Published in final edited form as:

Alzheimers Dement. 2018 May ; 14(5): 644–651. doi:10.1016/j.jalz.2017.11.004.

A prognostic model of Alzheimer's disease relying on multiple longitudinal measures and time-to-event data

Kan Li^a, Richard O'Brien^b, Michael Lutz^b, Sheng Luo^{a,*}, and the Alzheimer's Disease Neuroimaging Initiative¹

^aDepartment of Biostatistics, The University of Texas Health Science Center at Houston, Houston, TX, USA

^bDepartment of Neurology, Duke University School of Medicine, Durham, NC, USA

Abstract

INTRODUCTION—Characterizing progression in Alzheimer's disease (AD) is critically important for early detection and targeted treatment. The objective was to develop a prognostic model, based on multivariate longitudinal makers, for predicting progression-free survival in patients with mild cognitive impairment.

METHODS—The information contained in multiple longitudinal makers was extracted using multivariate functional principle components analysis and used as predictors in the Cox regression models. Cross-validation was used for selecting the best model based on ADNI-1. External validation was conducted on ADNI-2.

RESULTS—Model comparison yielded a prognostic index computed as the weighted combination of historical information of five neurocognitive longitudinal markers that are routinely collected in observational studies. The comprehensive validity analysis provided solid evidence of the usefulness of the model for predicting AD progression.

DISCUSSION—The prognostic model was improved by incorporating multiple longitudinal makers. It is useful for monitoring disease and identifying patients for clinical trial recruitment.

Keywords

mild cognitive impairment; multivariate functional component analysis; prediction; external validation; ADNI

*Corresponding author: Sheng Luo, PhD, Department of Biostatistics, School of Public Health, The University of Texas Health Science Center at Houston, 1200 Herman Pressler Dr, Rm E815, Houston, TX 77030, USA. Tel: +1-713-500-9554; sheng.t.luo@uth.tmc.edu.

¹Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu>). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Conflicts of interest: none

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

The growing public health threat posed by Alzheimer's disease (AD) has raised the urgency to discover and assess prognostic markers for the early detection of the disease. In this regard, a great deal of effort has been dedicated to building models for predicting AD based on a single marker, or a combination of multiple markers, which captures the heterogeneity among subjects and detects the disease progression of subjects at risk [1].

Since mild cognitive impairment (MCI) is a risk state for AD, existing research has identified a number of biomarkers that predict clinical changes of MCI patients [2–7], including neurocognitive markers, neuroimaging markers, genetics, and cerebrospinal fluid (CSF)-based markers. Although these studies showed an enhancement of the prognostic value when multivariate markers are considered, the prediction was either qualitative change (i.e., conversion from MCI to AD as a binary response) or quantitative change (i.e., cognitive scores) in the next few years. These studies did not examine the prediction of time from MCI to AD, which is usually the primary outcome in prevention clinical studies of AD [8].

Relatively few recent studies have investigated the timing from MCI to AD over the duration of follow-up based on Cox regression models [9–11]. These studies assessed the predictive utility of various candidate prognostic markers independently or in combination, but only focused on the baseline measurements of the markers. Li, et al. [12] developed a joint modeling of longitudinal and time-to-event data technique to examine the comparative utility of the longitudinal markers in determining the risk of incident AD conversion at future time points. However, they analyzed each longitudinal marker independently because of the limitation of the current state-of-the-art joint modeling software. To our knowledge, no prior study has leveraged multiple longitudinal markers and time-to-event information jointly to investigate the prognosis of Alzheimer's disease.

The overarching goal of this study is to develop a prognostic model, which relies on serial measurements of multiple markers, for predicting progression-free survival in patients with MCI. The prognostic model uses several cutting-edge statistical methods, which enable it to facilitate clinical decision-making based on all the information collected. We evaluated the model using data from the Alzheimer's Disease Neuroimaging Initiative 1 (ADNI-1), a public dataset that is well suitable for this task because of its large sample size, breadth of markers, and prospective structure. The combined prognostic value of longitudinal neurocognitive tests, neuroimaging, genetics, and CSF markers was assessed using the prognostic model. External validation of the model is carried out on ADNI-2 to demonstrate the usefulness of the model across studies. The main output of the prognostic model is a prognostic index which can be updated over time as new measurements are available. Such an index is useful for monitoring disease progression for MCI patients and to enrich clinical trials with subjects likely to develop AD in the timeframe of the trial.

