# TRACE4AD: an AI model based on machine learning predicting the subject risk of Alzheimer's disease dementia from 3T MRI-T1 brain study and neuropsychological assessment

Matteo Interlenghi[1], Christian Salvatore[1], Isabella Castiglioni[2]

## Affiliations

1. **DeepTrace Technologies S.R.L., Milan, Italy**
   Matteo Interlenghi, Christian Salvatore

2. **Dipartimento di Fisica G. Occhialini, University of Milan-Bicocca, Milan, Italy**
   Isabella Castiglioni

## Corresponding author

Correspondence to Christian Salvatore, email address: salvatore@deeptracetech.com

Date: 10 March 2021

## Abstract

Background: To evaluate the performance of an AI model (TRACE4AD), based on machine learning, in predicting the subject risk level of Alzheimer's disease (AD) dementia from his/her 3T MRI-T1 brain study and neuropsychological assessment (NPA).
Methods: A cohort including a total of 73 patients who underwent 3T MRI-T1 brain examination (by different neuroimaging centers and 3T MRI systems) at baseline, with known stable clinical diagnosis at a follow-up of at least 24-months, were considered to test the AI model TRACE4AD (DeepTrace Technologies S.R.L., Milan, Italy). The model implemented in the software tool TRACE4AD was used in predicting the risk of AD dementia (low risk, high risk) by the automatic processing of the MRI-T1 brain images (for the whole cohort), in combination with NPA report (for a subcohort of 60 patients). The performance of the model was assessed in terms of accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV).
Results: TRACE4AD showed 78% accuracy [67-87%, 95% confidence interval], 78[62-89]% sensitivity, 79[61-91]% specificity, 82[66-92]% PPV, 74[57-88]% NPV based on MRI analysis only (whole cohort), and 83[71-92]% accuracy, 87[70-96]% sensitivity, 79[60-92]% specificity, 82[65-93]% PPV, 85[66-96]% NPV, based on the combined analysis of MRI and NPA (subcohort).
Conclusions: The AI tool TRACE4AD tool is efficace in predicting the individual risk of AD dementia within 24 months, supporting clinical decision making of neurologists and MRI reporting of neuroradiologists.

**Keywords (5)**

**Artificial Intelligence; Machine Learning; Alzheimer's Disease; Neuropsychological assessment, Dementia.**

## AIM

In this study, we evaluated the performance of the AI tool TRACE4AD when used for the automatic processing of single subject T1-weighted brain images from 3T Magnetic Resonance Imaging (MRI) systems and neuropsychological assessment (NPA) report to predicting the subject risk of Alzheimer's disease (AD) dementia (low risk, high risk). The tool is intended to be used by specialized clinicians as a decision support system for a personalized early diagnosis and prognosis for individual patients. We tested the diagnostic and prognostic performance of such software on a patients' cohort of 73 patients (automatically analyzing their MRI brain studies) and a subcohort of 60 patients (automatically analyzing their MRI brain studies in combination with their NPA report), being such testing cohorts fully independent from patients' cohorts used for the development and training of the TRACE4AD system.

## METHODS

### Study design and study population

This is a multi-center, observational, retrospective clinical study. The study population included a cohort of subjects obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI)[1]. The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), and the Food and Drug Administration (FDA), as a 5-year public private partnership, led by the principal investigator, Michael W. Weiner, MD. The primary goal of ADNI was to test whether serial MRI, positron emission tomography (PET), other biological markers, and clinical and NPA subjected to participants could be combined to measure the progression of mild cognitive impairment (MCI) and early AD – see www.adni-info.org.

As specified in the ADNI protocol[2], each participant was willing, spoke either English or Spanish, was able to perform all test procedures described in the protocol and had a study partner able to provide an independent evaluation of functioning. Inclusion criteria were the following:

1. Cognitively normal (CN) subjects having a Mini Mental State Examination (MMSE) (Folstein et al., 1975) score between 24 and 30, clinical dementia rating (CDR) of zero (Morris, 1993), and absence of depression, MCI and dementia;
2. MCI patients, having MMSE scores between 24 and 30, CDR of 0.5, objective memory loss measured by education-adjusted scores on the Logical Memory II subtest of the Wechsler Memory Scale (Wechsler, 1987), absence of significant levels of impairment in other cognitive domains, and absence of dementia;

[1] adni.loni.usc.edu

[2] http://www.adni-info.org/Scientists/ADNIStudyProcedures.html

3. AD patients with MMSE scores between 20 and 26, CDR of 0.5 or 1.0, and criteria for probable AD as defined by the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) e by the Alzheimer's Disease and Related Disorders Association (ADRDA) (McKhann et al., 1984; Dubois et al., 2007).

Serial MRI studies were performed to participants from baseline, covering a follow-up period of several years. Each participant was diagnosed at each time point of serial MRI studies.

In the present work, a cohort of 73 subjects were selected from the ADNI database, consisting into 20 subjects with a stable clinical diagnosis of CN state over 24 months of follow up, 13 subjects with a stable clinical diagnosis of MCI over 24 months of follow up, 20 subjects with a stable probable clinical diagnosis of AD over 24 months of follow up, and 20 subjects with an initial clinical diagnosis of MCI who showed a progression to AD within the 24-months follow-up. The choice of the subjects sample was guided by the desire to include all MCI subjects stable at 24 months follow-up (13) and all subjects converted to Alzheimer's dementia within 24 months (20) maintaining a balanced number in the choice of AD (20) and CN (20) subjects stable at 24 months follow-up, chosen randomly among the available sample

The probable clinical diagnosis at the follow-up interval was considered as the reference standard to estimate the accuracy of the AI model in predicting the risk of AD dementia. Thus, subjects with a stable probable clinical diagnosis of AD or with an initial clinical diagnosis of MCI who showed a progression to AD within the 24-months follow-up were considered as subjects with a high risk level of AD dementia; subjects with a stable clinical diagnosis of CN or subjects with a stable clinical diagnosis of MCI were considered as subjects with a low risk level of AD dementia.

## Brain MRI and neuropsychological studies

For each subject of the cohort considered in the study (73), his/her 3T T1-weighted magnetic resonance (MR) images were obtained from the ADNI data repository and downloaded on a local repository in 3D NIfTI format. According to the ADNI acquisition protocol used for all brain MRI examinations of ADNI subjects (Jack et al., 2008), 3T T1-weighted MR images were acquired with MPRAGE protocol of approximately 10 min duration, thus fully compliant with the MPRAGE MRI brain acquisition protocol required for MR images to be automatically processed by TRACE4AD.

NPA data were also obtained from the ADNI data repository for a subcohort of 60 subjects (NPA scores were not available for 13 subjects). According to the ADNI NPA protocol used for all NPA assessment of ADNI subjects, NPA data included both scores and subscores of seven neuropsychological tests, namely the Rey Auditory Verbal Learning Test (AVLT), the Trail Making Test A-B (TMT A-B), the Clock Test, the Digit Span (DS), the Category Fluency Tests (Animals and Vegetables), the Boston Naming Test (BNT), and the Functional Assessment Questionnaire (FAQ). The full list of NPA scores and subscores used in this study is reported in the Supplementary Table S1. Also in this case, NPA tests and scores were fully compliant with the NPA tests and scores required to be automatically processed by TRACE4AD.

