

Published in final edited form as:

Am J Alzheimers Dis (Columbia). 2013 ; 2(1): 12–28. doi:10.7726/ajad.2013.1002.

Variation in Variables that Predict Progression from MCI to AD Dementia over Duration of Follow-up

Shanshan Li^{1,*}, Ozioma Okonkwo², Marilyn Albert³, and Mei-Cheng Wang¹

¹Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

²Department of Medicine, University of Wisconsin, Madison, Wisconsin, USA

³Department of Neurology, Johns Hopkins School of Medicine, Baltimore, Maryland, USA

Abstract

The purpose of this paper is to investigate the relative utility of using neuroimaging, genetic, cerebrospinal fluid (CSF), and cognitive measures to predict progression from mild cognitive impairment (MCI) to Alzheimer's disease (AD) dementia over a follow-up period. The studied subjects were 139 persons with MCI enrolled in the Alzheimer's Disease Neuroimaging Initiative. Predictors of progression to AD included brain volume, ventricular volume, hippocampal volume, APOE $\epsilon 4$ two alleles, $A\beta_{42}$, p-tau₁₈₁, p-tau₁₈₁/ $A\beta_{42}$, memory, language, and executive function. We employ a combination of Cox regression analyses and time-dependent receiver operating characteristic (ROC) methods to assess the prognostic utility and performance stability of candidate biomarkers. In a demographic-adjusted multivariable Cox model, seven measures—brain volume, hippocampal volume, ventricular volume, APOE $\epsilon 4$ two alleles, $A\beta_{42}$, Memory composite, Executive function composite—predicted progression to AD. Time-dependent ROC revealed that this multivariable model had an area under the curve of 0.832, 0.788, 0.794, and 0.757 at 12, 18, 24, and 36 months respectively. Supplemental Cox models with time of origin set differentially at 12, 18, 24 and 36 months showed that six measures were significant predictors at 12 months whereas only memory and executive function predicted progression to AD at 18 and 24 months. The authors concluded that baseline volumetric MRI and cognitive measures selectively predict progression from MCI to AD, with cognitive measures remaining predictive even late in the follow-up period. These findings may inform case selection for AD clinical trials.

Keywords

Alzheimer's disease; Cox models; Mild cognitive impairment; memory; ROC analysis

1. Introduction

Mild cognitive impairment (MCI) is a clinical syndrome that can presage Alzheimer's dementia (AD) (Petersen and Negash, 2008). However, MCI patients as a group can be heterogeneous. Both clinical and epidemiological studies have shown that a variety of pathophysiological processes may explain the clinical syndrome among those who progress to dementia. Moreover, some individuals remain stable while others revert to normal cognitive functioning suggesting heterogeneity of outcome in this population (Gauthier et

© The author(s) 2013.

*Corresponding: shli@jhsph.edu. Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA.

al., 2006; Petersen et al., 2009). Accordingly, a great deal of effort has been dedicated to identifying a biomarker, or combinations of biomarkers, that increase the likelihood that an AD pathophysiological process is present and will likely lead to dementia within a few years (Devanand et al., 2008; Petersen and Negash, 2008; Mattsson et al., 2009)

Investigations of the risk of progressing from MCI to AD dementia have largely focused on measures from the following categories: a) cognition (Dickerson et al., 2007), b) genetics (Petersen et al., 1995), c) neuroimaging (Jack Jr et al., 2005), or d) cerebrospinal fluid (CSF) (Mattsson et al., 2009). The majority of these studies have examined the predictive utility of such measures independently. While some studies have examined these measures in tandem, most have not used statistical approaches that account for censoring, and take into account the passage of time over which the prediction is being made (Devanand et al., 2008; Fleisher et al., 2008; Landau et al., 2010; Ewers et al., 2012; Zhang et al., 2012). Thus, relatively little is known about the comparative utility of these measures in determining the risk of incident AD dementia in persons with MCI, especially whether there is a change in their relative utility over time. There remains a need to develop a parsimonious and robust multidomain predictive model of the risk of progression to AD dementia among persons with MCI, taking progression of time into account. Such a model would potentially be useful for the design and evaluation of MCI clinical trials.

In this paper, we employ a combination of Cox regression analyses and time-dependent receiver operating characteristic (ROC) methods to investigate the utility of various candidate measures, alone or in combination, for predicting progression to AD dementia in a large, well-characterized group of persons with MCI. These statistical methods were selected because they provide a method for choosing among a set of biomarkers the ones that might be most predictive, but also permit a comparison of their predictive utility with progression of disease. The specific measures chosen cover the domains of cognitive function, genetics, neuroimaging, molecular markers in CSF, and relevant demographic variables (i.e., age, education, and sex) were employed as covariates given their potential effect on disease progression in AD (Fleisher et al., 2007)

2. Materials and Methods

2.1 Participants

The analyses presented here were based on data from the Alzheimer's Disease Neuroimaging Initiative (ADNI; <http://adni.loni.ucla.edu/>). The ADNI was launched in 2004 by the National Institute on Aging and other entities (see Acknowledgments) as a 5-year public-private partnership. Enrollment target was 800 participants—200 normal controls, 400 patients with amnesic MCI, and 200 patients with mild AD—at 58 sites in the United States and Canada. Participants were enrolled on a rolling basis, and evaluated at six-month intervals. Further details about ADNI, including participant selection procedures and complete study protocol, have been presented elsewhere (Weiner et al., 2010), and may be found online at: <http://www.alzheimers.org/clinicaltrials/fullrec.asp?PrimaryKey=208>

The present analyses included all 139 amnesic MCI patients who had complete cognitive, genetic (i.e., apolipoprotein E [APOE]), magnetic resonance imaging (MRI), and cerebrospinal fluid (CSF) data when data download occurred in March 2012. Informed consent was obtained from study participants, and the study was approved by the local institutional review board at each participating site.

2.2 Cognitive measures

Participants completed the following neuropsychological measures: Story A from the Logical Memory test (Wechsler, 1987), Category fluency test (Animals and Vegetables)

(Butters et al. 1987), the Boston Naming test (Kaplan et al., 1983), Clock drawing test (Goodglass and Kaplan, 1983), Digit Span test (Wechsler, 1981), Digit Symbol Substitution test (Wechsler, 1981), Trail Making test (A and B) (Reitan, 1958), and the Rey Auditory Verbal Learning test (Rey, 1964). For our analyses, participants' scores on these measures were represented as composite measures.

First, we transformed the raw scores to z-scores using group means and standard deviations at baseline. Next, we averaged these z-scores to derive composite scores that represented five cognitive domains: Memory, Language, Executive function, Spatial ability, and Attention. The Memory composite consisted of scores on the Logical Memory and Rey Auditory Verbal Learning tests. The Language composite consisted of scores on the Category fluency and Boston Naming tests. The Executive function composite consisted of scores on the Trail Making, Digit Symbol Substitution, and spontaneous Clock drawing tests. The Spatial ability composite consisted of scores on the copy trial of the Clock drawing test. The Attention composite consisted of scores on the Digit Span test.

2.3 APOE genotyping

Participants provided Ethylenediaminetetraacetic acid (EDTA) blood samples, on which APOE genotyping was performed by the ADNI Biomarker Core, using TaqMan assays as previously described (Shaw et al., 2009).

2.4 MRI procedure

Participants underwent 1.5T structural brain MRI scans using a standardized protocol (Jack Jr et al., 2008). In our analyses, we used total brain, ventricular and bilateral hippocampal volumes. Brain and ventricular volumes were derived using the boundary shift integral technique (Freeborough et al., 1997; Leung et al., 2010) and hippocampal volumes were obtained using a commercially available high dimensional brain-mapping tool (Medtronic Surgical Navigation Technologies, Louisville, CO) (Hsu et al., 2002; Schuff et al., 2009). All three measures were deemed to have passed quality control by the respective labs that processed the raw data (Nick Fox [UCL London, for brain and ventricular measures] and Michael Weiner [UCSF, for hippocampal measure]).