2. Materials and Methods

2.1 Study population

Alzheimer's Disease Neuroimaging Initiative (ADNI) is a longitudinal observational study investigating whether serial brain imaging, clinical and neuropsychological assessments can be combined to measure the progression of Alzheimer's disease. Detailed information regarding the ADNI study procedures, including participant inclusion and exclusion criteria and the complete study protocol can be found at <http://www.adni-info.org>. The ADNI-1 dataset included 379 patients with amnesic MCI at baseline evaluation who had at least one follow-up visit. Criteria for MCI diagnosis were the same as defined by Petersen et al. [13]. As part of the ADNI-1, subjects were assessed at baseline, 6, 12, 18, 24 and 36 months, and continued follow-ups were conducted annually as part of the ADNI-2. The ADNI-2 study had the same overall goals as ADNI-1. ADNI-2 enrolled an additional group of 424 patients in early MCI, late MCI, and significant memory concern, with at least one follow-up visit and collected over 4-year worth of longitudinal data. We considered all these subjects as MCI patients in our analysis. All subjects were given a written informed consent at the time of enrollment, and the study has been approved by the local institutional review board at all participating sites. The data are publically available at <http://ida.loni.ucla.edu> and were downloaded on April 15, 2017.

In the present study, the candidate prognostic factors were the longitudinal neurocognitive and imaging markers with the strongest predictive utility identified by Li, et al. [12], and are available in both ADNI-1 and ADNI-2 datasets: Alzheimer Disease Assessment Scale–Cognitive 13 items (ADAS-Cog 13); Rey Auditory Verbal Learning Test (RAVLT immediate: sum of 5 trials; RAVLT learning: trial 5 – trial 1); Functional Assessment Questionnaire (FAQ); Mini Mental State Examination (MMSE); volumetric data of middle temporal gyrus (MidTemp) and hippocampus from structural magnetic resonance imaging; and fluorine 18 fluorodeoxyglucose (FDG) positron emission tomography (PET). We also considered cerebrospinal fluid (CSF) biomarkers, total tau (t-tau) and amyloid-beta 1–42 peptide ($A\beta_{42}$), which have been identified as being amongst the most promising and informative AD biomarkers [14,15]. Relevant demographic and genetic variables (i.e., age, gender, education and *APOE* genotype) were employed as other prognostic variables given their potential effects on disease progression in AD [16]. Sample size and descriptive statistics for key variables at baseline are shown in Table 1.

2.2 Statistical analysis

We adopted a novel method to jointly model the historical observations of markers as well as time-to-event of interest (e.g. conversion from MCI diagnosis to AD diagnosis) using a functional principal component (FPC) approach and Cox regression model [17,18]. The observed longitudinal values of a specific marker were assumed to come from a latent longitudinal process. The FPC analysis (FPCA) was employed to extract the trajectory pattern of the process [19]. The overall trend and changing patterns of the marker were estimated from the entire sample. Each marker trajectory of an individual was summarized as a set of FPC scores. This approach has the advantage of handling missed, irregularly observed data, measurement errors, and no pre-specified form of the longitudinal trajectory

[19]. Because the longitudinal markers of interest were all associated with AD and are likely to be highly correlated, a non-negligible correlation may exist between the FPC scores that derived individually from each longitudinal marker. The correlated scores can lead to interpretation difficulty and multicollinearity issues in the regression analysis. To address this issue, we adopted a multivariate FPC (MFPC) approach [20], which extends the FPC approach by accounting for the correlation among multiple longitudinal markers. We used the MFPC scores, which capture the essential information in the multiple longitudinal markers, as new prognostic factors for predicting risk of event in a survival model. The statistical approach is detailed in the Appendix A.

We first used the ADNI-1 dataset to investigate whether multivariate longitudinal markers, compared to baseline values of the markers, could enhance the predictive performance of AD conversion in terms of discriminating patients with high and low risk. We did not thoroughly explore every possible combination of the markers, but mainly focused on the five that have strong comparative predictive value and are commonly measured in studies of Alzheimer's disease, i.e., ADAS-Cog 13, RAVLT immediate, RAVLT learning, FAQ, and MMSE. Two preplanned Cox regression models were specified, with the first model (Model 1) including only baseline observation of the five markers, and the second model (Model 2) incorporated the longitudinal information of the five markers by including the multivariate FPC scores derived from the longitudinal markers of interest. We selected the first 12 MFPC scores which account for at least 95% total variation in the five longitudinal markers. We also included baseline age, gender, education years and number of the *APOE* $\epsilon 4$ alleles as covariates in both models. Since the exact date of conversion from MCI to AD was not known, the survival endpoint was defined as the date of the first follow-up with an AD diagnosis. Subjects that did not convert were censored at the date of their last interview. The two models were specified as

$$h_i(t) = h_0(t) \exp\{\gamma_1 \text{baselineAge}_i + \gamma_2 \text{gender}_i + \gamma_3 \text{education}_i + \gamma_4 \text{APOE}\epsilon 4_i + \gamma_5 \text{ADAS} - \text{Cog13}_i + \gamma_6 \text{RAVLT} . \text{imme}_i + \gamma_7 \text{RAVLT} . \text{learn}_i + \gamma_8 \text{FAQ}_i + \gamma_9 \text{MMSE}_i\}; \quad (1)$$

$$h_i(t) = h_0(t) \exp\{\gamma_1 \text{baselineAge}_i + \gamma_2 \text{gender}_i + \gamma_3 \text{education}_i + \gamma_4 \text{APOE}\epsilon 4_i + \sum_{p=1}^{12} \beta_p \text{MFPCscores}_{ip}\}, \quad (2)$$