Thus, due to the lack of NPA data for 13 subjects, the cohort of 73 patients had baseline 3T T1 MRI brain studies, the subcohort of 60 patients had both baseline 3T T1 MRI brain studies and NPA scores.

## The AI TRACE4AD tool

TRACE4AD is a stand-alone software based on medical image analytics, data mining and machine learning classifiers. It is based on the association of high risk of AD dementia to the early automatic identification of atrophy features in the brain gray matter (GM) following neuronal death caused by AD. The combination of GM features with measures of early cognitive impairment typical of AD dementia allows a more accurate risk prediction.

With TRACE4AD, brain GM features of a subject are automatically extracted from his/her MRI-T1 images, possibly integrated and mined with his/her cognitive measures (when available), then automatically analyzed by support vector machines, a family of machine learning classifiers that, once properly trained, allow automatically distinguishing between two classes of subjects of interest, in this case:

-subjects being affected by AD dementia or subjects with MCI converting to AD dementia within 24 months vs

-subjects with MCI not converting to AD dementia within 24 months or CN subjects.

TRACE4AD uses support vector machines to learn, from the brain GM features of the above-described subjects, how to find the support vectors fixing the best “hyperplane” able to separate the subjects from the two classes. The training of support vector machines is supervised by the probable clinical diagnosis of the subjects, AD dementia, stable MCI, CN, according to their clinical follow up at least 24 months. Thus, subjects with a stable probable clinical diagnosis of AD or with an initial clinical diagnosis of MCI who show a progression to AD within the 24-months follow-up are considered as subjects with a high risk level of AD dementia; subjects with a stable clinical diagnosis of CN or subjects with a stable clinical diagnosis of MCI are considered as subjects with a low risk level of AD dementia.

A model of support vector machines has been previously trained and tested as a binary classifier to construct the TRACE4AD predictive model based on a set of MRI T1 selected features from the brain of subjects with known stable diagnosis, thus known stable risk (low risk, high risk), called training dataset, with and without a set of NPA measures from the same subjects [Salvatore et al., 2018].

Patterns of morphological abnormalities localized in the temporal pole and medial-temporal cortex of these subject’s brains and acknowledged as neuroimaging biomarkers of clinical progression and evolution of AD were already observed by the trained model on MRI brain studies 24 months before AD diagnosis [Salvatore et al., 2018]. Moreover, neuropsychological predictors mainly including measures of functional abilities, memory and learning, working memory, language, visuoconstructional reasoning, and complex attention, with a particular focus on some of the sub-scores of the FAQ and AVLT tests, when used in combination with the above-mentioned brain predictors improved the TRACE4AD ability in predicting the risk of AD dementia [Salvatore et al., 2018].

When the brain MRI-T1 images of a subject at risk of AD dementia as indicated by neurologists, possibly in combination with the subject’s neuropsychological scores, are

given as input to TRACE4AD, the software uses the fixed hyperplane to classify that MRI-T1 GM image in one of the two classes (subjects being affected by AD dementia or subjects with MCI converting to AD dementia within 24 months vs subjects with MCI not converting to AD dementia within 24 months or CN subjects), assigning the corresponding level of risk (high risk of AD dementia within 24 months, low risk of AD dementia within 24 months).

TRACE4AD has been implemented in a stand-alone software tool that can be accessed by target users on cloud through authorization and authentication (http://www.trace4ad.com), or installed on premise on a dedicated Microsoft Windows workstation. For the on-premise version, the workstation should have the following minimum characteristics: operative system windows 10, processor Intel i5 x86 64, RAM of 8 GB and 30 GB of free space on the hard drive.

## Automatic analysis of brain MRI and neuropsychological studies by TRACE4AD

For each subject of the cohort considered in the study (73), his/her 3D brain MRI-T1 NIfTI study was uploaded in the TRACE4AD tool for automatic processing. For each subject of the subcohort of 60 subjects, his/her NPA scores were reported in the TRACE4AD pdf report (Figure 1) that was uploaded in the TRACE4AD tool. The datasets were associated with the patient ID.

# Summary of the scores of the neuropsychological examination

**Fill in the gray fields** with the scores of the neuropsychological exam and **save the document in PDF format.**
The score of M28, M43 and M54 will be calculated automatically by the software.

## Test 1. Rey Auditory Verbal Learning Test (AVL)

| Measure | Description [Score range] | score |
|---|---|---|
| M1 | First repetition - Total of correct answers [0->15] | 10 |
| M2 | First repetition - Total intrusions [0->X]* | 0 |
| M3 | Second repetition - Total of correct answers [0->15] | 11 |
| M4 | Second repetition - Total intrusions [0->X]* | 0 |
| M5 | Third repetition - Total of correct answers [0->15] | 13 |
| M6 | Third repetition - Total intrusions [0->X]* | 1 |
| M7 | Forth repetition - Total of correct answers [0->15] | 11 |
| M8 | Forth repetition - Total intrusions [0->X]* | 0 |
| M9 | Fifth repetition - Total of correct answers [0->15] | 13 |
| M10 | Fifth repetition - Total intrusions [0->X]* | 0 |
| M11 | Sixth repetition (B list) - Total of correct answers [0->15] | 7 |
| M12 | Sixth repetition (B list) - Total intrusions [0->X]* | 1 |
| M13 | Recall Test - Total of correct answers [0->15] | 11 |
| M14 | Recall Test - Total intrusions [0->X]* | 0 |
| M15 | Recognition – Total of recognized words [0->15] | 11 |
| M16 | Recognition – Total of false recognitions [0->31] | 0 |

** X = any number.*

## Test 2. Trail Making Test A-B (TMT A-B)

| Measure | Description [Score range] | Score |
|---|---|---|
| M17 | TMT A - Time to complete (in seconds) [0-150] | 31 |
| M18 | TMT A - Number of errors linking numbers [0->X]* | 0 |
| M19 | TMT A - Number of omission errors of numbers [0->X]* | 0 |
| M20 | TMT B - Time to complete (in seconds) [0-300] | 65 |
| M21 | TMT B - Number of errors linking numbers [0->X]* | 0 |
| M22 | TMT B - Number of omission errors of numbers [0->X]* | 0 |

** X = any number.*

## Test 3. Clock test

| Measure | Description [Score range] | Score |
|---|---|---|
| M23 | Approximately circular face [0->1]* | 1 |
| M24 | Symmetry and correct order of number placement [0->1]* | 1 |
| M25 | Correctness of numbers (1 to 12) [0->1]* | 1 |
| M26 | Presence of the two hands [0->1]* | 1 |
| M27 | Time indicated at six forty-five in some way [0->1]* | 1 |
| M28 | Total score [0->5] | / |

** 0 = wrong answer, 1 = right answer.*

**Test 4a. Digit Span (DS)**

| Measure | Description [Score range] | Score |
|---|---|---|
| M29 | Length of numbers sequence repeated correctly forward [0->9] | 7 |
| M30 | Length of the number sequence correctly repeated backwards [0->8] | 5 |

