Journal of Alzheimer's Disease 85 (2022) 837–850
DOI 10.3233/JAD-215266
IOS Press

Survival Analysis in Cognitively Normal Subjects and in Patients with Mild Cognitive Impairment Using a Proportional Hazards Model with Extreme Gradient Boosting Regression

Boshra Khajehpiri[a], Hamid Abrishami Moghaddam[a], Mohamad Forouzanfar[a,b], Reza Lashgari[c], Jaime Ramos-Cejudo[e], Ricardo S. Osorio[d,e] and Babak A. Ardekani[d,*]; for the Alzheimer's Disease Neuroimaging Initiative[1]

[a]*Machine Vision and Medical Image Processing (MVMIP) Laboratory, Faculty of Electrical and Computer Engineering, K. N. Toosi University of Technology, Tehran, Iran*
[b]*Department of Systems Engineering, École de Technologie Supérieure, Université du Québec, Montreal, Quebec, Canada*
[c]*Institute of Medical Science and Technology, Shahid Beheshti University, Tehran, Iran*
[d]*The Nathan S. Kline Institute for Psychiatric Research, Orangeburg, NY, USA*
[e]*Department of Psychiatry, New York University (NYU) Grossman School of Medicine, New York, NY, USA*

Handling Associate Editor: Ali Ezzati

Accepted 22 October 2021
Pre-press 2 December 2021

Abstract.
Background: Evaluating the risk of Alzheimer's disease (AD) in cognitively normal (CN) and patients with mild cognitive impairment (MCI) is extremely important. While MCI-to-AD progression risk has been studied extensively, few studies estimate CN-to-MCI conversion risk. The Cox proportional hazards (PH), a widely used survival analysis model, assumes a linear predictor-risk relationship. Generalizing the PH model to more complex predictor-risk relationships may increase risk estimation accuracy.
Objective: The aim of this study was to develop a PH model using an Xgboost regressor, based on demographic, genetic, neuropsychiatric, and neuroimaging predictors to estimate risk of AD in patients with MCI, and the risk of MCI in CN subjects.
Methods: We replaced the Cox PH linear model with an Xgboost regressor to capture complex interactions between predictors, and non-linear predictor-risk associations. We endeavored to limit model inputs to noninvasive and more widely available predictors in order to facilitate future applicability in a wider setting.

*Correspondence to: Babak A. Ardekani, PhD, Center for Brain Imaging and Neuromodulation, The Nathan S. Kline Institute for Psychiatric Research, 140 Old Orangeburg Road, Orangeburg, NY 10962, USA. E-mail: ardekani@nki.rfmh.org.

[1]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (https://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: https://adni.loni.usc.edu/wpcontent/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

ISSN 1387-2877/$35.00

Results: In MCI-to-AD ($n = 882$), the Xgboost model achieved a concordance index (C-index) of 84.5%. When the model was used for MCI risk prediction in CN ($n = 100$) individuals, the C-index was 73.3%. In both applications, the C-index was statistically significantly higher in the Xgboost in comparison to the Cox PH model.

Conclusion: Using non-linear regressors such as Xgboost improves AD dementia risk assessment in CN and MCI. It is possible to achieve reasonable risk stratification using predictors that are relatively low-cost in terms of time, invasiveness, and availability. Future strategies for improving AD dementia risk estimation are discussed.

Keywords: Alzheimer's disease, brain, hippocampal atrophy, machine learning, magnetic resonance imaging, mild cognitive impairment, proportional hazards model, survival analysis, Xgboost

INTRODUCTION

The course of Alzheimer's disease (AD) typically spans three broad stages: a preclinical phase when patients may appear cognitively normal (CN) [1]; mild cognitive impairment (MCI) [2]; and dementia which progresses from mild to moderate to severe [3]. The ability to detect AD before the dementia stage is extremely important for early treatment [4–6], sample enrichment in clinical trials [7, 8], and disease management.

Multiple indicators of incipient AD dementia have been identified [9–11]. In recent years, research has focused on developing machine learning algorithms that are capable of combining multiple "weak" predictors of dementia measured simultaneously to obtain a more accurate risk assessment. The predominant machine learning approach has been the use of classification algorithms, where, for example, MCI subjects are dichotomized as stable or progressive [12–16]. Classification methods, however, cannot properly handle censored information and definitions of class membership can be subjective. More recently, survival analysis (SA) has been utilized as a more appropriate data analytic approach for AD risk stratification in MCI [17–21]. As compared to MCI-to-AD progression, few studies have researched conversion from normal cognition to MCI using SA methods [22, 23].