2.5 CSF procedure

At each ADNI site, subjects underwent lumbar puncture in the morning following an overnight fast. The extracted CSF samples were then shipped overnight to the ADNI Biomarker Core where aliquots (0.5 ml) were prepared from the samples and stored in bar code-labeled polypropylene vials at -80°C . T-tau, $\text{A}\beta_{42}$, and p-tau₁₈₁ were assayed from these aliquots using the multiplex xMAP Luminex platform (Luminex Corp., Austin, TX) with Innogenetics immunoassay kit-based reagents (INNO-BIA AlzBio3; Ghent, Belgium; for research use-only reagents) (Shaw et al., 2009).

2.6 Data analyses

All data analyses were performed using R version 2.15.1. The basic analytic approach involved a linking of Cox regression models and time-dependent receiver operating characteristic (ROC) methods to assess the prognostic utility and performance stability of candidate biomarkers. Variables that emerged significantly in univariate Cox models were considered as covariates of interest for multivariate Cox models. Two approaches were implemented for Cox model analysis with further variable selection, adjusting for age, education and gender. One approach used p-value criterion, where predictors with p-value greater than 0.05 were removed from the model. The other approach used LASSO (Tibshirani, 1996; Tibshirani, 1997), a penalized likelihood based method that is designed

for purpose of variable reduction. Risk scores—for use in subsequent ROC analyses—were derived by fitting a new Cox model that included selected variables from p-value criterion or LASSO.

Although ROC analysis has been previously utilized in interrogating AD-related data, including data from ADNI (De Meyer et al., 2010; Landau et al., 2010), neither the complication of censoring nor the component of time (given the longitudinal nature of these studies) has been optimally handled. Therefore, the results of these analyses might have been biased. Assuming t (year) is the time point of interest, we focused on subjects with follow-up time of at least t years. To account for censoring, we used the reduced sample technique to extract a representative data sample. That is, participants who dropped out before t years were excluded from the analysis. The classic ROC analysis was therefore carried out based on the reduced sample. Although this results in loss of data information (Kaplan and Meier, 1958), this reduced sample technique produces a representative sample that incorporates censoring. To fully utilize information from all available data, we implemented an approach (i.e., time-dependent ROC) that permits an evaluation of the diagnostic performance of biomarkers at all time points of interest.

Time-dependent ROC is an extension of the classic ROC. It was first introduced by Etzioni et al. (1999) to accommodate settings where disease status changes with time. Instead of a binary outcome (e.g., “converter” vs. “non-converter”), conversion status is allowed to change over time. For example, a subject who is unimpaired at an earlier time point can be a case (i.e., impaired) at a later time. Sensitivity is defined as $P(M_i > c | T_i = t)$, and specificity is defined as $P(M_i < c | T_i > t)$, where T_i is the event time, M_i is the risk score, and c is a cut-off point for classification. Thus sensitivity measures the expected fraction of subjects with marker value greater than c among the sub-population of individuals who are diagnosed with AD at time t , while specificity measures the fraction of subjects with marker values less than or equal to c among those whose failure time is greater than t (Heagerty and Zheng, 2005). This dynamic status parallels the multiple contributions that a subject can make to the risk set that underlies the Cox partial likelihood function (Heagerty and Zheng, 2005). Time-dependent ROC analysis and Cox regression form the ideal marriage when the proportional hazard assumption is (approximately) satisfied. The area under the ROC curve (AUC) was calculated to measure the intrinsic ability of the biomarkers to discriminate between the converters and non-converters (Komori and Eguchi, 2010).

Though there is extensive research in identifying biomarkers for Alzheimer’s disease, cognitive measures are still the gold standard for clinical assessment of MCI and AD (Albert et al., 2011). It is useful to evaluate the predictive accuracy of the model with measures selected from cognitive domain only (i.e. Memory composite and Executive function composite) and compare it to the model with measures selected from all domains. For this purpose, we compared the time-dependent ROC curves for the two models and calculated their corresponding AUCs. To test the improvement in prediction performance by adding the additional variables, we tested whether the difference in the AUCs of the two models was statistically significant (Pepe et al., 2013).

To further evaluate the stability of the predictor set derived from our Cox model, we modified Cox regression to allow biomarker combinations to change over time and to identify temporal patterns in biomarker combination for predicting AD. The time-varying effects were evaluated by refitting the Cox model after removing from the dataset, at predetermined time points (e.g., 12, 18, 24 months), participants who had developed AD or been censored by those time points.

3. Results

Table 1 shows some demographic summaries of the studied participants. Table 2 lists the summary statistics for 13 covariates of the studied sample. In univariate Cox models, ten variables were associated with increased risk of progression to AD at 0.1 p-value level, as listed in Table 3. These included brain volume, hippocampal volume, ventricular volume, APOE $\epsilon 4$ 2 alleles, $A\beta_{42}$, p-tau₁₈₁, p-tau₁₈₁/ $A\beta_{42}$, Memory, Language, and Executive function. (Note that dose effect was found for the APOE genotype: the indicator of Apoe $\epsilon 4$ 2 alleles was significant while the indicator of Apoe $\epsilon 4$ 1 allele was not significant.) These measures were then entered into a multivariable Cox model, with age, education, and gender included as covariates. In the multivariable Cox model, brain volume, hippocampal volume, ventricular volume, APOE $\epsilon 4$ 2 alleles, Memory composite, and Executive function composite emerged as significant predictors with p-value less than 0.05. The LASSO method was also performed for variable selection, and the output from LASSO contained one more variable $A\beta_{42}$ in addition to the six variables selected by p-value criterion. The LASSO approach selected the optimal model as a whole, while the p-value criterion selected each variable individually with values of other variables fixed, so the results from LASSO were preferred from perspective of model selection. Note that the variables selected by p-value criterion were a subset of the variables selected by LASSO, which suggested a level of robustness of the model selected by LASSO. The seven variables from LASSO output then entered the multivariate Cox model (see Table 4). The score function derived from the model was employed as the classifier in subsequent ROC analyses and its predictability was studied.

The classifier was first used in a classic ROC analysis. To account for censoring, we chose $t = 24$ month as the cut-off. Of the 139 participants who were MCI at baseline, 22 had dropped out of the study before 24 months. Therefore, they were excluded from the analysis, leaving us with a “reduced sample” of 117 participants, of whom 45 had developed AD by month 24 (“converters”) with the other 72 remaining MCI (“non-converters”). Fig. 1 provides an assessment of the classifier in distinguishing converters vs. non-converters over the whole range of potential cutoff values. The AUC for this curve was 0.82. The ROC curve evaluates biomarker performance at different cutoff values, and different cutoff values yield different true positive rate and false positive rate. Controlling the false positive rate at 0.2 led to a true positive rate of approximately 0.6. A stricter false positive rate of 0.1 resulted in a sensitivity of approximately 0.4.

Fig. 2 displays time-dependent ROC curves to assess the discriminant performance of the model score at different time points. It revealed that the model’s predictive accuracy diminished somewhat with increasing time since baseline. The AUCs at 12, 18, 24 and 36 months were respectively 0.832, 0.788, 0.794 and 0.757. The 95% confidence intervals for the AUCs, calculated using bootstrapping (Efron, 1979), were respectively: (0.826, 0.832) at 12 month, (0.783, 0.788) at 18 month, (0.786, 0.794) at 24 month, and (0.757, 0.757) at 36 month. Therefore, the AUC at 12 month (and AUC at 36 month) was significantly different from AUCs at other time points, but the AUCs at 18 month and 24 month were not statistically different. Controlling the false positive rate at 0.2 led to an incident sensitivity of 0.70 at 12 months, decreasing to 0.61, 0.61, and 0.54 at 18, 24 and 36 months respectively.

