181/tau ratio as biomarker for TDP pathology in frontotemporal dementia. Amyotroph. Lateral Scler. Front. Degener. 2015;16(1–2):86–91.

34. Boccardi M, Altomare D, Ferrari C, et al. Assessment of the Incremental Diagnostic Value of Florbetapir F 18 Imaging in Patients With Cognitive Impairment: The Incremental Diagnostic Value of Amyloid PET With [18F]-Florbetapir (INDIA-FBP) Study. [Internet]. JAMA Neurol. 2016;73(12):1417–1424.Available from: http://archneur.jamanetwork.com/article.aspx?doi=10.1001/jamaneurol.2016.3751

35. Groppa S, Oliviero A, Eisen A, et al. A practical guide to diagnostic transcranial magnetic stimulation: report of an IFCN committee. [Internet]. Clin. Neurophysiol. 2012;123(5):858–882.Available from: http://linkinghub.elsevier.com/retrieve/pii/S1388245712000569

36. Kujirai T, Caramia MD, Rothwell JC, et al. Corticocortical inhibition in human motor cortex. [Internet]. J. Physiol. 1993;471(1):501–519.Available from: http://doi.wiley.com/10.1113/jphysiol.1993.sp019912

37. Ziemann U, Rothwell JC, Ridding MC. Interaction between intracortical inhibition and facilitation in human motor cortex. [Internet]. J. Physiol. 1996;496(3):873–881.Available from: http://doi.wiley.com/10.1113/jphysiol.1996.sp021734

38. Valls-Solé J, Pascual-Leone A, Wassermann EM, Hallett M. Human motor evoked responses to paired transcranial magnetic stimuli. [Internet]. Electroencephalogr. Clin. Neurophysiol. 1992;85(6):355–364.Available from: http://linkinghub.elsevier.com/retrieve/pii/016855979290048G

39. Tokimura H, Di Lazzaro V, Tokimura Y, et al. Short latency inhibition of human hand motor cortex by somatosensory input from the hand. [Internet]. J. Physiol. 2000;523 Pt 2(2):503–513.Available from: http://doi.wiley.com/10.1111/j.1469-7793.2000.t01-1-00503.x

40. Troyanskaya O, Cantor M, Sherlock G, et al. Missing value estimation methods for DNA microarrays. Bioinformatics 2001;17(6):520–525.

41. Breiman L. Random Forests [Internet]. Mach. Learn. 2001;45(1):5–32.Available from: http://link.springer.com/article/10.1023/A:1010933404324

42. Powers DMW, Ailab. Evaluation: From Precision, Recall and F-Measure To Roc, Informedness, Markedness & Correlation [Internet]. J. Mach. Learn. Technol. ISSN 2011;2(1):2229–3981.Available from: http://www.bioinfo.in/contents.php?id=51

43. Liaw A, Wiener M. Classification and Regression by randomForest. R News 2003;2/3(December 2002):18–22.

44. Banerjee M, Ding Y, Noone AM. Identifying representative trees from ensembles. Stat. Med. 2012;31(15):1601–1616.

45. Slawski M, Daumer M, Boulesteix AL. CMA - A comprehensive Bioconductor package for supervised classification with high dimensional data. BMC Bioinformatics 2008;9:1–17.

46. Brier GW. Verification of forecasts expersses in terms of probaility. Mon. Weather Rev. 1950;78(1):1–3.

47. Di Lazzaro V, Pilato F, Dileone M, et al. In vivo cholinergic circuit evaluation in frontotemporal and Alzheimer dementias. [Internet]. Neurology 2006;66(7):1111–1113.Available from: http://www.neurology.org/cgi/doi/10.1212/01.wnl.0000204183.26231.23

48. Koch G, Di Lorenzo F, igrave SB, et al. Dopaminergic Modulation of Cortical Plasticity in Alzheimer's Disease Patients [Internet]. Neuropsychopharmacology 2014;1–8.Available from: http://dx.doi.org/10.1038/npp.2014.119

49. Khundakar AA, Hanson PS, Erskine D, et al. Analysis of primary visual cortex in dementia with Lewy bodies indicates GABAergic involvement associated with recurrent complex visual hallucinations. [Internet]. Acta Neuropathol. Commun. 2016;4(1):66.Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=27357212&retmode=ref&cmd=prlinks

50. Murley AG, Rowe JB. Neurotransmitter deficits from fronto temporal lobar degeneration [Internet]. Brain 2018;141(5):1263–1285.Available from: https://academic.oup.com/brain/article/141/5/1263/4823510