**Test 4b. Digit Symbol**

| Measure | Description [Score range] | Score |
|---|---|---|
| M31 | Number of correct symbol-number associations (in 90 seconds) [0->70] | 65 |

**Test 5. Category Fluency Test (Animals and Vegetables)**

| Measure | Description [Score range] | Score |
|---|---|---|
| M32 | Number of recalled animals (in 1 minute) [0->X]* | 25 |
| M33 | Number of perseverations while recalling animals (in 1 minute) [0->X]* | 0 |
| M34 | Number of intrusions while recalling animals (in 1 minute) [0->X]* | 0 |
| M35 | Number of recalled vegetables/fruits (in 1 minute) [0->X]* | 29 |
| M36 | Number of perseverations while recalling vegetables/fruits (in 1 minute) [0->X]* | 0 |
| M37 | Number of intrusions while recalling vegetables/fruits (in 1 minute) [0->X]* | 0 |

** X = any number.*

**Test 6. Boston Naming Test (BNT)**

| Measure | Description [Score range] | Score |
|---|---|---|
| M38 | Number of spontaneously given correct responses [0->30] | 29 |
| M39 | Number of semantic cues given [0->30] | 1 |
| M40 | Number of correct responses following a semantic cue [0->30] | 1 |
| M41 | Number of phonemic cues given [0->30] | 0 |
| M42 | Number of correct responses following a phonemic cue [0->30] | 0 |
| M43 | Total number of correct answers [0->30] | / |

**Test 7. Functional Assessment Questionnaire (FAQ)**

| Measure | Description [Score range] | Score |
|---|---|---|
| M44 | Writing checks, paying bills, or balancing checkbook [0->5]* | 0 |
| M45 | Assembling tax records, business affairs, or other papers [0->5]* | 0 |
| M46 | Shopping alone for clothes, household necessities, or groceries [0->5]* | 0 |
| M47 | Playing a game of skill such as bridge or chess, working on a hobby [0->5]* | 0 |
| M48 | Heating water, making a cup of coffee, turning off the stove [0->5]* | 0 |
| M49 | Preparing a balanced meal [0->5]* | 0 |
| M50 | Keeping track of current events [0->5]* | 0 |
| M51 | Paying attention to and understanding a TV program, book, or magazine [0->5]* | 0 |
| M52 | Remembering appointments, family occasions, holidays, medications [0->5]* | 0 |
| M53 | Traveling out of the neighborhood, driving, or take public transportation [0->5]* | 0 |
| M54 | Total score [0->30] | / |

* *0 = normal, 1 = activities that are not usually carried out but that the patient would be able to perform, 2 = activity that is not carried out habitually and which, if performed, would require assistance, 3 = activity carried out independently but with difficulty, 4 = activity carried out, but only with assistance, 5 = complete dependency*

**Fig. 1** TRACE4AD pdf report with scores of NPA examination for a representative patient

Figure 2 shows the TRACE4AD user interface to upload MRI and NPA data.

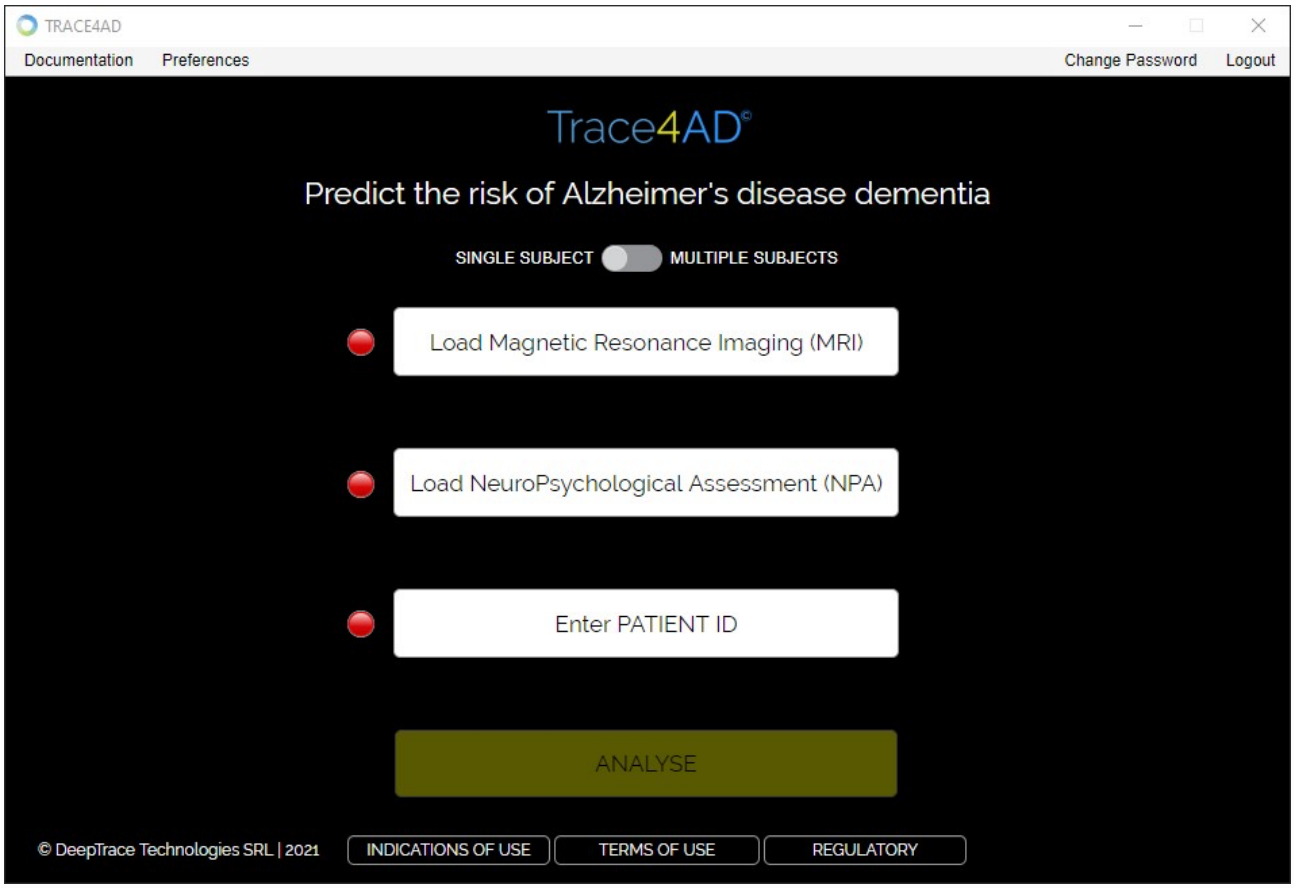



**Fig. 2**. TRACE4AD user interface for the MRI and NPA data upload by users

Once uploaded in TRACE4AD tool, the 3D brain MRI-T1 images of each subject underwent the following preprocessing steps: (1) image re-orientation; (2) cropping; (3) deskulling; (4) image normalization to the Montreal Neurological Institute (MNI) standard space by means of co-registration to the MNI template (MNI152 T1 1 mm brain) (Grabner et al., 2006; O'Hanlon et al., 2013). MR images were segmented into GM tissue probability maps, and smoothed using an isotropic Gaussian kernel with Full Width at Half Maximum (FWHM) ranging from 2 to 12mm$^3$, with a step of 2mm$^3$. After this phase, all MR images resulted to be GM images of size 121 by 145 by 121 voxels.