The most commonly used SA method in this context has been the Cox proportional hazards (PH) model [24]. However, the Cox PH model makes two restrictive assumptions: firstly, that the hazard functions of different individuals have the same shape in time and differ only by a proportionality factor; and secondly, that the proportionality factor is a function of a linear combination of the predictor features. We hypothesize that relaxing these assumptions will lead to more flexible models with better prediction performance.

In the current work, we generalize the linear regression assumption of the Cox PH model by extreme gradient boosting regression (Xgboost) [25]. Xgboost is a recently developed and very successful machine learning method that is applied to learn a mapping from the space of the feature vectors to the space of positive real numbers, representing the proportionality factor in the PH model. Studies have shown that CN individuals with apolipoprotein E (*APOE*) $\varepsilon 4$ allele are more likely to convert to MCI/AD, but sex modifies this effect by influencing the conversion risk more strongly in women [26]. The hippocampal atrophy rates have been shown to be faster in women than in men, but only among those with underlying AD pathology [27]. The right hippocampal parenchymal fraction (HPF) value has been shown to be greater than the left HPF in CN, and moreover, this right > left asymmetry is more pronounced in men than in women [28]. The HPF also non-linearly decreases with age [28]. Scores in memory tests are higher in CN and MCI females compared to males but this sex difference fades with disease progression. This may suggest a delay in the onset of cognitive decline or diagnosis and/or a faster trajectory of cognitive decline in females [29]. Irrespective of diagnosis, cerebrospinal fluid (CSF) tau-pathology is disproportionately elevated in female carriers of *APOE* $\varepsilon 4$ compared to males. In contrast, male carriers of *APOE* $\varepsilon 4$ have reduced levels of CSF amyloid-β compared to females. We postulated that the decision tree-based Xgboost learning algorithm may be able to more accurately capture interactions and non-linear relationships between AD predictors and risk. Therefore, we hypothesized that replacing the restrictive linear regression assumption of the Cox PH model by gradient boosting regression would lead to significantly more accurate risk assessments.

In this study, we focused on a set of predictors that are not prohibitively invasive, expensive to measure, or limited in availability. Since hippocampal neuronal

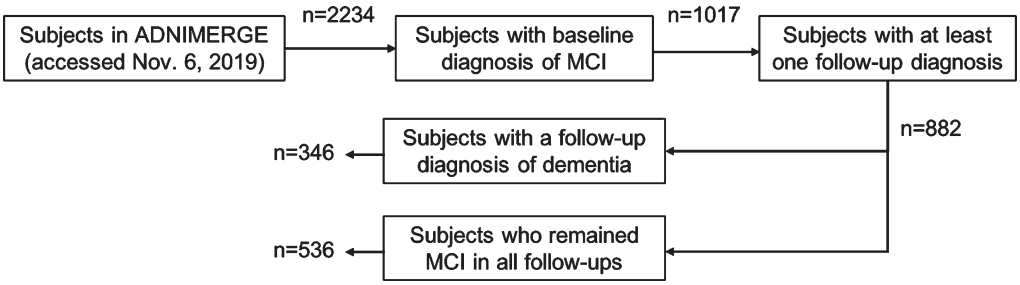

Fig. 1. Cohort selection flowchart.

injury is one the earliest indicators cognitive decline in CN, estimated to start over a decade before MCI [30], these predictors included the bilateral HPF, a marker of hippocampal structural integrity derived from 3D structural MRI [27, 31]. HPF can be computed efficiently (seconds) and reproducibly from raw MRI scans without need for any pre-processing, image analysis expertise, or manual intervention.

In summary, we used a novel SA method based on Xgboost regression to stratify the risk of progression from MCI to AD dementia, as well as the risk of conversion from CN to MCI. Relatively very few studies have considered the latter problem. We also utilize the HPF, a novel marker of hippocampal neuronal injury, which has not been previously used as a marker in SA. Finally, we endeavored to limit our set of predictors to those that are relatively less expensive, non-invasive, and more widely available.

METHODS

Study subjects

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (https://adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. For up-to-date information, see https://www.adni-info.org.

The method used for our cohort selection is shown in Fig. 1. The file ADNIMERGE.csv was accessed on November 6, 2019. It included 14,451 records from 2,234 participants. We only retained the records from the 1,017 participants with a baseline diagnosis of MCI. We further restricted the cohort to the 882 participants with at least one follow-up assessment. The final cohort consisted of 346 individuals (211 males; 135 females) with a subsequent diagnostic progression to AD dementia and 536 individuals (316 males; 220 females) whose diagnosis remained MCI.