Fig. 3 shows the comparison of the ROC curves for the model using cognitive domain only and the model using measures from all domains. It is observed that the ROC curves for the model with all selected covariates dominate the ROC curves for the model with selected cognitive measures only. This is as expected, since adding more significant covariates would increase the predictive accuracy of the model score. However, it is also seen that the selected

cognitive measures only model had adequate prognostic utility: the AUCs at 12, 18, 24 and 36 months were 0.745 (CI: 0.741, 0.746), 0.725 (CI: 0.721, 0.725), 0.740 (CI: 0.732, 0.740) and 0.698 (CI: 0.698, 0.698). Comparing with the 95% confidence intervals for the AUCs of the model with all selected covariates: (0.826, 0.832) at 12 month, (0.783, 0.788) at 18 month, (0.786, 0.794) at 24 month, and (0.757, 0.757) at 36 month, we concluded that, the model with all selected covariates had significantly higher AUC, and additional measures did improve on the predictive utility of the cognitive measures.

To further evaluate the stability of the predictor set derived from our Cox model, we refitted the regression model, choosing as our “time of origin” 12, 18, 24 and 36 months respectively, after excluding those participants who had developed AD or been censored by those time points. Table 5 shows the results of these analyses. The reduced sample size at 12, 18, 24, and 36 months were 123, 102, 85 and 46 respectively. Five measures were significant predictors at 12 month; Memory and Executive function composites were significant at 18 and 24 month, whereas only Memory composite remained significant at all four time points. The sign for each predictor’s regression estimate remained unchanged at all time points, suggesting that the predictors’ mean effect was essentially stable over time even if attenuated.

Fig. 4 displays ROC curves for the time-varying model scores. At each time point, we compared the diagnostic performance of the model score derived from the time-invariant Cox model fitted using all available data (“model 1,” listed in Table 4 and Fig. 2) to the model scores derived from the time-varying Cox models fitted using “updated” data sets (“models 2–5,” listed in Tables 5 respectively). At 12 months, the AUCs for model 1 and model 2 were 0.832 (CI: 0.826, 0.832) and 0.851 (CI: 0.845, 0.851) respectively. At 18 months, the AUCs for model 1 and model 3 were 0.788 (CI: 0.783, 0.788) and 0.784 (CI: 0.777, 0.784) respectively. At 24 months, the AUCs for model 1 and model 4 were 0.794 (CI: 0.786, 0.794) and 0.912 (CI: 0.904, 0.912) respectively. At 36 months, the AUCs for model 1 and model 5 were 0.757 (CI: 0.757, 0.757) and 0.907 (CI: 0.907, 0.907) respectively. Thus, the diagnostic performance of the combination derived from model 1 was optimal at earlier time points but reduced at later times. This suggested that the proportional hazard assumption did not fully hold (i.e., the effects of the 4 measures were not constant over time) and thus a time-varying coefficient model was preferred.

4. Discussion

In this paper, we utilized a combination of demographics-adjusted Cox regression analyses and time-dependent ROC curves to examine univariate and multivariable predictors of progression from MCI to AD dementia. This permitted an examination of the relative utility of these measures over time. In our univariate models, we found that (i) all three MRI measures, brain, ventricular and hippocampal volumes predicted progression to AD, (ii) APOE $\epsilon 4$ 2 alleles predicted progression to AD, (iii) among the CSF measures, $A\beta_{42}$, p-tau₁₈₁ and p-tau₁₈₁/ $A\beta_{42}$, predicted progression to AD, and finally (iv) among the cognitive composite measures, Memory, Language, and Executive function composites, predicted progression to AD. Two model selection approaches were applied to those significant biomarkers identified by univariate Cox models, and the results only differed by one variable $A\beta_{42}$. The output from LASSO was preferred, as it was optimal from perspective of model selection. The selected predictors included brain volume, ventricular volume, hippocampal volume, APOE $\epsilon 4$ 2 alleles, $A\beta_{42}$, Memory, and Executive. An ancillary demographics-adjusted multivariable Cox model was then fitted, using only the selected variables, for the purpose of obtaining risk scores that were then employed in time-dependent ROC analyses that assessed the discriminant performance of the regression model at different time points. These ROC analyses revealed that our multivariable model had high

performance AUCs of 0.832, 0.788, 0.794 and 0.757 at 12, 18, 24, and 36 months respectively. Of note, by accounting for censoring and the passage of time, our ROC analysis modified the conventional ROC approach and resulted in more informative measures for evaluating markers.

Although prior studies have examined biomarker combinations as predictors of incident AD among persons with MCI (Fleisher et al., 2008; Landau et al., 2010; Ewers et al., 2012; Zhang et al., 2012), there have been very few references on using time-dependent ROC analyses to assess model performance across time. Derby et al. (2013) applied time-dependent ROC analyses in the investigation of progression to AD dementia using data from the Einstein Aging Study (EAS), where they studied the predictive ability of a few cognitive measures and they assumed proportional hazards model for these measures. In our work, we considered measures from CSF and neuroimaging domains in addition to the cognitive domain and we found that additional measures did improve on the predictive utility of the cognitive measures. Furthermore, we examined the stability of biomarkers over time by refitting the multivariable Cox regression model at specific time points of interest (i.e., 12, 18, and 24 months) after discarding data from MCI cases who had already received a diagnosis of AD or had been censored at those respective time points. These analyses revealed that at 12 months, five measures (i.e., brain volume, ventricular volume, APOE $\epsilon 4$ 2 alleles, Memory composite, and Executive Function composite) remained significant predictors of prospective progression to AD. At 18 and 24 months, only Memory and Executive Function were significant predictors of progression to AD in this sample. Finally, at 36 months, only Memory remained significant. These findings suggest that whereas MRI and cognitive measures might generally be equally useful in prognosticating risk of incident AD among persons with MCI, cognitive measures retain their predictability for longer time during the disease progression, whereas MRI measures become less predictive with the passage of time. This supposition would be consistent with a recently proposed hypothetical model of AD pathological cascade that situates cognitive impairments in the later phases of the AD spectrum, with MRI abnormalities occurring earlier (Jack Jr et al., 2010).

Baseline measures of brain and, especially, hippocampal volumes have been shown in past studies to be useful in predicting progression from MCI to AD dementia (Jack Jr et al., 1999; Devanand et al., 2007; Smith et al., 2008) even when the statistical model included other measures. Similarly, impairments in neuropsychological measures of episodic memory and executive functions are recognized as the most common features of AD even at its earliest stages (Grady et al., 1988; Storandt and Hill, 1989). Not surprisingly then, several studies have demonstrated that baseline performance on measures that tap memory and executive functions are effective in predicting the transition from a state of mild impairments to clinically-manifest AD dementia (Albert et al., 2001; Dickerson et al. 2007; Fleisher et al. 2007; Landau et al., 2010; Ewers et al., 2012). Thus, overall, our finding that progression to AD dementia in persons with MCI varied as a function of baseline brain, ventricular and hippocampal volumes, baseline A β_{42} , APOE genotype, and scores on composite measures of Memory and Executive function is consistent with prior reports in the literature. Furthermore, our findings suggest that combinations of measures that reflect putatively earlier pathognomonic AD processes (i.e., hippocampal atrophy and memory loss) and those that perhaps reflect more extensive disease (i.e., whole brain atrophy and executive dysfunction) might be optimal for predicting which persons with MCI would go on to develop clinical AD.

As noted earlier, a major strength of this paper is the use of statistical methodologies—specifically time-dependent ROC curves—that allowed us to examine progression to AD in a manner that has not been accomplished by prior studies. Our approach revealed that, overall, the performance of our predictor set remained stable over time although there was

some tapering of effect at later time periods. In addition, it demonstrated that cognitive tests of memory and executive function retain their predictive utility even late in the disease process whereas MRI measures become less useful with the passage of time. It would be of interest to see whether our predictor set replicates when additional biomarkers such as measures that tag inflammation, neuronal injury, glial activation, or oxidative stress are included in the model.