If the pdf report on the NPA examination was available and uploaded in the TRACE4AD tool, scores of NPA were combined with the GM images and used by the model to automatically classify the subject (with blinded prognosis) as belonging to one of the two risk classes. If not available the model used only GM images. The whole analysis process is completely integrated by the TRACE4AD tool and the user received a report on the results of TRACE4 analysis in a few minutes (from 15 up to 20 minutes).

## Statistical analysis

Data are presented as frequencies and percentage for class characteristics and disease status. Diagnostic and prognostic performances of TRACE4AD were reported in terms of sensitivity, specificity, accuracy, positive predictive value (PPV), negative predictive value (NPV), and corresponding 95% confidence interval (CI), comparing the results of the

TRACE4AD tool (low vs. high risk) *versus* risk level based on the probable clinical diagnosis at the follow-up interval as the reference standard.

## RESULTS

### Population

From the cohort, we retrospectively tested 73 patients with available 3T MRI-T1 brain images and probable clinical diagnosis at baseline and at follow-up. Of these, 20 patients had a baseline probable clinical diagnosis of AD, 33 of MCI, and 20 were CN. Of the 33 patients with baseline diagnosis of MCI, 20 (60.60%) had a probable clinical diagnosis of AD at 24-months follow-up, thus showing high-risk level of AD dementia; 13 (39.39%) had a stable probable clinical diagnosis of MCI at 24-months follow-up, thus showing low-risk level of AD dementia.

The characteristics of the three groups with respect to the baseline probable clinical diagnosis (AD, MCI, CN) for this first cohort are summarized in Table 1.

**Table 1.** Cohort population

| | # patients | Probable clinical diagnosis at baseline | High risk based on 24-months follow-up (%) |
|---|---|---|---|
| Cohort (73 patients) | 20 | AD | 20/20 (100%) |
| | 33 | MCI | 20/33 (60.60%) |
| | 20 | CN | 0/20 (0%) |

The second cohort is a subcohort of the first cohort. As above, we retrospectively tested 60 patients with available 3T MRI-T1 brain images, results of the neuropsychological assessment, and probable clinical diagnosis at baseline and at follow-up. Of these, 11 patients had a baseline probable clinical diagnosis of AD, 33 of MCI, and 16 were CN. Of the 33 patients with baseline diagnosis of MCI, 20 (60.60%) had a probable clinical diagnosis of AD at 24-months follow-up, thus showing high-risk level of AD dementia; 13 (39.39%) had a stable probable clinical diagnosis of MCI at 24-months follow-up, thus showing low-risk level of AD dementia. Figure 5 reports a summary of the scores of the NPA examination for a representative patient.

The characteristics of the three groups with respect to the baseline probable clinical diagnosis (AD, MCI, CN) for this second cohort are summarized in Table 2.

**Table 2.** Subcohort population

| | # patients | Probable clinical diagnosis at baseline | High risk based on 24-months follow-up (%) |
|---|---|---|---|
| Subcohort (60 patients) | 11 | AD | 11/11 (100%) |
| | 33 | MCI | 20/33 (60.60%) |
| | 16 | CN | 0/16 (0%) |

Characteristics of the patients included in this multi-center study for both cohort and sub-cohort are summarized in Table 3.

**Table 3**. Characteristics of patients included in the multi-center study

| **Probable clinical diagnosis at baseline** | **Probable clinical diagnosis at 24m follow-up** | **Risk level of AD dementia at baseline based on 24-m follow-up (low risk, high risk)** | **Number of patients** | **Number of involved centers** |
|---|---|---|---|---|
| **COHORT** | | | | |
| AD | AD | High risk | 20 | 11 |
| MCI | AD | | 20 | 12 |
| | MCI | Low risk | 13 | 11 |
| CN | CN | | 20 | 12 |
| **SUB-COHORT** | | | | |
| AD | AD | High risk | 11 | 8 |
| MCI | AD | | 20 | 12 |
| | MCI | Low risk | 13 | 11 |
| CN | CN | | 16 | 11 |

## The AI TRACE4AD tool performance

The performances of TRACE4AD in predicting the risk level of AD dementia are shown in Table 4 for specificity, sensitivity, accuracy, PPV, NPV. For the cohort of 73 patients the performances are based on the analysis of sole 3T MRI-T1 brain images. For the subcohort of 60 patients the performances are based on the analysis of the sole 3T MRI-T1 brain and in combination with NPA scores. Table 5 shows the single-subject risk as predicted by the tool for each subject of the cohort and subcohort.

**Table 4.** Performances of TRACE4AD in predicting the risk level of AD dementia for the cohort of 73 patients and subcohort of 60 patients. 95% confidence interval is indicated in square brackets.

| | | MRI | MRI + NPA |
|---|---|---|---|
| Cohort (73 patients) | Sensitivity | (31/40) 77.50% [61.55 - 89.16] | - |
| | Specificity | (26/33) 78.79% [61.09 - 91.02] | - |
| | Accuracy | (57/73) 78.08% [66.86 - 86.92] | - |
| | PPV | (31/38) 81.58% [65.67 - 92.26] | - |
| | NPV | (26/35) 74.29% [56.74 - 87.51] | - |
| Sub-cohort (60 patients) | Sensitivity | (24/31) 77.42% [58.90 - 90.41] | (27/31) 87.10% [70.17 - 96.37] |
| | Specificity | (22/29) 75.86% [56.46 - 89.70] | (23/29) 79.31% [60.28 - 92.01] |
| | Accuracy | (46/60) 76.67% [63.96 - 86.62] | (50/60) 83.33% [71.48 - 91.71] |
| | PPV | (24/31) 77.42% [58.90 - 90.41] | (27/33) 81.82% [64.54 - 93.02] |
| | NPV | (22/29) 75.86% [56.46 - 89.70] | (23/27) 85.19% [66.27 - 95.81] |

**Table 5.** Single-subject characteristics and risk level of being affected or converting to AD dementia within 24 months as predicted by TRACE4AD for subjects belonging to the cohort of 73 patients and sub-cohort of 60 patients.