where $h_0(t)$ was the baseline hazard function, γ_i and β_p were regression coefficients. There was no error term in Cox model as the randomness was implicit to the survival process. The MFPCscores_i were calculated from the subject i 's longitudinal measures that collected prior to the visit of AD diagnosis or censoring. It is worth mentioning that MFPCscores_i can be calculated for subjects with only baseline observation, but repeated measures may provide more precise information for risk prediction. For a comprehensive comparison, we also investigated the predictive ability of baseline imaging markers by including hippocampal

volume, MidTemporal volume, and FDG-PET in addition to the five baseline neurocognitive markers in Model 1. We refer this model as Model 1a.

We assessed the performance of each prognostic model using an integrated area under the time-dependent receiver operating characteristic curve (iAUC; range 0 ~ 1, higher is better) [21]. We consider iAUC = 0.78 to be an acceptable predictive performance as suggested in the previous study [23]. We also used integrated Brier score [22] (lower is better) as a calibration tool to assess the agreement between the predicted and true risks. To avoid overestimation of the predictive performance, we conducted a 10-fold cross validation (CV) for survival model [24]. The 10-fold CV process was repeated 100 times to account for the variation due to data splitting. To provide a benchmark for comparison, internal validation for ADNI-1 and ADNI-2 was performed using repeated 10-fold CV. For the external validation, parameter estimations from ADNI-1 were used to calculate the performance measures for ADNI-2. To conduct comparable MFPC analyses, the longitudinal trajectories in ADNI-1 and ADNI-2 datasets were defined on the same time domain with the largest follow-up time in ADNI-1 as the upper bound.

In the second analysis, the structural form of Model 2 was used to evaluate changes to the performance of the model as additional longitudinal markers were added. In addition to the five neurocognitive-based longitudinal markers used in the first analysis, we included three longitudinal imaging markers: hippocampal volume, MidTemporal volume, FDG-PET, to calculate MFPC scores. Because only a subset of subjects had all the markers, we conducted model comparisons by restricting subjects to those with all eight markers. We also investigated the additional predictive value of CSF biomarkers based on subjects whose CSF data were collected in addition to imaging markers. We only considered the baseline CSF biomarkers in the analysis, because only a small number of subjects in ADNI-1 and ADNI-2 had at least two CSF observations.

Based on the estimated Cox model (2), prognostic index (PI_{ADNI}) for each patient was computed as the weighted combination of the regression coefficients and the prognostic variables (Equation 3).

$$PI_{ADNI} = \gamma_1 baselineAge_i + \gamma_2 gender_i + \gamma_3 education_i + \gamma_4 APOE\epsilon 4_i \quad (3) \\ + \sum_{p=1}^{12} \beta_p MFPCscore_{ip}$$

In the external validation, the parameters estimated from ADNI-1 were used to compute the PI_{ADNI} in ADNI-2. Risk groups for conversion from MCI to AD were formed by the quartiles of the PI_{ADNI} that computed in ADNI-1. In addition, we demonstrated how the PI_{ADNI} for a hypothetical patient was updated over time as new clinical information became available. All the analyses were performed using R (version 3.2.1).

3. Results

The results of the first analysis were summarized in Table 2. The multivariate baseline markers (Model 1) showed that the internal predictive performance ($iAUC_{INT}$) on ADNI-1 dataset was 0.776, which was closed to the pre-specified threshold of 0.78. Model 1a had about 4 percentage point increases in mean $iAUC_{INT}$ compared to Model 1, by including baseline volumetric MRI (hippocampus, MidTemporal), and FDG-PET in the model. Model 2 had about 7 percentage point increases in mean $iAUC_{INT}$ to 0.846, compared to Model 1, by utilizing all information from the multivariate longitudinal markers. For ADNI-2, the external predictive performance ($iAUC_{EXT}$) was as good as internal performance in both models. For example, ADNI-2 had internal performance of $iAUC_{INT}=0.909$ and external performance of $iAUC_{EXT}=0.933$ for Model 2. The performance statistics favored Model 2 with largest mean $iAUC$ s and smallest Brier score in both internal and external validations, which indicated that the longitudinal neurocognitive markers had the best predictive value. Therefore, we considered Model 2 to be a more preferable model.