At this time, there are no federally-approved therapeutics for delaying the onset of dementia in persons with mild impairments. However, several targets are being evaluated in clinical trials (Neugroschl and Sano, 2009). The measures we have identified as predictive of progression from MCI to AD could be useful in case selection for these trials so that the trials are populated with persons who truly have underlying AD pathophysiology and are at greater risk of developing a dementia over a reasonable time frame. These measures could also play a role in identifying MCI cases who would benefit from treatment, when such treatments become available.

The caption should appear directly below the picture. Picture should be one space above/below the text. Labeling should always go as follows: Fig 1. Figure should be abbreviated with the proper number following. The period should always follow the number NOT the abbreviated word.

Acknowledgments

Data collection and sharing for this project was funded by the ADNI (Principal Investigator: Michael Weiner; NIH grant U01AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering (NIBIB), and through generous contributions from the following: Pfizer Inc., Wyeth Research, Bristol-Myers Squibb, Eli Lilly and Company, GlaxoSmithKline, Merck & Co. Inc., AstraZeneca AB, Novartis Pharmaceuticals Corporation, Alzheimer's Association, Eisai Global Clinical Development, Elan Corporation plc, Forest Laboratories, and the Institute for the Study of Aging, with participation from the U.S. Food and Drug Administration. Industry partnerships are coordinated through the Foundation for the National Institutes of Health. The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the AD Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory of Neuro Imaging at the University of California, Los Angeles.

These analyses were supported, in part, by U01 AG033655 (Principal Investigator: Marilyn S. Albert; NIA grant).

References

- Albert MS, Moss B, Tanzi R, Jones K. Preclinical prediction of AD using neuropsychological tests. *J. Int. Neuropsychol. Soc.* 2001; 7:631–639. <http://dx.doi.org/10.1017/S1355617701755105> PMID: 11459114. [PubMed: 11459114]
- Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, et al. The diagnosis of mild cognitive impairment due to Alzheimers disease: Recommendations from the National Institute on Aging-Alzheimers Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's and Dementia.* 2011; 7:270–279. <http://dx.doi.org/10.1016/j.jalz.2011.03.008> PMID:21514249 PMCID:3312027.
- Butters N, Granholm E, Salmon DP, Grant I, Wolfe J. Episodic and semantic memory: a comparison of amnesic and demented patients. *J. Clin. Exp. Neuropsychol.* 1987; 9:479–497. <http://dx.doi.org/10.1080/01688638708410764> PMID:2959682. [PubMed: 2959682]
- De Meyer G, Shapiro F, Vanderstichele H, Vanmechelen E, Engelborghs S, De Deyn PP, Coart E, Hansson O, Minthon L, Zetterberg H, et al. Diagnosis-independent Alzheimer disease biomarker signature in cognitively normal elderly people. *Archives of Neurology.* 2010; 67:949. <http://dx.doi.org/10.1001/archneurol.2010.179> PMID:20697045 PMCID:2963067. [PubMed: 20697045]
- Derby CA, Burns LC, Wang C, Katz MJ, Zimmerman ME, L'Italien G, Guo Z, Berman RM, Lipton RB. Screening for pre- dementia AD: time-dependent operating characteristics of episodic memory

- tests. *Neurology*. 2013; 80:1307–1314. <http://dx.doi.org/10.1212/WNL.0b013e31828ab2c9> PMID: 23468542. [PubMed: 23468542]
- Devanand D, Pradhaban G, Liu X, Khandji A, De Santi S, Segal S, Rusinek H, Pelton G, Honig L, Mayeux R, et al. Hippocampal and entorhinal atrophy in mild cognitive impairment Prediction of Alzheimer disease. *Neurology*. 2007; 68:828–836. <http://dx.doi.org/10.1212/01.wnl.0000256697.20968.d7> PMID:17353470. [PubMed: 17353470]
- Devanand D, Liu X, Tabert MH, Pradhaban G, Cuasay K, Bell K, de Leon M, Doty RL, Stern Y, Pelton GH. Combining early markers strongly predicts conversion from mild cognitive impairment to Alzheimer's disease. *Biol. Psychiatry*. 2008; 64:871–879. <http://dx.doi.org/10.1016/j.biopsych.2008.06.020> PMID:18723162 PMCID:2613777. [PubMed: 18723162]
- Dickerson BC, Sperling RA, Hyman BT, Albert MS, Blacker D. Clinical prediction of Alzheimer disease dementia across the spectrum of mild cognitive impairment. *Arch. Gen. Psychiatry*. 2007; 64:1443–1450. <http://dx.doi.org/10.1001/archpsyc.64.12.1443> PMID:18056553 PMCID:2581771. [PubMed: 18056553]
- Efron B. Bootstrap methods: Another look at the jackknife. *Ann. Statist.* 1979; 7:1–26. <http://dx.doi.org/10.1214/aos/1176344552>.
- Etzioni R, Pepe M, Longton G, Hu C, Goodman G. Incorporating the time dimension in receiver operating characteristic curves: a case study of prostate cancer. *Medical Decision Making*. 1999; 19:242–251. <http://dx.doi.org/10.1177/0272989X9901900303> PMID:10424831. [PubMed: 10424831]
- Ewers M, Walsh C, Trojanowski JQ, Shaw LM, Petersen RC, Jack CR Jr, Feldman HH, Bokde AL, Alexander GE, Scheltens P, et al. Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance. *Neurobiology of Aging*. 2012; 33:1203–1214. <http://dx.doi.org/10.1016/j.neurobiolaging.2010.10.019> PMID:21159408 PMCID:3328615. [PubMed: 21159408]
- Fleisher A, Sowell B, Taylor C, Gamst A, Petersen R, Thal L, et al. Clinical predictors of progression to Alzheimer disease in amnesic mild cognitive impairment. *Neurology*. 2007; 68:1588–1595. <http://dx.doi.org/10.1212/01.wnl.0000258542.58725.4c> PMID:17287448. [PubMed: 17287448]
- Fleisher A, Sun S, Taylor C, Ward C, Gamst A, Petersen R, Jack CR Jr, Aisen P, Thal L, et al. Volumetric MRI vs clinical predictors of Alzheimer disease in mild cognitive impairment. *Neurology*. 2008; 70:191–199. <http://dx.doi.org/10.1212/01.wnl.0000287091.57376.65> PMID: 18195264. [PubMed: 18195264]
- Freeborough PA, Fox NC. The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI. *IEEE Trans. Med. Imaging*. 1997; 16:623–629. <http://dx.doi.org/10.1109/42.640753> PMID:9368118. [PubMed: 9368118]
- Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K, Broich K, Belleville S, Brodaty H, Bennett D, Chertkow H, Cummings JK, de Leon M, Feldman H, Ganguli M, Hampel H, Scheltens P, Tierney M, Whitehouse P, Winblad B. Mild cognitive impairment. *Lancet*. 2006; 367:1262–1270. [http://dx.doi.org/10.1016/S0140-6736\(06\)68542-5](http://dx.doi.org/10.1016/S0140-6736(06)68542-5). [PubMed: 16631882]
- Goodglass, H.; Kaplan, E. *The assessment of aphasia and related disorders*. Philadelphia: Lea & Febiger; 1983.
- Grady CL, Haxby J, Horwitz B, Sundaram M, Berg G, Schapiro M, Friedland R, Rapoport S. Longitudinal study of the early neuropsychological and cerebral metabolic changes in dementia of the Alzheimer type. *Journal of Clinical and Experimental Neuropsychology*. 1988; 10:576–596. <http://dx.doi.org/10.1080/01688638808402796> PMID:3265710. [PubMed: 3265710]
- Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics*. 2005; 61:92–105. <http://dx.doi.org/10.1111/j.0006-341X.2005.030814.x> PMID:15737082. [PubMed: 15737082]
- Hsu Y-Y, Schuff N, Du A-T, Mark K, Zhu X, Hardin D, Weiner MW. Comparison of automated and manual MRI volumetry of hippocampus in normal aging and dementia. *Journal of Magnetic Resonance Imaging*. 2002; 16:305–310. <http://dx.doi.org/10.1002/jmri.10163> PMID:12205587 PMCID:1851676. [PubMed: 12205587]
- Jack CR Jr, Petersen RC, Xu YC, O'Brien PC, Smith GE, Ivnik RJ, Boeve BF, Waring SC, Tangalos EG, Kokmen E. Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. *Neurology*. 1999; 52:1397–1397. <http://dx.doi.org/10.1212/WNL.52.7.1397> PMID: 10227624 PMCID:2730146. [PubMed: 10227624]