Baseline diagnostic criteria for individuals with MCI were: Mini-Mental State Examination (MMSE) scores between 24–30 (inclusive), a subjective memory concern reported by subject, informant, or clinician, objective memory loss measured by education adjusted scores on delayed recall of one paragraph from the Wechsler Memory Scale Logical Memory II, a Clinical Dementia Rating (CDR) of 0.5, absence of significant levels of impairment in other cognitive domains, essentially preserved activities of daily living, and an absence of dementia. A potential progression of diagnosis from MCI to AD dementia was initially triggered by a site physician, confirmed by a clinical monitor, and finalized by the consensus of the ADNI Conversion Committee. The diagnostic criteria for AD were: MMSE scores between 20–26 (inclusive), CDR of 0.5 or 1.0, and meeting the NINCDS/ADRDA criteria for probable AD.

In addition, we selected 100 subjects (53 males; 47 females) with a baseline diagnosis of CN with at least one follow-up assessment. The diagnosis of 50 subjects remained CN in all their follow-up visits (right-censored subjects). The remaining 50 subjects converted to MCI sometime after their baseline visit. This cohort was used to assess the performance of our SA model in predicting the risk of CN-to-MCI conversion.

Predictor variables

In addition to their survival time and diagnosis status, 14 predictor variables were associated with each participant at baseline. These are summarized in

Table 1
Elements of the 14-dimensional feature space compared between progressive and right-censored MCI participants

		progressive ($n=346$)	right-censored ($n=536$)	p
1	Age (y)	73.9 ± 7.14	72.5 ± 7.67	$<0.01^{\ddagger}$
2	Sex (M/F)	211/135	316/220	$0.59^{\dagger}$
3	Education (y)	15.90 ± 2.76	15.98 ± 2.87	$0.52^{\dagger}$
4	APOE $\varepsilon 4$ (+/−)	224/122	224/312	$<0.001^{\dagger}$
5	RHPF	0.75 ± 0.10	0.82 ± 0.09	$<0.001^{\ddagger}$
6	LHPF	0.74 ± 0.10	0.81 ± 0.09	$<0.001^{\ddagger}$
7	B_0 (1.5 T/3 T)	210/136	288/248	$<0.05^{\dagger}$
8	CDR-SB	1.89 ± 0.95	1.25 ± 0.73	$<0.001^{\ddagger}$
9	FAQ	5.09 ± 4.72	1.91 ± 3.04	$<0.001^{\ddagger}$
10	ADAS-13	20.37 ± 6.31	14.38 ± 5.91	$<0.001^{\ddagger}$
11	ADNI-MEM	-0.21 ± 0.53	0.41 ± 0.63	$<0.001^{\ddagger}$
12	ADNI-EF	-0.12 ± 0.86	0.41 ± 0.85	$<0.001^{\ddagger}$
13	ADNI-VS	-0.17 ± 0.77	0 ± 0.73	$<0.001^{\ddagger}$
14	ADNI-LAN	-0.07 ± 0.74	0.33 ± 0.76	$<0.001^{\ddagger}$

Non-categorical variable values are shown as mean ± standard deviation. *APOE* $\varepsilon 4$, presence (+) or absence (–) of an $\varepsilon 4$ allele of the Apolipoprotein E gene; RHPF, right hippocampal parenchymal fraction; LHPF, left hippocampal parenchymal fraction; B_0, scanner magnetic field strength; CDR-SB, Clinical Dementia Rating – sum of boxes; FAQ, Functional Activities Questionnaire; ADAS-13, 13-item AD Assessment Scale–cognitive subscale; ADNI-MEM, composite score for memory; ADNI-EF, composite score for executive function; ADNI-VS, composite score for visuospatial functioning; ADNI-LAN, composite score of language. †chi-squared; ‡Mann Whitney U Test