Results for the second analysis which added longitudinal imaging markers to the collection of neurocognitive markers are shown in Table 3. All four models had the same functional forms as Model 2, except that the 12 MFPC scores were derived from different sets of multivariate longitudinal markers. Specifically, the MFPC scores in Model 2a incorporated the longitudinal information of the five neurocognitive markers in the first analysis. From Model 2b to Model 2d, longitudinal hippocampus, MidTemporal, and FDG-PET were incorporated successively and in combination into MFPC score calculation. The models had similar predictive performance (Model 2b through Model 2d), with slightly higher mean $iAUC_{INT}$ as the additional imaging markers were added to the model. This suggests that predictive performance can be improved by including additional longitudinal imaging markers, but the improvement was modest when the neurocognitive measures and genetic variable were already in the model. In general, ADNI-2 had higher internal and external performance compared with ADNI-1. This may be explained by a high variability in progression status in ADNI-2 population, i.e., the larger differences in multiple clinical outcomes between individuals progressed to AD and those who did not make it easier to classify the two groups. Moreover, the additional predictive value of CSF markers was small when neurocognitive and imaging markers had already in the model. Based on a small subset of ADNI-2 ($n = 177$) and 10-fold CV, $iAUC_{INT}$ increased one percentage point by adding baseline total tau (t-tau) and $A\beta_{42}$ as predictors into Model 2d. We did not further consider the model with imaging and CSF markers because of the small improvement in predictive performance. Model 2a was potentially more suitable for clinical settings because the cognitive and functional markers can be collected at minimal cost and without specific equipment.

We selected Model 2a as the prognostic model for further analysis and determined the prognostic index PI_{ADNI} for both datasets based on the parameters derived from ADNI-1. The cohort in each dataset was categorized into four risk groups based on the quartiles of the PI_{ADNI} computed from the ADNI-1 data. Figure 1 shows the Kaplan-Meier survival curves for the AD diagnosis risk groups. Survival curves for ADNI-1 were similar to the survival curves for ADNI-2 in the sense that the risk curves were in the same order (from top to

bottom: low, mid-low, mid-high, and high). This indicates that the PI_{ADNI} quartiles which categorize four groups for ADNI-1 can be well applied to ADNI-2. The survival curves in the validation ADNI-2 dataset generally agree well with those calculated from the ADNI-1 dataset. For example, in both studies, 50% in the high risk cohort are expected to have an AD diagnosis around 1 year, and 80% are expected to have an AD diagnosis by 2 years. Although the follow up periods of ADNI-1 and ADNI-2 were different, these two cohorts were reasonably similar, and thus the two datasets were combined, and used to estimate final parameters in MFPCA and Cox regression.

Based on the estimated MFPC scores for each subject and the Cox regression coefficients (Table B.1), the PI_{ADNI} can be computed. Large values indicate greater risk of conversion from MCI to AD and the PI_{ADNI} quartiles (based on pooled data; 25%: -0.78 ; 50%: 0.29 ; 75%: 1.56) can help to estimate the risk for an individual and enhance interpretation based on survival curves. In addition, when new marker information becomes available for an individual, his/her MFPC scores, as well as PI_{ADNI} , can be updated based on the longitudinal trajectories with new data. In Appendix B, we selected one target MCI patient from ADNI-1 population to demonstrate how PI_{ADNI} was updated over time. In short, for this patient, PI_{ADNI} based on the baseline observations was -0.65 , which categorized the patient into mid-low risk group. By incorporating the additional measurements that were collected at three follow-up visits, the value of PI_{ADNI} increased to 0.34 (mid-high risk group), suggesting that the patient was at increased risk of conversion from MCI to AD and further intervention could be initiated as necessary. Based on the Kaplan-Meier curves in Figure 1, we can get an estimation of patients' survival based on the risk groups they are in. For any future time point, we can also approximate the baseline risk function $h_0(t)$ in Model 2 with splines or parametric functions and calculate the survival risk of AD diagnosis for this patient.

4. Discussion

In this paper, we assessed the performance of multivariate longitudinal markers for the prediction of the timing of AD diagnosis (conversion) for MCI patients. We proposed a novel statistical method to jointly analyze time-to-AD from MCI as well as the history of multiple longitudinal markers of cognitive, functional/behavioral, neuroimaging and CSF. A prognostic model for AD diagnosis based on multiple longitudinal measurements was developed. Such prognostic models with multivariate longitudinal markers have not been developed before.

One strength of this study was that the prognostic value of the model was evaluated via both internal and external validations. External validation is the gold standard for reproducibility [25], and the good predictive performance in ADNI-2 external study underscores the general usefulness of the proposed prognostic model. Our results suggest that including the history of multivariate longitudinal markers, in addition to baseline variables, can greatly enhance prognostic performance. The prognostic index generated from the prognostic model has classified individuals into four risk groups of AD diagnosis by its quantiles. Both ADNI-1 and ADNI-2 studies indicated that 50% of subjects in the cohort of high risk remained prediagnosis at year one, and 50% of subjects in the cohort of mid-high risk group remained

prediagnosis at year three. The emphasis of the proposed PI_{ADNI} was that it efficiently utilized the longitudinal information of multiple markers manifested by the MFPC scores, and can be updated dynamically to reflect the patient's latest prognosis.