| CENTER ID | ADNI ID STUDY | Clinical diagnosis at baseline (Probable AD, MCI, CN) | Clinical diagnosis at follow-up (Probable AD, MCI, CN) | Risk level based on 24-months follow-up | COHORT Risk level predicted by TRACE4AD using MRI | SUBCOHORT Risk level predicted by TRACE4AD using MRI | SUBCOHORT Risk level predicted by TRACE4AD using MRI + NPA |
|---|---|---|---|---|---|---|---|
| 002 | 002_S_0816 | Probable AD | Probable AD | HR | LR | LR | LR |
| 002 | 002_S_5018 | Probable AD | Probable AD | HR | HR | - | - |
| 006 | 006_S_4153 | Probable AD | Probable AD | HR | HR | - | - |
| 006 | 006_S_4192 | Probable AD | Probable AD | HR | HR | - | - |
| 006 | 006_S_4546 | Probable AD | Probable AD | HR | HR | - | - |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 006 | 006_S_4867 | Probable AD | Probable AD | HR | HR | - | - |
| 007 | 007_S_1304 | Probable AD | Probable AD | HR | HR | HR | HR |
| 009 | 009_S_5027 | Probable AD | Probable AD | HR | HR | - | - |
| 009 | 009_S_5037 | Probable AD | Probable AD | HR | HR | - | - |
| 011 | 011_S_4827 | Probable AD | Probable AD | HR | LR | - | - |
| 011 | 011_S_4912 | Probable AD | Probable AD | HR | LR | - | - |
| 018 | 018_S_0633 | Probable AD | Probable AD | HR | HR | HR | HR |
| 023 | 023_S_1262 | Probable AD | Probable AD | HR | HR | HR | HR |
| 023 | 023_S_1289 | Probable AD | Probable AD | HR | HR | HR | HR |
| 031 | 031_S_1209 | Probable AD | Probable AD | HR | HR | HR | HR |
| 032 | 032_S_1101 | Probable AD | Probable AD | HR | HR | HR | HR |
| 067 | 067_S_0812 | Probable AD | Probable AD | HR | HR | HR | HR |
| 067 | 067_S_1185 | Probable AD | Probable AD | HR | HR | HR | HR |
| 067 | 067_S_1253 | Probable AD | Probable AD | HR | LR | LR | HR |
| 131 | 131_S_0691 | Probable AD | Probable AD | HR | HR | HR | HR |
| 002 | 002_S_0954 | MCI | Probable AD | HR | HR | HR | HR |
| 016 | 016_S_0769 | MCI | Probable AD | HR | HR | HR | LR |
| 016 | 016_S_1121 | MCI | Probable AD | HR | HR | HR | HR |
| 016 | 016_S_1138 | MCI | Probable AD | HR | LR | LR | HR |
| 016 | 016_S_1326 | MCI | Probable AD | HR | LR | LR | LR |
| 023 | 023_S_0030 | MCI | Probable AD | HR | LR | LR | HR |
| 023 | 023_S_0855 | MCI | Probable AD | HR | LR | LR | LR |
| 023 | 023_S_1126 | MCI | Probable AD | HR | HR | HR | HR |
| 027 | 027_S_0835 | MCI | Probable AD | HR | HR | HR | HR |
| 031 | 031_S_0568 | MCI | Probable AD | HR | HR | HR | HR |
| 031 | 031_S_1066 | MCI | Probable AD | HR | LR | LR | HR |
| 032 | 032_S_0187 | MCI | Probable AD | HR | HR | HR | HR |
| 037 | 037_S_1225 | MCI | Probable AD | HR | HR | HR | HR |
| 116 | 116_S_0649 | MCI | Probable AD | HR | HR | HR | HR |
| 116 | 116_S_0752 | MCI | Probable AD | HR | HR | HR | HR |
| 128 | 128_S_1148 | MCI | Probable AD | HR | HR | HR | HR |
| 131 | 131_S_1389 | MCI | Probable AD | HR | HR | HR | HR |
| 133 | 133_S_0913 | MCI | Probable AD | HR | HR | HR | HR |
| 141 | 141_S_0915 | MCI | Probable AD | HR | HR | HR | HR |
| 141 | 141_S_1004 | MCI | Probable AD | HR | HR | HR | HR |
| 037 | 037_S_0377 | MCI | MCI | LR | HR | HR | LR |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 041 | 041_S_1418 | MCI | MCI | LR | LR | LR | LR |
| 051 | 051_S_1072 | MCI | MCI | LR | LR | LR | LR |
| 067 | 067_S_0290 | MCI | MCI | LR | HR | HR | HR |
| 068 | 068_S_0802 | MCI | MCI | LR | LR | LR | LR |
| 082 | 082_S_0928 | MCI | MCI | LR | HR | HR | HR |
| 100 | 100_S_0296 | MCI | MCI | LR | LR | LR | HR |
| 128 | 128_S_1088 | MCI | MCI | LR | LR | LR | HR |
| 133 | 133_S_0792 | MCI | MCI | LR | LR | LR | LR |
| 133 | 133_S_1031 | MCI | MCI | LR | HR | HR | LR |
| 137 | 137_S_0800 | MCI | MCI | LR | LR | LR | LR |
| 137 | 137_S_1414 | MCI | MCI | LR | LR | LR | HR |
| 141 | 141_S_1052 | MCI | MCI | LR | LR | LR | LR |
| 002 | 002_S_4213 | CN | CN | LR | LR | - | - |
| 002 | 002_S_4225 | CN | CN | LR | LR | - | - |
| 002 | 002_S_4262 | CN | CN | LR | LR | - | - |
| 002 | 002_S_4270 | CN | CN | LR | LR | - | - |
| 005 | 005_S_0602 | CN | CN | LR | LR | LR | LR |
| 007 | 007_S_1222 | CN | CN | LR | LR | LR | LR |
| 018 | 018_S_0369 | CN | CN | LR | LR | LR | LR |
| 018 | 018_S_0425 | CN | CN | LR | HR | HR | HR |
| 023 | 023_S_0963 | CN | CN | LR | LR | LR | LR |
| 023 | 023_S_1190 | CN | CN | LR | LR | LR | LR |
| 032 | 032_S_0677 | CN | CN | LR | LR | LR | LR |
| 032 | 032_S_1169 | CN | CN | LR | HR | HR | LR |
| 051 | 051_S_1123 | CN | CN | LR | LR | LR | LR |
| 082 | 082_S_1256 | CN | CN | LR | LR | LR | LR |
| 094 | 094_S_1241 | CN | CN | LR | LR | LR | LR |
| 094 | 094_S_1267 | CN | CN | LR | LR | LR | LR |
| 126 | 126_S_0405 | CN | CN | LR | LR | LR | LR |
| 126 | 126_S_0605 | CN | CN | LR | LR | LR | LR |
| 127 | 127_S_0622 | CN | CN | LR | HR | HR | LR |
| 131 | 131_S_1301 | CN | CN | LR | LR | LR | LR |