Table 2
Elements of the 14-dimensional feature space compared between converted and right-censored CN participants

		converted ($n=50$)	right-censored ($n=50$)	p
1	Age (y)	76.0 ± 4.59	75.5 ± 4.42	$0.48^{\ddagger}$
2	Sex (M/F)	27/23	26/24	$1^{\dagger}$
3	Education (y)	15.94 ± 2.51	15.80 ± 3.14	$0.81^{\ddagger}$
4	APOE $\varepsilon 4$ (+/−)	15/35	11/39	$0.49^{\dagger}$
5	RHPF	0.80 ± 0.07	0.84 ± 0.05	$0.003^{\ddagger}$
6	LHPF	0.79 ± 0.07	0.83 ± 0.06	$0.007^{\ddagger}$
7	B_0 (1.5 T/3 T)	47/3	50/0	$0.24^{\dagger}$
8	CDR-SB	0.01 ± 0.07	0.02 ± 0.10	$0.57^{\ddagger}$
9	FAQ	0.40 ± 1.11	0.02 ± 1.14	$0.007^{\ddagger}$
10	ADAS-13	10.68 ± 4.32	8.31 ± 3.71	$0.006^{\ddagger}$
11	ADNI-MEM	0.86 ± 0.52	1.16 ± 0.54	$0.007^{\ddagger}$
12	ADNI-EF	0.68 ± 0.71	0.75 ± 0.79	$0.58^{\ddagger}$
13	ADNI-VS	0.31 ± 0.57	0.12 ± 0.64	$0.11^{\ddagger}$
14	ADNI-LAN	0.58 ± 0.61	0.89 ± 0.72	$0.08^{\ddagger}$

Non-categorical variable values are shown as mean ± standard deviation. *APOE* $\varepsilon 4$, presence (+) or absence (–) of an $\varepsilon 4$ allele of the Apolipoprotein E gene; RHPF, right hippocampal parenchymal fraction; LHPF, left hippocampal parenchymal fraction; B_0, scanner magnetic field strength; CDR-SB, Clinical Dementia Rating – sum of boxes; FAQ, Functional Activities Questionnaire; ADAS-13, 13-item AD Assessment Scale–cognitive subscale; ADNI-MEM, composite score for memory; ADNI-EF, composite score for executive function; ADNI-VS, composite score for visuospatial functioning; ADNI-LAN, composite score of language. †chi-squared; ‡Mann Whitney U Test.

Table 1 for the MCI subjects and can be divided into three groups: (1–4) demographic predictors; (5–7) MRI measures; and (8–14) neuropsychiatric/clinical test scores. The demographic variables were age, sex, years of education, and presence/absence of the $\varepsilon 4$ allele of the *APOE* gene. The MRI and neuropsychiatric/clinical predictors are described below.

Summary statistics by diagnosis status of the 14 predictors are given in Table 1 for the MCI, and in Table 2 for the CN subjects. Chi-squared tests were used to determine significant differences in the observed frequencies for the categorical variables (sex, *APOE* $\varepsilon 4$, B_0). Mann–Whitney U tests were used to compare the distributions of the non-categorical variables between groups. Statistical tests were performed using R (version 3.6.3).

MRI-based features

All participants had two back-to-back three-dimensional (3D) T1-weighted structural MP-RAGE MRI scans acquired at baseline. The HPF was computed using the KAIBA software of the Automatic Registration Toolbox (ART) (https://www.nitrc.org/projects/art). For each participant, the right and left HPF (RHPF and LHPF) were computed from both scans and averaged separately to obtain a RHPF and a LHPF for each patient. Details of the HPF computation method have been described elsewhere [27, 31]. Briefly, a standardized volume of interest (VOI) is defined in the vicinity of each hippocampus using automatic registration and landmark detection algorithms. Then, the histogram of the voxel intensities within the VOI is automatically analyzed to estimate the HPF as the brain tissue fraction that is contained in the VOI. The MRI scanner's magnetic field strength (B_0) was used as an additional feature since it has been shown to slightly affect the HPF [28].

Neuropsychiatric/clinical features

Seven neuropsychiatric/clinical test scores were downloaded directly from ADNI for each participant: 1) Clinical Dementia Rating Sum of Boxes (CDR-SB) [32]; 2) Functional Activities Questionnaire (FAQ) [33]; 3) the 13-item Alzheimer Disease Assessment Scale-Cognitive Subscale (ADAS-13) [34]; 4–7) composite scores related to memory (ADNI-MEM) [35], executive functioning (ADNI-EF) [36], visuospatial functioning (ADNI-VS), and

language (ADNI-LAN) [37], which were derived from the ADNI neuropsychological battery using item response theory (IRT). The composite scores indicate the general competency of patients in each of the four different cognitive domains. IRT can be considered as a feature selection method in machine learning applications [38]. The composite scores integrate select elements from all cognitive and neuropsychiatric measures available in ADNI, briefly: ADNI-MEM (MMSE, Logical Memory I and II, Auditory Verbal Learning, ADAS-Cog); ADNI-EF (Digit Span, Category Fluency, Digit Symbol, Trails A & B, Clock Drawing); ADNI-VS (MMSE, Clock Drawing, ADAS-Cog); ADNI-LAN (MMSE, Category Fluency, Boston Naming Test, ADAS-Cog). We did not perform a separate neuropsychiatric/clinical feature selection in this research.