51. Benussi A, Alberici A, Buratti E, et al. Toward a Glutamate Hypothesis of Frontotemporal Dementia. [Internet]. Front. Neurosci. 2019;13(March):304.Available from: https://www.frontiersin.org/article/10.3389/fnins.2019.00304/full

52. Kar S, Slowikowski SPM, Westaway D, Mount HTJ. Interactions between beta-amyloid and central cholinergic neurons: implications for Alzheimer's disease. [Internet]. J. Psychiatry Neurosci. 2004;29(6):427–441.Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=15644984&retmode=ref&cmd=prlinks

53. Di Lazzaro V, Pilato F, Dileone M, et al. Functional evaluation of cerebral cortex in dementia with Lewy bodies. [Internet]. Neuroimage 2007;37(2):422–9.Available from: papers3://publication/doi/10.1016/j.neuroimage.2007.05.003

54. Marra C, Quaranta D, Profice P, et al. Central cholinergic dysfunction measured "in vivo" correlates with different behavioral disorders in Alzheimer's disease and dementia with Lewy body. [Internet]. Brain Stimul. 2012;5(4):533–8.Available from: http://www.ncbi.nlm.nih.gov/pubmed/22019082

55. Di Lazzaro V. Emergence of neurophysiological biomarkers of Alzheimer disease. J. Neurol. Neurosurg. Psychiatry 2018;89(12):1235.

56. Barker WW, Luis CA, Kashuba A, et al. Relative Frequencies of Alzheimer Disease, Lewy Body, Vascular and Frontotemporal Dementia, and Hippocampal Sclerosis in the State of Florida Brain Bank [Internet]. Alzheimer Dis. Assoc. Disord. 2002;16(4):203–212.Available from: http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00002093-200210000-00001

Table 1. Demographic and clinical characteristics of included patients.

Variable	AD	FTD	DLB	HC
Patients (number)	273	207	67	147
Age, years	70.7±7.0§*	65.9±8.7†‡*	72.9±6.1§*	57.7±17.6†§‡
Gender, F%	50.9‡	45.4‡	22.4†*§	57.8‡
Age at onset, years	68.1±7.0§	62.8±8.3†‡	70.0±6.2§	-
Education, years	9.6±4.4*	10.5±4.5	9.7±4.7*	11.6±4.6†‡
MMSE scores	23.3±4.8*	22.2±7.9*	23.4±3.5*	29.1±1.6†§‡
CSF $A\beta_{42}$, pg/ml (n)	461.2±203.1 (93)§‡*	839±314.3 (111)†*	744.7±352.0 (9)†*	1140.0±169.3 (7)†§‡
CSF Tau, pg/ml (n)	725.5±394.6 (93)§‡*	381.0±248.4 (111)†	230.6±6.8 (9)†	204.0±64.3 (7)†
CSF pTau, pg/ml (n)	99.5±124.2 (93)§	54.8±38.7 (111)†	41.9±22.1 (9)	40.1±9.5 (7)
PET amy, positive % (n)	98.6 (69)§‡*	7 (27)†	0 (3)†	0 (12)†
TMS parameters				
RMT (% MSO)	0.43±0.08§	0.46±0.09†‡	0.42±0.09§	0.45±0.08
SICI	0.57±0.26§‡*	0.76±0.31†*	0.76±0.41†*	0.47±0.27†§‡
ICF	1.33±0.33§‡*	0.96±0.34†‡*	1.13±0.46†§*	1.44±0.3†§‡
SAI	0.81±0.19§*	0.54±0.17†‡	0.83±0.26§*	0.51±0.14†‡
LICI	0.45±0.26§*	0.69±0.40†‡*	0.43±0.35§	0.31±0.26†§

Demographic and clinical characteristics, and neurophysiological parameters are expressed as mean ± standard deviation; SICI, ICF, LICI and SAI are represented as ratio of mean conditioned and unconditioned (i.e. control) motor evoked potential (MEP) amplitude.

AD = Alzheimer's Disease; FTD = Frontotemporal Dementia; HC = healthy controls; n = number; F = female; MMSE=Mini-Mental State Examination; CSF = cerebrospinal fluid; PET amy = amyloid Positron Emission Tomography; TMS = Transcranial Magnetic Stimulation; RTM = resting motor threshold; SICI = mean short interval intracortical inhibition (1, 2, 3, 5 ms); ICF = mean intracortical facilitation (7, 10, 15 ms); SAI = mean short latency afferent inhibition (0, +4 ms); LICI = mean long interval intracortical inhibition (50, 100, 150 ms).