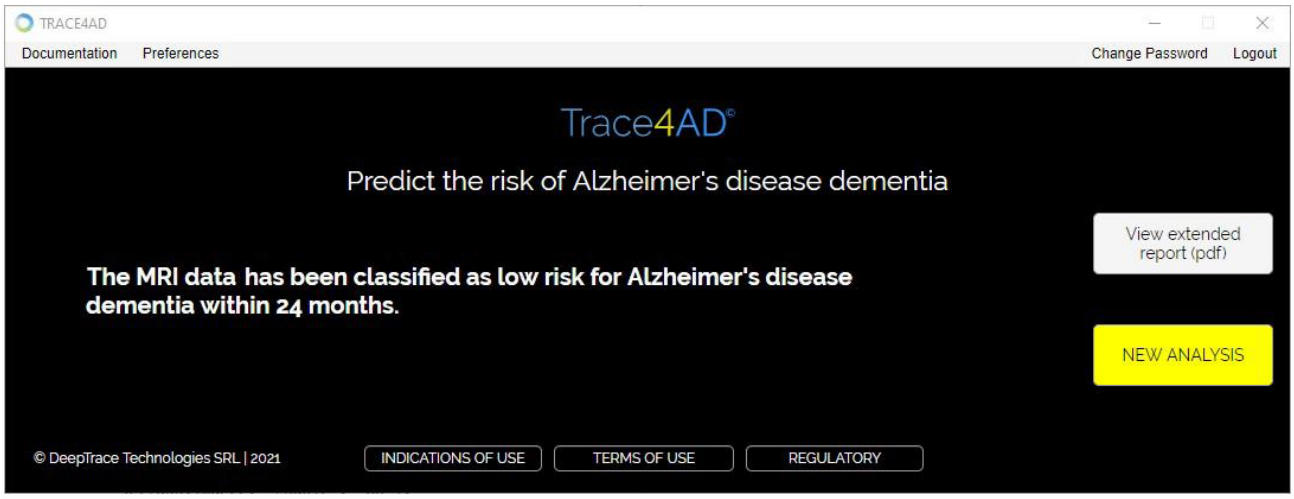



**Fig 3.** TRACE4AD interface as output of the process showing the predicted risk level resulting from the automatic analysis of MRI-T1 data for a representative subject of the cohort

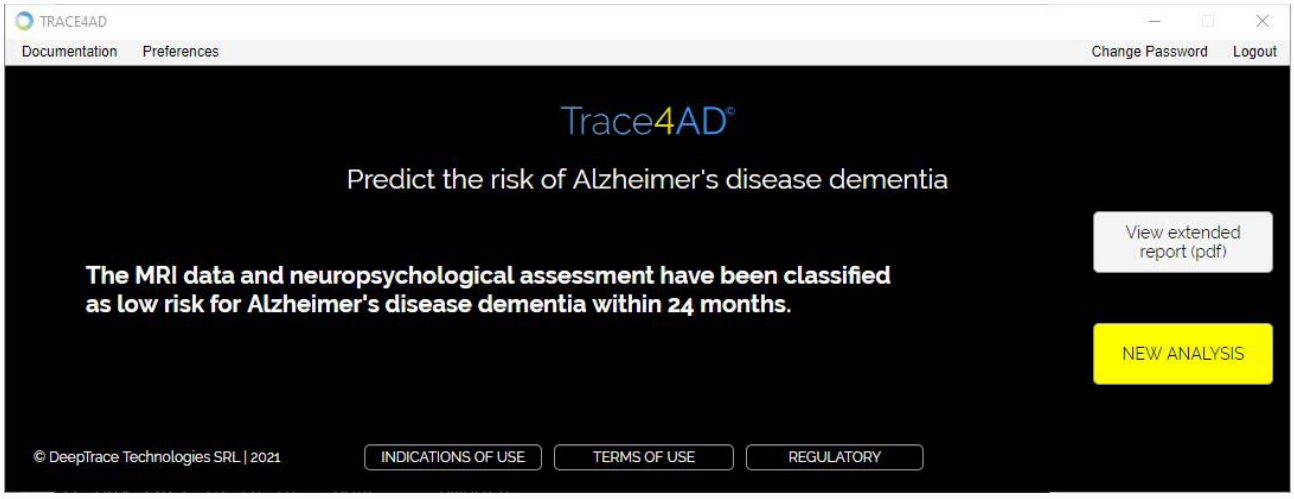


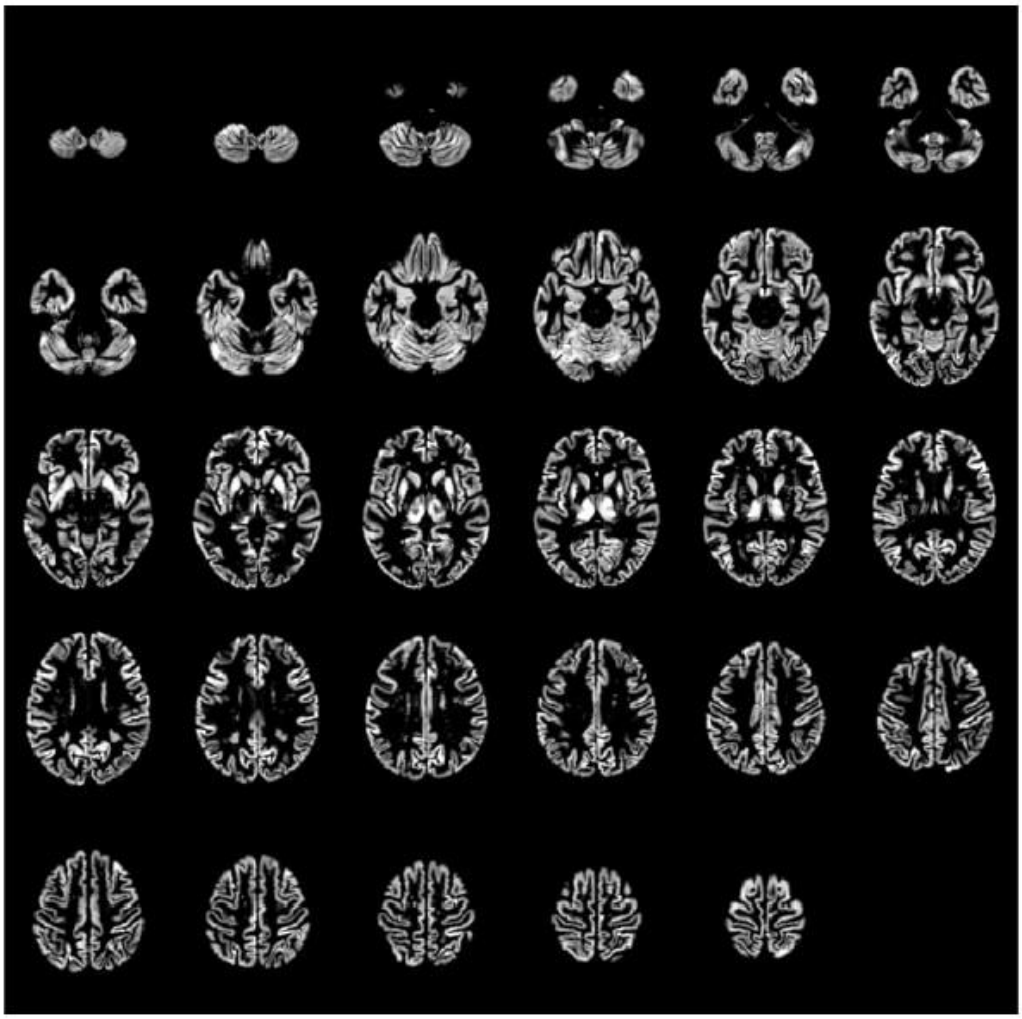

**Fig 4.** TRACE4AD interface as output of the process showing the predicted risk level resulting from the automatic analysis of MRI-T1 and NPA for a representative subject of the subcohort

File-creation date: 12-May-2021 15:57:55

## *TRACE4AD | Report DTT-AD__P001.01MNP*

The MRI data and neuropsychological assessment have been classified as low risk for Alzheimer's disease dementia within 24 months.