- Jack CR Jr, Shiung MM, Weigand SD, OBrien PC, Gunter JL, Boeve BF, Knopman DS, Smith GE, Ivnik RJ, Tangalos EG, et al. Brain atrophy rates predict subsequent clinical conversion in normal elderly and amnestic. *MCI. Neurology*. 2005; 65:1227–1231. <http://dx.doi.org/10.1212/01.wnl.0000180958.22678.91> PMID:16247049 PMCID:2753547.
- Jack CR Jr, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, Borowski B, Britson PJJ, Whitwell J, Ward C, et al. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging*. 2008; 27:685–691. <http://dx.doi.org/10.1002/jmri.21049> PMID:18302232 PMCID:2544629. [PubMed: 18302232]
- Jack CR Jr, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, Petersen RC, Trojanowski JQ. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet neurology*. 2010; 9:119–128. [http://dx.doi.org/10.1016/S1474-4422\(09\)70299-6](http://dx.doi.org/10.1016/S1474-4422(09)70299-6). [PubMed: 20083042]
- Landau S, Harvey D, Madison C, Reiman E, Foster N, Aisen P, Petersen R, Shaw L, Trojanowski J, Jack CR Jr, et al. Comparing predictors of conversion and decline in mild cognitive impairment. *Neurology*. 2010; 75:230–238. <http://dx.doi.org/10.1212/WNL.0b013e3181e8e8b8> PMID:20592257 PMCID:2906178. [PubMed: 20592257]
- Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *Journal of The American Statistical Association*. 1958; 53:457–481. <http://dx.doi.org/10.1080/01621459.1958.10501452>.
- Kaplan, E.; Goodglass, H.; Weintraub, S. *Boston Naming Test*. Philadelphia: Lea & Febiger; 1983.
- Komori O, Eguchi S. A boosting method for maximizing the partial area under the ROC curve. *BMC Bioinformatics*. 2010; 11:314. <http://dx.doi.org/10.1186/1471-2105-11-314> PMID:20537139 PMCID:2898798. [PubMed: 20537139]
- Leung KK, Clarkson MJ, Bartlett JW, Clegg S, Jack CR Jr, Weiner MW, Fox NC, Ourselin S. Robust atrophy rate measurement in Alzheimer's disease using multi-site serial MRI: Tissue-specific intensity normalization and parameter selection. *Neuroimage*. 2010; 50:516–523. <http://dx.doi.org/10.1016/j.neuroimage.2009.12.059> PMID:20034579 PMCID:2828361. [PubMed: 20034579]
- Mattsson N, Zetterberg H, Hansson O, Andreason N, Parnetti L, Jonsson M, Herukka S-K, van der Flier WM, Blankenstein MA, Ewers M, et al. CSF biomarkers and incipient Alzheimer disease in patients with mild cognitive impairment. *JAMA: the journal of the American Medical Association*. 2009; 302:385–393. <http://dx.doi.org/10.1001/jama.2009.1064> PMID:19622817. [PubMed: 19622817]
- Neugroschl J, Sano M. An update on treatment and prevention strategies for Alzheimer's disease. *Current Neurology and Neuroscience Reports*. 2009; 9:368–376. <http://dx.doi.org/10.1007/s11910-009-0054-1> PMID:19664366. [PubMed: 19664366]
- Pepe MS, Kerr KF, Longton G, Wang Z. Testing for improvement in prediction model performance. *Stat. Med*. 2013 Epub ahead of print. <http://dx.doi.org/10.1002/sim.5727>.
- Petersen RC, Smith GE, Ivnik RJ, Tangalos EG, Schaid DJ, Thibodeau SN, Kokmen E, Waring SC, Kurland LT. Apolipoprotein E status as a predictor of the development of Alzheimer's disease in memory-impaired individuals. *JAMA: the journal of the American Medical Association*. 1995; 273:1274–1278. <http://dx.doi.org/10.1001/jama.1995.03520400044042> PMID:7646655. [PubMed: 7646655]
- Petersen RC, Negash S. Mild cognitive impairment: An overview. *CNS Spectrums*. 2008; 13:45–53. PMID:18204414. [PubMed: 18204414]
- Petersen RC, Roberts RO, Knopman DS, Boeve BF, Geda YE, Ivnik RJ, Smith GE, Jack CR Jr. Mild cognitive impairment: ten years later. *Archives of Neurology*. 2009; 66:1447–1455. <http://dx.doi.org/10.1001/archneurol.2009.266> PMID:20008648:PMCID:3081688. [PubMed: 20008648]
- Reitan RM. Validity of the trail making test as an indicator of organic brain damage. *Percept. Mot. Skills*. 1958; 8:271–276.
- Rey, A. *L'examen clinique en psychologie*. Paris: Presses Universitaires de France; 1964.
- Schuff N, Woerner N, Boreta L, Kornfield T, Shaw L, Trojanowski J, Thompson P, Jack CR Jr, Weiner M, et al. MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. *Brain*. 2009; 132:1067–1077. <http://dx.doi.org/10.1093/brain/awp007> PMID:19251758 PMCID:2668943. [PubMed: 19251758]

- Shaw LM, Vanderstichele H, KnapiKczajka M, Clark CM, Aisen PS, Petersen RC, Blennow K, Soares H, Simon A, Lewczuk P, et al. Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. *Annals of neurology*. 2009; 65:403–413. <http://dx.doi.org/10.1002/ana.21610> PMID:19296504 PMCID:2696350. [PubMed: 19296504]
- Smith EE, Egorova S, Blacker D, Killiany RJ, Muzikansky A, Dickerson BC, Tanzi RE, Albert MS, Greenberg SM, Guttman CR. Magnetic resonance imaging white matter hyperintensities and brain volume in the prediction of mild cognitive impairment and dementia. *Archives of neurology*. 2008; 65:94–100. <http://dx.doi.org/10.1001/archneurol.2007.23> PMID:18195145. [PubMed: 18195145]
- Storandt M, Hill RD. Very mild senile dementia of the Alzheimer type. II. Psychometric test performance *Arch. Neurol*. 1989; 46:383–386. <http://dx.doi.org/10.1001/archneur.1989.00520400037017> PMID:2705897.
- Tibshirani R. Regression shrinkage and selection via the lasso. *J. Royal. Statist. Soc B*. 1996; 58:267–288.
- Tibshirani R. The lasso method for variable selection in the Cox model. *Statistics in Medicine*. 1997; 16:385–395. [http://dx.doi.org/10.1002/\(SICI\)1097-0258\(19970228\)16:4<385::AID-SIM380>3.0.CO;2-3](http://dx.doi.org/10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3). [PubMed: 9044528]
- Wechsler D. Wechsler Adult Intelligence Scale - Revised edition. San Antonio: Psychological Corporation. 1981
- Wechsler D. Wechsler Memory Scale - Revised edition. San Antonio: Psychological Corporation. 1987 PMCID:255998.
- Weiner MW, Aisen PS, Jack CR Jr, Jagust WJ, Trojanowski JQ, Shaw L, Saykin AJ, Morris JC, Cairns N, Beckett LA, et al. The Alzheimer's disease neuroimaging initiative: progress report and future plans. *Alzheimer's & dementia: the journal of the Alzheimer's Association*. 2010; 6:202–211.
- Zhang D, Shen D. Alzheimer's Disease Neuroimaging Initiative. Predicting Future Clinical Changes of MCI Patients Using Longitudinal and Multimodal Biomarkers. *PLoS ONE*. 2012; 7(3):e33182. <http://dx.doi.org/10.1371/journal.pone.0033182>. [PubMed: 22457741]