A model based on cognitive and functional assessments as prognostic factors can be widely used in clinical setting because of their broad availability. Our analyses showed that the improvement in predictive power by adding imaging biomarkers and baseline CSF markers is limited when cognitive and functional variables were already in the model. The proposed prognostic index PI_{ADNI} has the following advantages: (1) it is easy to compute with commonly measured variables; (2) it is useful to estimate when an individual will convert from MCI to AD; (3) it can identify candidate individuals for clinical trials from those who have registered in ongoing observational studies and may already have a history of biomarker observations. For example, if a preventative trial requires participants to be many years from AD diagnosis, the bottom percentage (e.g. 25%) of individuals with the smallest PI_{ADNI} (based on the longitudinal measures until the time point) might be selected. Alternatively, to increase the probability of detecting significant differences between treatment and control patients, a trial might recruit from the cohort with sufficient progression in the time frame (e.g. 3 years) of a trial. In this case, the untreated mid-high risk group may be more appropriate, as indicated by the corresponding survival curve in Figure 1. Individuals from whose PI_{ADNI} fall within this risk group might be candidates for recruitment.

We applied multivariate functional principle analysis (MFPCA) to multiple longitudinal marker data. The MFPCA can effectively deal with measurement errors and capture the temporal correlation in the markers as well as correlations among markers. It does not require a pre-defined functional form for the longitudinal trajectory and allows missing observations and different time visits for subjects. The computed MFPC scores may lack clinical interpretability, and not all scores are significantly associated with the likelihood of AD diagnosis (Table B.1). We investigated the variable selection for the MFPC scores. However, a more parsimonious model, which excluded the insignificant MFPC scores, gave no improvement in predictive performance for PI_{ADNI} . The approaches used in this study can be broadly applied to longitudinal studies of neurodegenerative diseases. A calculator for computing MFPC scores and prognostic index is available upon request from the authors. Although the methods can be implemented in available software, a well-documented and further validated web-based calculator is needed to facilitate their use in clinical settings.

There are several limitations in our study. First, although MFPC score can be calculated even if some subjects had only one observation, we required majority of subjects in the model building dataset to have at least two observations to ensure the accuracy of the parameter estimation in MFPCA. Second, for simplicity and without loss of generality, we assume the survival time is continuous. However, the AD conversion in ADNI study were diagnosed about every 6 month at follow-up visits. Further improvement of the model could be made by considering survival time as observed in grouped form or considering AD conversion as an interval censoring problem. Third, death is treated as a non-informative right-censoring event in our analyses. Cox proportional hazard model may produce biased parameter

estimates of the predictor effects when failing to account for the competing risk of death in elder population. The competing risk approaches could be adopted to improve our method depending on the study of interest. Fourth, we did not thoroughly explore the prognostic performance of all combinations of markers. A potentially more powerful prognostic index could be developed using a combination of other imaging, clinical, genetic or fluid biomarker variables. In addition, participants who volunteered for studies such as ADNI may not represent the more general clinical population and the potential selection bias exists. We used ADNI-2 dataset for external validation because of its breadth of cognitive, imaging and CSF biomarkers, and similarity to ADNI-1. If only commonly collected neurocognitive and demographic variables are of interest, external validation on other studies with large sample sizes, such as the National Alzheimer's Coordinating Center dataset, would provide additional validation and replication for our findings. Moreover, although the proposed method can easily incorporate multiple longitudinal markers and are computational efficient as compared with the joint modeling approach adopted in a previous paper [12], it may not well address the informative censoring issue for the longitudinal marker. For example, subjects with more serious disease status may be more likely to experience the event or withdraw from the study earlier than healthier individual, leading to fewer longitudinal measurements. To investigate the impact of the informative censoring, we conducted additional simulation studies to compare the proposed approach with joint models and found that the informative censoring had very limited effects on risk prediction. However, it is important to consider the issue in future methodology development.

Finally, given the lack of disease-modifying treatments for AD and recent failures of several promising drugs in clinical trials, it is believed that treatments may need to be initiated early in the disease process to be effective. Our finding demonstrated how a single risk score based on multivariate longitudinal variables is useful and valuable for early diagnosis, and for targeted recruitment for clinical trials.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank two anonymous reviewers for the thoughtful suggestions that help us improve the quality of our manuscript. This work was supported by the National Institute of Neurological Disorders and Stroke [grant number R01NS091307, 5U01NS043127]. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Funding: This work was supported by the National Institute of Neurological Disorders and Stroke [grant number R01NS091307, 5U01NS043127].