$^{*}p<0.05$ vs HC; $^{\dagger}p<0.05$ vs AD, $^{\ddagger}p<0.05$ vs DLB; $^{\S}p<0.05$ vs FTD using one-way ANOVA or chi-square tests, as appropriate (*post hoc* tests with Bonferroni correction for multiple comparisons, only after a significant interaction).

Table 2. Classification accuracy, precision, recall and F1-score of the two-group classifiers (5-fold cross-validation), after outliers with Brier score >1 were removed; training and validation performed on the whole dataset with: (A) unadjusted regression parameters, (B) adjusted for center, age and sex regression parameters, (C) patients with a biomarker-supported diagnosis and (D) in a single random independent validation set.

Two-group classifier	***First classier*** **Cases vs HC**	***Second classifier*** **FTD vs non-FTD**	***Third classifier*** **AD vs DLB**
A) All patients (unadjusted)	n=645	n=504	n=305
Accuracy	0.92	0.91	0.90
Precision	0.93	0.91	0.92
Recall	0.97	0.95	0.96
F1-score	0.95	0.93	0.94
B) All patients (adjusted)	n=649	n=508	n=320
Accuracy	0.89	0.89	0.92
Precision	0.90	0.86	0.92
Recall	0.97	0.93	0.98
F1-score	0.94	0.89	0.95
C) Patients with biomarkers	n=454	n=297	n=171
Accuracy	0.87	0.89	0.96
Precision	0.89	0.88	0.97
Recall	0.93	0.91	0.99
F1-score	0.91	0.90	0.98
D) Single random independent validation set	n=322	n=278	n=169
Accuracy	0.89	0.87	0.94
Precision	0.97	0.86	0.97
Recall	0.91	0.94	0.96
F1-score	0.94	0.89	0.97

AD = Alzheimer Disease; FTD = Frontotemporal Dementia; DLB = Dementia with Lewy Bodies; HC = Healthy Controls; Cases = AD, FTD or DLB; non-FTD =AD or DLB.

Legend to figures:

Figure 1. Flow diagram of the study.

After the index test, results were sorted on the basis of the reference standard (McKhann criteria for AD, Rascovsky and Gorno-Tempini criteria for FTD, and McKeith criteria for DLB).

AD = Alzheimer disease; FTD = frontotemporal dementia; DLB = Dementia with Lewy bodies; HC = healthy controls.

Figure 2. TMS connectivity parameters according to diagnostic groups.

(A) SICI at ISI of 1, 2, 3, and 5 ms and ICF at ISI of 7, 10, and 15 ms, (B) SAI at ISI of $^{-}4$, 0, $^{+}4$, and $^{+}8$ ms, and (B) LICI at ISI of 50, 100, and 150 ms in patients with AD, FTD, DLB and in HC. Data are presented as a ratio to the unconditioned motor evoked potential amplitude; error bars represent standard errors.

AD = Alzheimer disease; FTD = frontotemporal dementia; DLB = Dementia with Lewy bodies; HC = healthy controls; ICF = intracortical facilitation; ISI = interstimulus interval; LICI=long-interval intracortical inhibition; MEP = motor evoked potential; SAI = short-latency afferent inhibition; SICI = short-interval intracortical inhibition.

$^{*}p<0.05$ vs HC; $^{\dagger}p<0.05$ vs AD, $^{\ddagger}p<0.05$ vs DLB; $^{\S}p<0.05$ vs FTD using one-way ANOVA (*post hoc* tests with Bonferroni correction for multiple comparisons).

Figure 3. Source data (A), time trend parametrization (B), and Random Forest model (C).

(A) source data evaluation (TMS connectivity measures), according to previous literature. Mean scores at each interstimulus interval (ISI) and curve trend were considered; (B) regression analysis for each TMS connectivity measure, according to source data and each curve shape. Regressions take the general form: y ~ poly(t); i.e., the indicator y is predicted as a polynomial function of time; (C) Random Forest (RF) learning and classification.

AD = Alzheimer disease; FTD = frontotemporal dementia; DLB = Dementia with Lewy bodies; HC = healthy controls.

Figure 4. Multi-dimensional scaling (MDS) plots (panel A-C) according to each of the three classifiers.

MDS of classification forests and decisions tress of A: first classifier (y=1 cases, y=0 HC), B: second classifier (y=1 FTD, y=0 non-FTD), and C: third classifier (y=1 AD, y=0 DLB).

AD = Alzheimer disease; FTD = frontotemporal dementia; DLB = Dementia with Lewy bodies; HC = healthy controls; cases = AD or FTD or DLB; non-FTD = AD or DLB.

MDS plots visualize the proximity matrix accumulated for the training data by out-of-bag (OOB) observations, i.e. predicted values for each observation in the training dataset that are not included in the bootstrap samples.