***MRI gray-matter image used by TRACE4AD to predict the risk level of Alzheimer's disease dementia (axial view)***



a

***TRACE4AD Imaging Features***

Figure below shows MRI gray-matter imaging features detected by TRACE4AD as best predictors of the risk level of Alzheimer's disease dementia within 24 months. Such imaging features have been revealed by TRACE4AD during training on MRI images of low- and high-risk of Alzheimer's disease dementia and mapped onto a standard 3D-MRI atlas for your convenience, based on significant differences in their expression levels. Significance is expressed in terms of imaging-feature importance and associated to the color scale on the right of the figure.

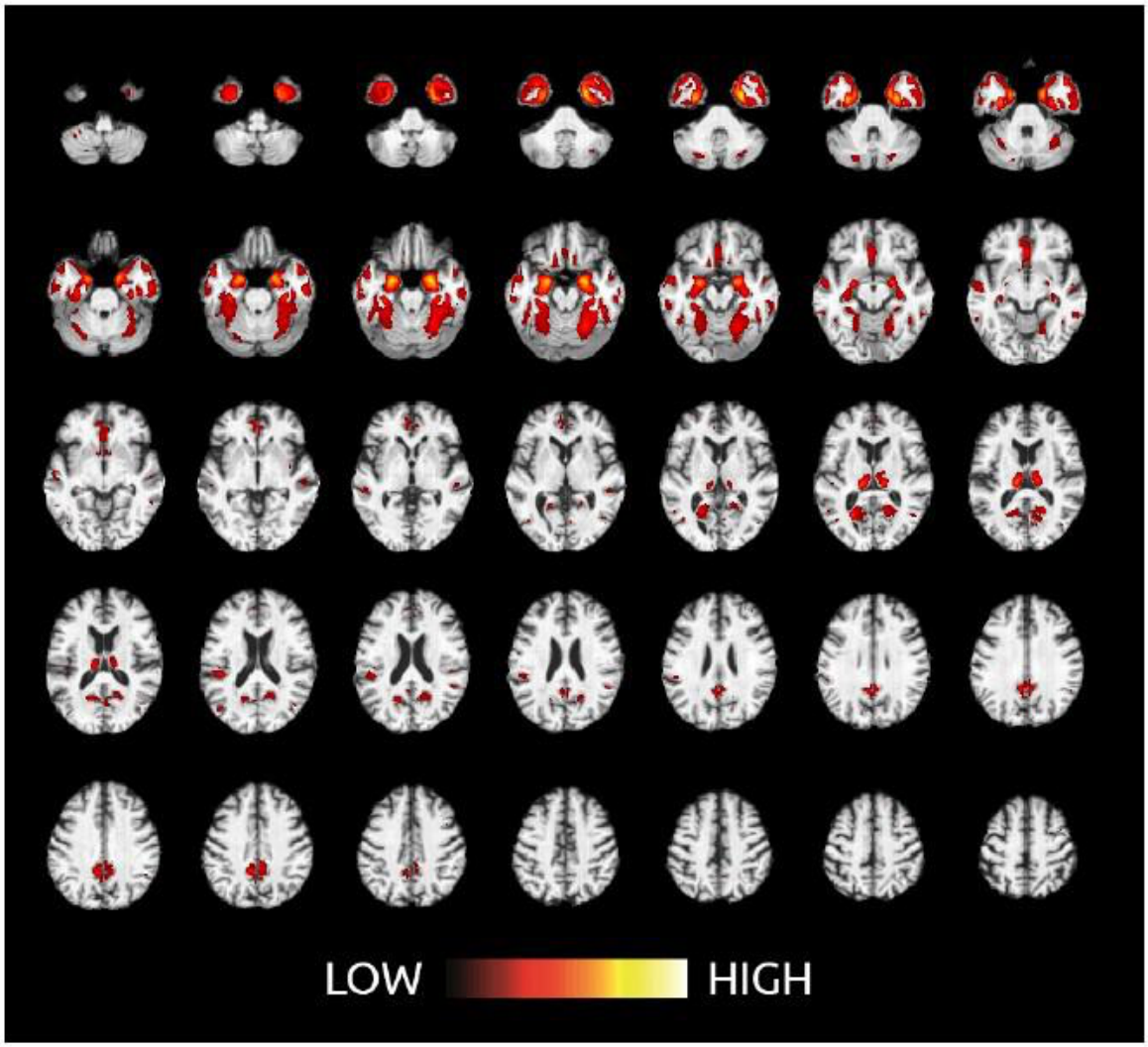



**Fig 5.** TRACE4AD final report showing (a) the single-subject predicted risk level of AD dementia, the MRI single-subject GM image used by TRACE4AD to predict the risk level, and (b) the MRI GM imaging features detected by TRACE4AD as best predictors of the risk level of AD dementia.

## DISCUSSION

In this study we have retrospectively tested and validated a machine learning model, TRACE4AD, in the task to predict the subject level of risk (low risk or high risk) of being

affected by or progressing to AD dementia within 24 months from the baseline examination, by the analysis of the subject's brain MRI study, possibly in combination with a NPA examination. The rationale of such model, as demonstrated by Salvatore et al. in 2018 is that patterns of morphological abnormalities localized in the temporal pole and medial-temporal cortex of a subject's brain and acknowledged as neuroimaging biomarkers of clinical progression and evolution of AD can be already observed by a trained AI reader on MRI brain studies 24 months before AD diagnosis. Moreover, neuropsychological predictors mainly including measures of functional abilities, memory and learning, working memory, language, visuoconstructional reasoning, and complex attention, with a particular focus on some of the sub-scores of the FAQ and AVLT tests, can be used in combination with the above-mentioned brain predictors to improve the TRACE4AD ability in predicting the risk of AD dementia.

For this purpose, MRI T1-weighted brain studies from a retrospective cohort of 73 patients and a subcohort of 60 patients were analysed using the stand-alone software tool TRACE4AD, in which the model was implemented, allowing fully automatic analysis and classification of risk at the level of single-subject. Moreover, follow-up evaluations longer than 24 months and probable clinical diagnosis was available for all patients and allowed to evaluate the performance of the tool in early diagnosis and prognosis within 24 awith a clinical outcome that is considered a reference standard.

Compared to our results on the model published in 2018, that were referred to 1,5 T MRI brain studies [Salvatore et al., 2018], in this work we have tested the model on 3T MRI brain studies, that are less common in radiology departments but are emerging as available instrumental examinations in particular for neurological studies, since recently approved in clinical practise by several national authorities. From the tests provided in this work, our model was shown to be suitable as support to specialized clinicians for the early diagnosis and prognosis of subjects at risk of AD. TRACE4AD showed good performance not only for the subcohort of patients with available both MRI and NPA studies (83% accuracy), with an improvement in sensitivity of 12% when NPA measures were analyzed in combination with MRI, but it was found a valid support also for the cohort with only MRI studies (78%), considering the very complex task for neurologists to predict the prognosis for subject at risk of AD. This performance shows a good generalization of the model with respect to different magnetic fields for MRI acquisitions (1,5 T in Salvatore et a., 2018, 3T in this study). This advantage is warranted by the use of the same pre-processing procedures applied to the input MRI studies, including image re-orientation, deskulling, normalization, coregistriation, segmentation and proper smoothing procedures of MRI brain images, fully provided by TRACE4AD, thus not performed by the operators using external software tools available for similar purpose.