Classic ROC analysis

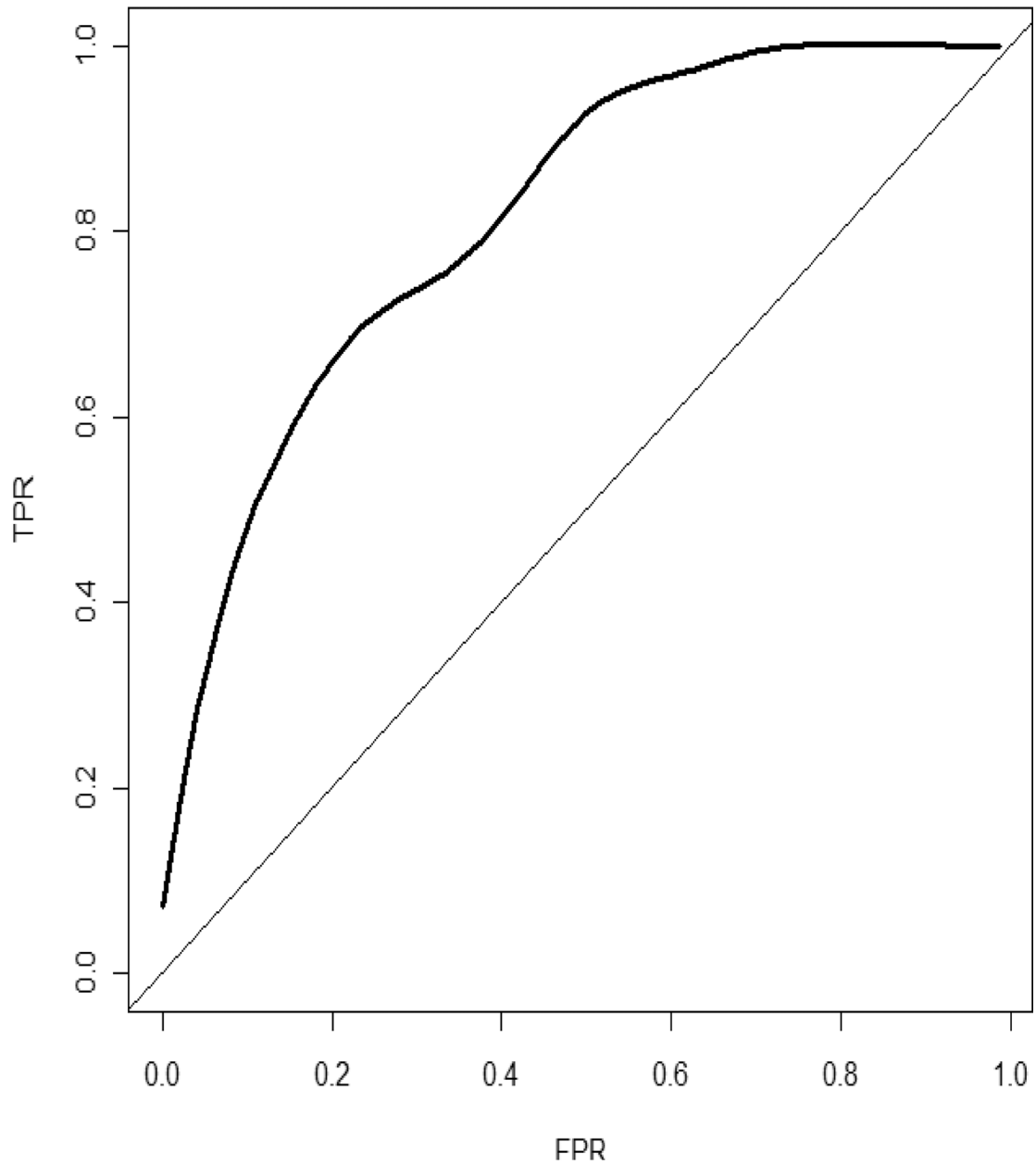
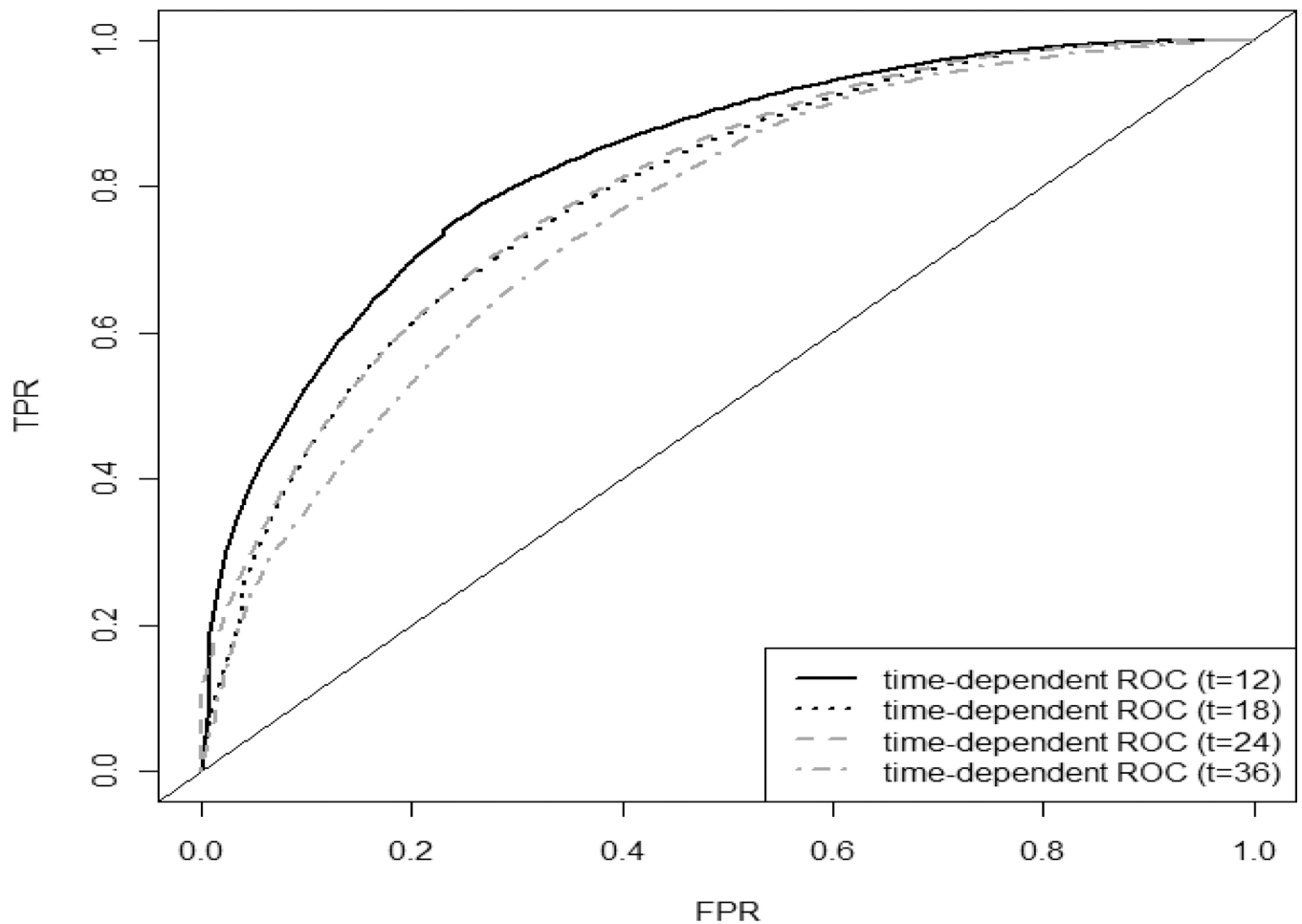


Fig 1. Classic ROC analysis. The model score is a combination of multiple markers: brain volume, hippocampal volume, ventricular volume, APOE ϵ 4 two alleles, $A\beta_{42}$, memory composite, executive function composite, age, education, and gender.