References

1. Weiner MW, Veitch DP, Aisen PS, Beckett LA, Cairns NJ, Green RC, et al. The Alzheimer's Disease Neuroimaging Initiative: A review of papers published since its inception. *Alzheimers Dement J Alzheimers Assoc.* 2012; 8:S1–68. DOI: 10.1016/j.jalz.2011.09.172
2. Devanand DP, Liu X, Tabert MH, Pradhaban G, Cuasay K, Bell K, et al. Combining early markers strongly predicts conversion from mild cognitive impairment to Alzheimer's disease. *Biol Psychiatry.* 2008; 64:871–9. DOI: 10.1016/j.biopsych.2008.06.020 [PubMed: 18723162]
3. Gomar JJ. Utility of Combinations of Biomarkers, Cognitive Markers, and Risk Factors to Predict Conversion From Mild Cognitive Impairment to Alzheimer Disease in Patients in the Alzheimer's Disease Neuroimaging Initiative. *Arch Gen Psychiatry.* 2011; 68:961.doi: 10.1001/archgenpsychiatry.2011.96 [PubMed: 21893661]
4. Cui Y, Liu B, Luo S, Zhen X, Fan M, Liu T, et al. Identification of Conversion from Mild Cognitive Impairment to Alzheimer's Disease Using Multivariate Predictors. *PLOS ONE.* 2011; 6:e21896.doi: 10.1371/journal.pone.0021896 [PubMed: 21814561]
5. Lo RY, Hubbard AE, Shaw LM, Trojanowski JQ, Petersen RC, Aisen PS, et al. Longitudinal change of biomarkers in cognitive decline. *Arch Neurol.* 2011; 68:1257–66. DOI: 10.1001/archneurol.2011.123 [PubMed: 21670386]
6. Zhang D, Shen D, Initiative ADN. Predicting Future Clinical Changes of MCI Patients Using Longitudinal and Multimodal Biomarkers. *PLOS ONE.* 2012; 7:e33182.doi: 10.1371/journal.pone.0033182 [PubMed: 22457741]
7. Ewers M, Walsh C, Trojanowski JQ, Shaw LM, Petersen RC, Jack CR, et al. Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance. *Neurobiol Aging.* 2012; 33:1203–1214e2. DOI: 10.1016/j.neurobiolaging.2010.10.019 [PubMed: 21159408]
8. Gauthier S, Wu L, Rosa-Neto P, Jia J. Prevention strategies for Alzheimer's disease. *Transl Neurodegener.* 2012; 1:13.doi: 10.1186/2047-9158-1-13 [PubMed: 23210473]
9. Li S, Okonkwo O, Albert M, Wang M-C. Variation in Variables that Predict Progression from MCI to AD Dementia over Duration of Follow-up. *Am J Alzheimers Dis Columbia Mo.* 2013; 2:12–28. DOI: 10.7726/ajad.2013.1002
10. Barnes DE, Cenzer IS, Yaffe K, Ritchie CS, Lee SJ. Alzheimer's Disease Neuroimaging Initiative. A point-based tool to predict conversion from mild cognitive impairment to probable Alzheimer's disease. *Alzheimers Dement J Alzheimers Assoc.* 2014; 10:646–55. DOI: 10.1016/j.jalz.2013.12.014
11. Kong D, Giovanello KS, Wang Y, Lin W, Lee E, Fan Y, et al. Predicting Alzheimer's Disease Using Combined Imaging-Whole Genome SNP Data. *J Alzheimers Dis JAD.* 2015; 46:695–702. DOI: 10.3233/JAD-150164 [PubMed: 25869783]
12. Li K, Chan W, Doody RS, Quinn J, Luo S. Prediction of Conversion to Alzheimer's Disease with Longitudinal Measures and Time-To-Event Data. *J Alzheimers Dis.* 2017; :1–11. Preprint. DOI: 10.3233/JAD-161201
13. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: Clinical characterization and outcome. *Arch Neurol.* 1999; 56:303–8. DOI: 10.1001/archneur.56.3.303 [PubMed: 10190820]
14. Shaw LM, Vanderstichele H, Knapik-Czajka M, Clark CM, Aisen PS, Petersen RC, et al. Cerebrospinal Fluid Biomarker Signature in Alzheimer's Disease Neuroimaging Initiative Subjects. *Ann Neurol.* 2009; 65:403–13. DOI: 10.1002/ana.21610 [PubMed: 19296504]
15. Hansson O, Zetterberg H, Buchhave P, Londos E, Blennow K, Minthon L. Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impairment: a follow-up study. *Lancet Neurol.* 2006; 5:228–34. DOI: 10.1016/S1474-4422(06)70355-6 [PubMed: 16488378]