A strength of our model is the full automatic preprocessing and analysis of MRI images and NPA reports. With respect in particular to the examiner reproducibility in the used of TRACE4AD, it should be noted that the dependence on the operator's experience could only affect the administration of the NPA tests to subjects that are precisely administered with pen and paper mode by neuropsychologists that had to take care to report the score in the TRACE4AD pdf report. However, following the reporting of the test scores in the pdf file and its upload together with the MRI study of the subject, the whole analysis process is completely integrated by the TRACE4AD tool and the user receives a report on the results of TRACE4 analysis in a few minutes.

Furthermore, another advantage of the model is that is able to classify patients into two classes (low risk and high risk) overcoming the problem of 'uncertain' class: the current recommendations for the “uncertain” class problem in AD is to assess the patient by second level imaging, Positron Emission Tomography, that is less common, more expensive and more invasive procedure than MRI study. With our predictive model, the subjects in the low class can be managed without intervention at least for 24 months, while the patients in the high risk class can be more strictly monitored by neurologists, eventually supported with anti-dementia drugs and psychosocial therapies, in particular focused to the enhancement of the cognitive reserve for a slowing of the disease and an improvement in the quality of life. This dichotomy certainly represents important decision support for less experienced neurologists or neuroradiologists examiners in MRI brain studies of patients in an early phase of AD.

Our work has some limitations due to the sample size of the considered cohort: the validation will indeed take advantage of cohorts of larger sample size. Of note, in this work the predictive values were quite balanced among low and high risk of being affected by or converting to AD dementia due to a predefined class prevalence of worst prognosis that was 54.79% (40/73) in the cohort of 73 patients and 51.66% (31/60) in the subcohort of 60 patients. This balanced prevalence is different in the general population [Limongi F et al., 2017], while can be not expected in the same proportion in the population at risk of AD dementia. Moreover, in this study, we did not perform a classification of MRI studies by neuroradiologists readers, as well as a classification studies of NPA by neuropsychologists, blinded with respect to TRACE4AD, into the two risk classes at the baseline time. Given that physicians' experience might be a key issue, for example, in interpreting findings from the subject brain concerning neurodegeneration in early phases, or subtle cognitive decline, this comparison could possibly highlight the fact that less experienced clinicians, such as those performing MRI examinations in centres with limited AD patients’ workflows, might benefit from our model to a greater extent. More so, it might be interesting to review whether having a human read parallel to the automatic classification system could possibly lead to an even higher accuracy, leaning towards a double-read system combining AI extracted features and human insight.

## CONCLUSIONS

TRACE4AD software is suitable to provide the subject level of risk (low risk or high risk) of being affected by or progressing to AD dementia within 24 months from the baseline TRACE4AD-processed examination, i.e., the subject's brain MRI investigation, possibly in combination with a NPA examination. Tested on a cohort of patients from several centres using different 3T MRI systems, the tool showed 78% accuracy [67-87%, 95% confidence interval], 78[62-89]% sensitivity, 79[61-91]% specificity, 82[66-92]% PPV, 74[57-88]% NPV based on MRI analysis only (whole cohort of 73 patients), and 83[71-92]% accuracy, 87[70-96]% sensitivity, 79[60-92]% specificity, 82[65-93]% PPV, 85[66-96]% NPV, based on the combined analysis of MRI and NPA (subcohort of 60 patients). However, it should be noted that TRACE4AD can be used as a support to the neurologists in their diagnosis and prognosis, and as a support to the radiology specialists with neuroradiological expertise in their reporting of brain MRI investigations, who have the sole decision-making responsibility.

## List of abbreviations

| | |
|---|---|
| 3T | 3 Tesla |
| AD | Alzheimer's Disease |
| ADNI | Alzheimer's Disease Neuroimaging Initiative |
| ADRDA | Alzheimer's Disease and Related Disorders Association |
| AI | Artificial intelligence |
| AVLT | Rey Auditory Verbal Learning Test |
| BNT | Boston Naming Test |
| CDR | Clinical Dementia Rating |
| CN | Cognitively normal |
| DS | Digit Span |
| FAQ | Functional Assessment Questionnaire |
| FWHM | Full width at half maximum |
| GM | Gray matter |
| HR | High risk |
| LR | Low risk |
| MCI | Mild cognitive impairment |
| MMSE | Mini-Mental State Examination |
| MNI | Montreal Neurological Institute |
| MPRAGE | Magnetization-prepared rapid acquisition with gradient echo |
| MR | Magnetic resonance |
| MRI | Magnetic Resonance Imaging |
| NIfTI | Neuroimaging Informatics Technology Initiative |
| NINCDS | National Institute of Neurological and Communicative Disorders and Stroke |
| NPA | Neuropsychological assessment |
| NPV | Negative predictive value |
| PPV | Positive predictive value |
| RAM | Random-access memory |
| TMT A-B | Trail Making Test A-B |

## Consent for publication

Not applicable.

## Availability of data and materials

The datasets used and/or analysed during the current study are available from the ADNI public data repository.

## Competing interests

Isabella Castiglioni and Matteo Interlenghi own DeepTrace Technologies S.R.L shares. Christian Salvatore is the CEO of DeepTrace Technologies S.R.L. DeepTrace Technologies S.R.L is a spin-off of Scuola Universitaria degli Studi Superiori IUSS, Pavia, Italy. The other authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

## Funding

No funding source was obtained.

## Authors' contributions

Conceptualization, I.C., C.S.; methodology and formal analysis, M.I., C.S.; data curation, M.I., C.S..; software C.S., writing, I.C., C.S.; review and editing, I.C., C.S., M.I.

## References

Grabner G, Janke A L, Budge M M, Smith D, Pruessner J, and Collins D L. Symmetric atlasing and model based segmentation: an application to the hippocampus in older adults. Med. Image Comput. Comput. Assist. Interv 2006:9, 58–66.

Jack C R, Bernstein M A, Fox N C, Thompson P, Alexander G, Harvey D, et al. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. J. Magn. Reson. Imaging 2008:27, 685–691. doi: 10.1002/jmri.21049

Limongi F, Siviero P, Noale M, Gesmundo A, Crepaldi G, Maggi S, for the Dementia Registry Study Group. Prevalence and conversion to dementia of Mild Cognitive Impairment in an elderly Italian population. 2017 Aging Clinical and Experimental Research 2017 29:361–370

O'Hanlon E, Newell F N, and Mitchell K J. Combined structural and functional imaging reveals cortical deactivations in graphemecolor synaesthesia. Front. Psychol. 2013 4:755. doi: 10.3389/fpsyg.2013.00755

Salvatore C, Cerasa A, Castiglioni I. MRI Characterizes the Progressive Course of AD and Predicts Conversion to Alzheimer's Dementia 24 Months Before Probable Diagnosis. Front Aging Neurosci. 2018 May 24;10:135. doi: 10.3389/fnagi.2018.00135. eCollection 2018.