ROC curves for the model score at different time points**Fig 2.**

Comparison of ROC curves for the model score in Table 4 at different time points (i.e., 12 months, 18 months, 24 months and 36 months). The model score is a combination of multiple markers: brain volume, hippocampal volume, ventricular volume, APOE $\epsilon 4$ two alleles, $A\beta_{42}$, memory composite, executive function composite, age, education, and gender.

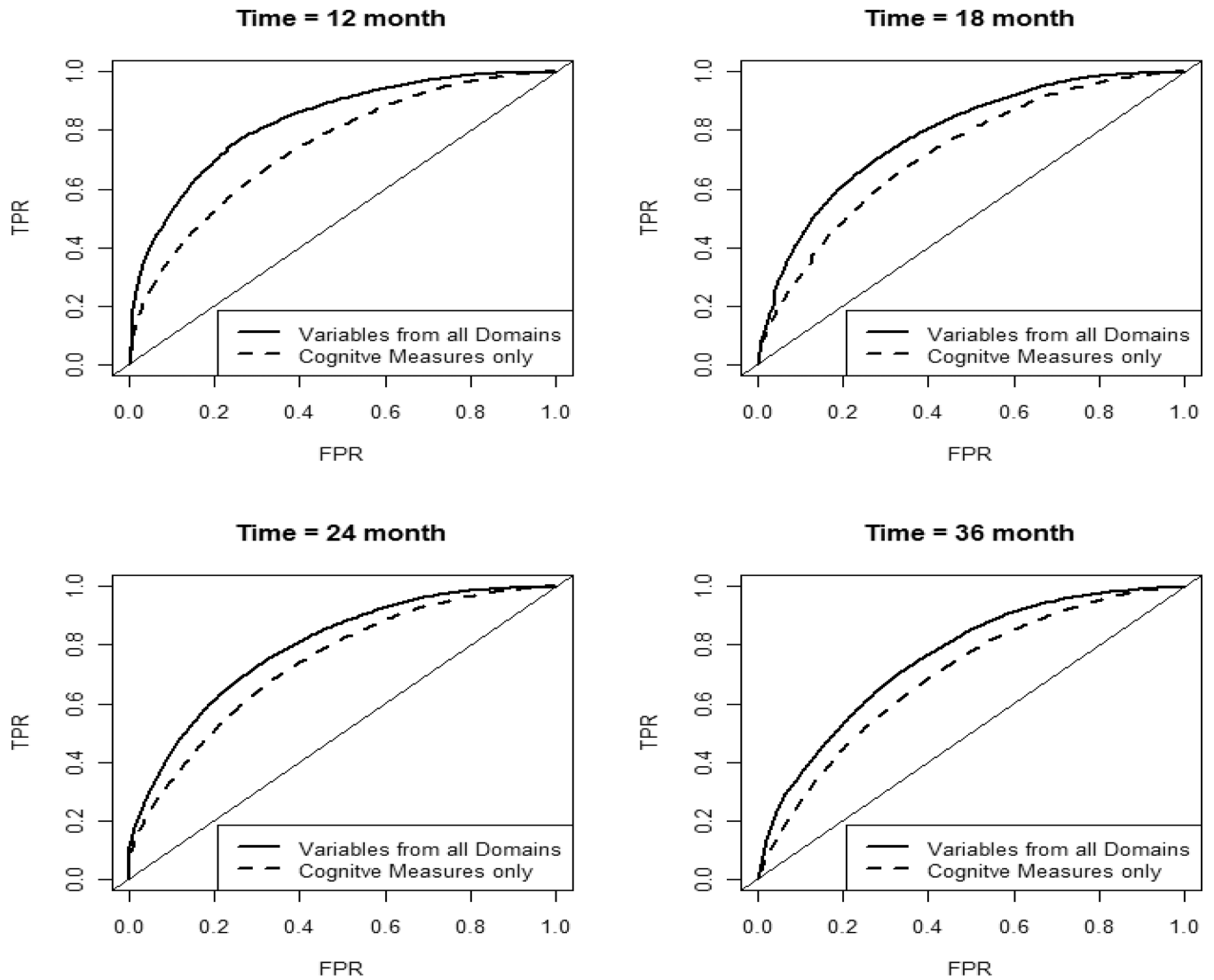


Fig 3.

Comparison of ROC curves for different models (model with variables from all domains as listed in Table 4 and model with cognitive measures only). The predictors from all domains include brain volume, hippocampal volume, ventricular volume, APOE $\epsilon 4$ two alleles, $A\beta_{42}$, memory composite, executive function composite. The predictors from cognitive measures include memory composite and executive function composite.