16. Fleisher AS, Sowell BB, Taylor C, Gamst AC, Petersen RC, Thal LJ, et al. Clinical predictors of progression to Alzheimer disease in amnesic mild cognitive impairment. *Neurology*. 2007; 68:1588–95. DOI: 10.1212/01.wnl.0000258542.58725.4c [PubMed: 17287448]
17. Holte SE, Randolph TW, Ding J, Tien J, McClelland RS, Baeten JM, et al. Efficient use of longitudinal CD4 counts and viral load measures in survival analysis. *Stat Med*. 2012; 31:2086–97. DOI: 10.1002/sim.5318 [PubMed: 22415871]
18. Fang H-B, Wu TT, Rapoport AP, Tan M. Survival analysis with functional covariates for partial follow-up studies. *Stat Methods Med Res*. 2016; 25:2405–19. DOI: 10.1177/0962280214523586 [PubMed: 24567438]
19. Yao F, Müller H-G, Wang J-L. Functional Data Analysis for Sparse Longitudinal Data. *J Am Stat Assoc*. 2005; 100:577–90. DOI: 10.1198/016214504000001745
20. Happ C, Greven S. Multivariate Functional Principal Component Analysis for Data Observed on Different (Dimensional) Domains. *J Am Stat Assoc*. 2017; 0:0–0. DOI: 10.1080/01621459.2016.1273115
21. Uno H, Cai T, Tian L, Wei LJ. Evaluating Prediction Rules for t-Year Survivors with Censored Regression Models. *J Am Stat Assoc*. 2007; 102:527–37.
22. Gerds TA, Schumacher M. Consistent Estimation of the Expected Brier Score in General Survival Models with Right-Censored Event Times. *Biom J*. 2006; 48:1029–40. DOI: 10.1002/bimj.200610301 [PubMed: 17240660]
23. Collins GS, de Groot JA, Dutton S, Omar O, Shanyinde M, Tajar A, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Med Res Methodol*. 2014; 14:40.doi: 10.1186/1471-2288-14-40 [PubMed: 24645774]
24. Simon RM, Subramanian J, Li M-C, Menezes S. Using cross-validation to evaluate predictive accuracy of survival risk classifiers based on high-dimensional data. *Brief Bioinform*. 2011; 12:203–14. DOI: 10.1093/bib/bbr001 [PubMed: 21324971]
25. Debray TPA, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KGM. A new framework to enhance the interpretation of external validation studies of clinical prediction models. *J Clin Epidemiol*. 2015; 68:279–89. DOI: 10.1016/j.jclinepi.2014.06.018 [PubMed: 25179855]

Highlights

- A novel and efficient statistical method for prognosis of progression-free survival in mild cognitive impairment patients using multiple longitudinal markers.
- The prognostic model was evaluated via both internal and external validations.
- The longitudinal profiles of multiple markers, all of which are routinely collected in observational studies, could enhance the prognostic accuracy of Alzheimer's disease, if they are properly included in the prognostic model
- A prognostic index produced by the prognostic model is useful for monitoring disease progression and identifying patients for clinical trial recruitment.

Research in context

Systematic review

Publications on prognostic and predictive modeling using longitudinal markers in mild cognitive impairment (MCI) patients were searched in PubMed and in references and citations of relevant publications. Previous research has identified a few biomarkers, e.g. neurocognitive, imaging, and CSF, using logistic regression or Cox regression. No previous methods leveraged both multiple longitudinal data and time-to-event data for predicting AD progression.

Interpretation

Our objective was to propose a prognostic model, which relies on multiple longitudinal markers, for prediction of progression-free survival in MCI patients. We assessed the prognostic performance of a wide breadth of longitudinal markers using both internal and external validations.

Future directions

There are still opportunities to further optimize prognostic models by using newly discovered AD related markers. For substantial advances beyond the work presented here, clinical trial data are encouraged to be made available for validating prognostic models.

Figure 1. Kaplan-Meier survival curves for risk groups of conversion from MCI diagnosis to AD diagnosis

PI_{ADNI} was calculated based on parameters estimated in Model 2a using the ADNI-1 dataset, with 12 MFPC scores derived from longitudinal information of ADAS-Cog 13, RAVLT immediate, FAQ, RAVLT learning, and MMSE. Shaded areas correspond to 95% confidence interval of the Kaplan-Meier curves.

Table 1
Descriptive statistics for variables measured at study entry of ADNI-1 and ADNI-2 participants with mild cognitive impairment (MCI)

	ADNI-1			ADNI-2		
	Progressed to AD during the study	Did not progress to AD during the study	Combined	Progressed to AD during the study	Did not progress to AD during the study	Combined
Total	198	181	379	94	330	424
Women	74 (37.37%)*	61 (33.7%)	135 (35.62%)	43 (45.74%)	159 (48.18%)	202 (47.64%)
Age (years)	74.49 (7.08)	75.06 (7.56)	74.76 (7.31)	72.56 (7.15)	71.51 (6.92)	71.74 (6.98)
Numb. of APOE ε4 alleles						
0	73 (36.87%)	102 (56.35%)	175 (46.17%)	29 (30.85%)	197 (59.70%)	226 (53.30%)
1	95 (47.98%)	62 (34.25%)	157 (41.42%)	50 (53.19%)	110 (33.33%)	160 (37.74%)
2	30 (15.15%)	17 (9.39%)	47 (12.4%)	15 (15.96%)	23 (6.97%)	38 (8.96%)
Education (years)	15.84 (2.86)	15.36 (3.19)	15.61 (3.03)	16.21 (2.52)	16.51 (2.64)	16.45 (2.62)
Time in study (years)	2.26 (1.74)	4.31 (2.9)	3.24 (2.57)	1.77 (1.1)	2.95 (1.16)	2.69 (1.24)
ADAS-Cog 13	20.85 (5.76)	16.32 (6.11)	18.67 (6.34)	21.24 (6.96)	12.21 (5.77)	14.22 (7.12)
RAVLT Immediate	28.19 (7.34)	33.69 (10.04)	30.82 (9.14)	29.48 (8.05)	40.75 (11.15)	38.25 (11.53)
RAVLT Learning	2.81 (2.23)	3.81 (2.42)	3.29 (2.37)	3.45 (2.60)	5.29 (2.53)	4.88 (2.65)
FAQ	5.20 (4.87)	2.34 (3.53)	3.83 (4.51)	5.77 (4.52)	1.40 (2.59)	2.36 (3.60)
MMSE	26.73 (1.74)	27.30 (1.78)	27 (1.78)	27.26 (1.80)	28.49 (1.57)	28.22 (1.70)
MidTemp (cm3)	17.82 (2.89)	19.62 (2.73)	18.68 (2.95)	19.25 (2.66)	20.58 (2.65)	20.28 (2.70)
Hippo (cm3)	6.11 (1.05)	6.70 (1.05)	6.39 (1.09)	6.41 (1.09)	7.33 (1.01)	7.13 (1.09)
FDG	5.86 (0.61)	6.14 (0.62)	60 (0.63)	1.17 (0.12)	1.30 (0.12)	1.27 (0.13)
t-tau	110.28 (52.25)	92.38 (55.28)	102.07 (54.26) [†]	119.09 (60.97)	71.50 (40.18)	82.29 (49.80)
Aβ ₄₂	149.16 (41.93)	178.97 (58.84)	162.82 (52.41)	147.58 (40.48)	187.47 (51.66)	178.43 (52.05)

* Data are mean (SD) or n (%).

[‡] CSF biomarkers were collected on a subset of ADNI-1 (n = 192) and ADNI-2 (n = 375) subjects.

Abbreviations: ADAS-Cog, Alzheimer's Disease Assessment Scale - Cognitive Subscale test; RAVLT, Rey Auditory Verbal Learning Test; FAQ, Functional Assessment Questionnaire; MMSE, Mini Mental State Examination; MidTemp, volume of middle temporal gyrus; Hippo, volume of hippocampus; FDG-PET, sum of mean glucose metabolism uptake in regions of angular, temporal, and posterior cingulate; t-tau, CSF level of total tau; A β ₄₂, CSF level of amyloid beta 1–42 peptide.

Table 2
Mean iAUC and integrated Brier score for 3 models with internal and external validation

Model 1: multivariate baseline markers including ADAS-Cog 13, RAVLT immediate, FAQ, MMSE, and RAVLT learning; Model 1a: multivariate baseline markers including ADAS-Cog 13, RAVLT immediate, FAQ, MMSE, RAVLT learning, volumetric MRI (hippocampus, MidTemporal), and FDG-PET; Model 2: multiple longitudinal neurocognitive markers defined by 12 multivariate FPC scores. All predictors were controlled by baseline age, gender, education and number of the *APOE* ϵ 4 alleles.

Study	Model 1			Model 1a			Model 2		
	iAUC _{INT}	iAUC _{EXT}	Brier _{INT}	iAUC _{INT}	iAUC _{EXT}	Brier _{INT}	iAUC _{INT}	iAUC _{EXT}	Brier _{EXT}
ADNI 1 (n = 379)	0.776			0.815			0.846		
ADNI 2 (n = 424)	0.862	0.882		0.875	0.885		0.909	0.933	
ADNI 1 (n = 379)			Brier _{INT}	Brier _{INT}	Brier _{EXT}	Brier _{INT}	Brier _{INT}		Brier _{EXT}
			0.155	0.138			0.130		
ADNI 2 (n = 424)	0.223	0.064		0.207	0.057		0.197	0.052	

Table 3

Model comparisons with neurocognitive and sets of imaging biomarkers.

Study	Model 2a [†]			Model 2b			Model 2c			Model 2d		
	iAUC _{INT}	iAUC _{EXT}	iAUC _{INT}	iAUC _{INT}	iAUC _{EXT}	iAUC _{INT}	iAUC _{INT}	iAUC _{EXT}	iAUC _{EXT}	iAUC _{INT}	iAUC _{INT}	iAUC _{EXT}
ADNI-1 (n = 120) *	0.859		0.870		0.869				0.866			
ADNI-2 (n = 201)	0.878	0.927	0.883	0.923	0.890	0.919			0.899			0.918

* Model testing was limited to subsets of individuals who had all markers of interest available (e.g. those in Model 2d).

[†] Base model with 12 MFPC scores derived from longitudinal information of ADAS-Cog 13, RAVLT immediate, RAVLT learning, FAQ, and MMSE. Model 2b through Model 2d, longitudinal information of volumetric MRI (hippocampus, MidTemporal), and FDG-PET were added to the MFPC score calculation successively and in combination, e.g., Model 2b included hippocampal volume, Model 2c included hippocampal and MidTemporal volumes, Model 2d included hippocampal, MidTemporal volumes and FDG-PET.