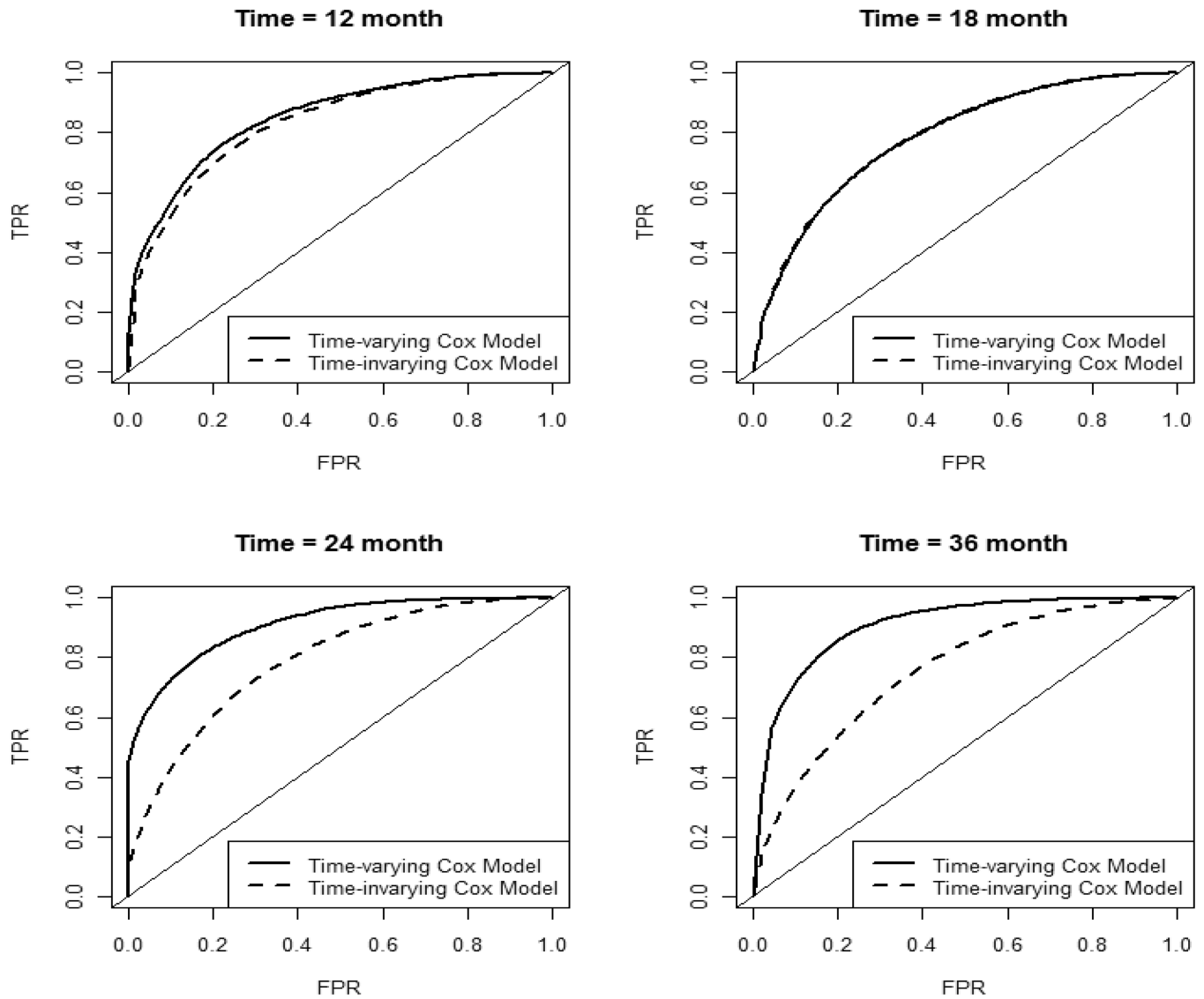


Fig 4.

Comparison of ROC curves for different models (the time varying Cox model and the time invariant Cox model). The predictors include brain volume, hippocampal volume, ventricular volume, APOE $\epsilon 4$ two alleles, $A\beta_{42}$, memory composite, executive function composite, age, education, and gender. The regression coefficients are different at different time points.

Table 1

Characteristics of Participants with complete covariates at Baseline in the ADNI study, United States and Canada, 2004–2009. Abbreviations: APOE ϵ 4 one allele = possession of one copy of apolipoprotein E ϵ 4 allele; APOE ϵ 4 two alleles = possession of two copies of apolipoprotein E ϵ 4 alleles; GDS = Geriatric Depression Scale; MMSE = Mini-Mental State Examination, CDR = Clinical Dementia Rating scale.

Variable	Value
Age, mean (SD)	74.63 (7.50)
Female, %	36.69
Caucasian, %	94.96
Education, mean (SD)	15.71 (3.04)
APOE ϵ 4 one allele, %	39.57
APOE ϵ 4 two allele, %	12.95
GDS, mean (SD)	1.55 (1.28)
MMSE, mean (SD)	26.89 (1.79)
CDR-global 0.5, %	100.00
CDR-sum of boxes, mean (SD)	1.58 (0.92)

Table 2

Summary Statistics for the baseline covariates of the analyzed subjects

Domain/Measure	Mean	SD	Min	Max
Neuroimaging				
Brain volume (cm ³)	1045.77	108.54	773.54	1332.73
Ventricular volume (cm ³)	45.84	22.71	9.44	133.52
Hippocampal volume (mm ³)	1812.06	345.59	959.95	2723.79
Cerebrospinal fluid				
T-tau (pg/mL)	104.65	55.69	28.16	350.78
A β 42 (pg/mL)	160.49	53.86	53.25	293.71
P-tau181 (pg/mL)	36.43	17.60	10.87	87.51
T-tau/A β 42 (pg/mL)	0.78	0.60	0.15	3.76
P-tau181/A β 42 (pg/mL)	0.27	0.18	0.04	0.89
Cognition^I				
Memory	-0.10	0.74	-1.45	2.19
Language	0.07	0.82	-2.70	1.74
Executive Function	0.01	0.72	-2.38	1.26
Spatial ability	-0.04	0.94	-3.87	0.53
Attention	0.05	1.05	-2.14	2.77

^IConverted to Z-scores (mean 0 and variance 1)

Table 3

Univariate Prediction of Time to Progression to AD

Domain/Measure	Estimate	SE	p-value
Neuroimaging			
Brain volume (cm ³)	-0.0052	0.0015	0.001
Ventricular volume (cm ³)	0.0154	0.0057	0.007
Hippocampal volume (mm ³)	-0.0018	0.0004	<0.001
Genetics ²			
APOE ε4 one allele	0.2965	0.2759	0.283
APOE ε4 two allele	0.7261	0.3509	0.039
Cerebrospinal fluid			
T-tau (pg/mL)	0.0033	0.0021	0.104
Aβ ₄₂ (pg/mL)	-0.0062	0.0023	0.008
P-tau ₁₈₁ (pg/mL)	0.0137	0.0066	0.039
T-tau/Aβ ₄₂ (pg/mL)	0.2311	0.1793	0.197
P-tau ₁₈₁ /Aβ ₄₂ (pg/mL)	1.1588	0.6412	0.071
Cognition ³			
Memory	-1.2270	0.2339	<0.001
Language	-0.2884	0.1556	0.064
Executive Function	-0.5195	0.1551	<0.001
Spatial ability	-0.1724	0.1204	0.152
Attention	-0.0522	0.1155	0.652

² Categorical measure (0 = APOE ε4 negative, 1 = APOE ε4 one allele, 2 = APOE ε4 two alleles)

³ Converted to Z-scores (mean 0 and variance 1)

Table 4

Multivariable Prediction of Time to Progression to AD with variables selected from LASSO

Measure	Estimate	SE	<i>p</i> -value
Brain volume (cm ³)	-0.0046	0.0018	0.009
Ventricular volume (cm ³)	0.0199	0.0065	0.002
Hippocampal volume (mm ³)	-0.0010	0.0005	0.059
APOE ε4 two allele	0.8739	0.3682	0.018
Aβ ₄₂	-0.0039	0.0030	0.190
Memory (z-score)	-1.1551	0.2352	< 0.001
Executive function (z-score)	-0.5435	0.1891	0.004

Table 5

Multivariable Prediction of Time to Progression to AD using Reduced Sample at 12, 18, 24 and 36 Month

Measure	Estimate	SE	p-value
T = 12 Month			
Brain volume (cm ³)	-0.0044	0.0019	0.023*
Ventricular volume (cm ³)	0.0217	0.0077	0.005*
Hippocampal volume (mm ³)	-0.0008	0.0006	0.164
APOE ε4 two alleles	1.1824	0.3996	0.003*
Aβ ₄₂	-0.0033	0.0031	0.285
Memory (z-score)	-1.2648	0.2607	< 0.001*
Executive function (z-score)	-0.5560	0.2215	0.012*
T = 18 Month			
Brain volume (cm ³)	-0.0025	0.0022	0.262
Ventricular volume (cm ³)	0.0138	0.0098	0.157
Hippocampal volume (mm ³)	-0.0011	0.0007	0.111
APOE ε4 two alleles	0.9421	0.5130	0.066
Aβ ₄₂	-0.0025	0.0037	0.504
Memory (z-score)	-1.1159	0.3182	< 0.001*
Executive function (z-score)	-0.7318	0.2690	0.006*
T = 24 Month			
Brain volume (cm ³)	-0.0019	0.0026	0.460
Ventricular volume (cm ³)	0.0237	0.0134	0.076
Hippocampal volume (mm ³)	-0.0012	0.0010	0.244
APOE ε4 two alleles	0.6097	0.6721	0.364
Aβ ₄₂	-0.0073	0.0055	0.178
Memory (z-score)	-1.8449	0.5009	< 0.001*
Executive function (z-score)	-1.2259	0.3774	0.001*
T = 36 Month			
Brain volume (cm ³)	-0.0037	0.0057	0.513
Ventricular volume (cm ³)	0.0352	0.0021	0.102
Hippocampal volume (mm ³)	-0.0028	0.0022	0.206
APOE ε4 two alleles	1.9182	1.3249	0.148
Aβ ₄₂	-0.0061	0.0077	0.430
Memory (z-score)	-1.9299	0.8563	0.024*
Executive function (z-score)	-1.4686	0.8344	0.078

* The indicates variables with p-value <0.05