Survival analysis

SA, also known as time-to-event analysis, is a collection of statistical techniques used to analyze data whose response variable is time until an event occurs [39]. This time is a random variable T referred to as the "survival time" representing the duration between the start of monitoring until a designated event is observed. In the current study, the survival time T was taken to be the time difference in months between a subject's ADNI baseline exam and the visit in which their diagnostic status changed from MCI to AD dementia, or from CN to MCI. In SA, the outcome random variable T is non-negative with a probability distribution $f(t)$ that is often skewed. The main objective of SA is to estimate this distribution using the survival data.

Survival data

The survival data in the current study comes from a cohort of 882 individuals who were diagnosed as amnestic MCI at their baseline ADNI visit and were followed longitudinally. From these, 346 individuals were diagnosed as having AD dementia at a follow-up visit. For these cases, the survival status $\delta_i = 1$, where i is the subject index, and the event-time is recorded as t_i and considered to be an observation of the random variable T. For the remaining 536 subjects, who at their last study visit still did not meet the diagnostic criteria for AD dementia, the survival status $\delta_i = 0$ and their survival time t_i was considered *right-censored*. That is, we do not know the actual time-to-event for these individuals and the only information that we have is that their time-to-event is greater than t_i. Regardless of their survival status, a 14-dimensional feature vector x_i was collected from each subject at their baseline visit. Thus, the entirety of the survival data in this study was a collection of ordered triplets $\{x_i, t_i, \delta_i)\}(i = 1, 2 \ldots, 882)$. Using these data, we estimated the *survival curves* under different conditions as well as fitting *proportional hazards models* as described in the following two subsections.

Survival curves

Survival curves $S(t)$ are functions of time that vary between 0 and 1. They represent the probability that the time-to-event is greater than or equal to t, that is: $S(t) = Pr(T \geq t)$. Therefore, in terms of the probability distribution of T, $f(t)$, the survival function is given by:

$$S(t) = \int_t^\infty f(\tau) d\tau$$

We used the non-parametric method of Kaplan-Meyer (KM) [40] to estimate the survival curve for the entire cohort. In addition, to investigate the effect of each risk factor on survival probability, we estimated KM survival curves for different categorical groups (e.g., *APOE* ε4+ versus *APOE* ε4–), and for subgroups obtained by dichotomizing according to specific threshold values of the continuous variable (e.g., age > 73 versus age ≤73). Group differences in KM curves were assessed using the non-parametric log-rank test [41].

Proportional hazards models

In SA, hazard models are used to estimate the influence of predictor variables x on survival probabilities. For this purpose, it is useful conceptually to consider the *conditional* probability distribution of the time-to-event random variable $f(t|x)$. Accordingly, the probability that the time-to-event is greater than or equal to t given predictors x at baseline is given by the *conditional* survival curve $S(t|x)$:

$$S(t|x) = \int_t^\infty f(\tau|x) d\tau$$

The *hazard function* $h(t|x)$ represents the instantaneous danger or hazard of the occurrence of the event

at time t given x and is defined as:

$$h(t|x) = \frac{f(t|x)}{S(t|x)} = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T \leq t + \Delta t | T \geq t, x)}{\Delta t}$$

The main assumption in the *proportional* hazards (PH) model is that $h(t|x)$ is separable into a product of two functions: one that is a function of t only, denoted by $h_0(t)$, and one that is a function of x only, given by $e^{\lambda(x)}$, that is: $h(t|x) = h_0(t)e^{\lambda(x)}$. In the Cox PH model [24], the function $\lambda(x)$ is modeled as linear combination of the predictor variables, that is, $\lambda(x) = \sum_k \beta_k x^{(k)}$ where $x^{(k)}$ represents the kth element of predictor array x. The model is fitted by estimating the regression parameters β_k using the survival data by maximizing the partial likelihood function. Other PH models can be obtained by using other models for $\lambda(x)$. In this work, in addition to fitting the Cox PH model to our data, we utilized the tree-based Xgboost model [25] to learn the mapping $\lambda(x)$ and refer to the resulting algorithm as the Xgboost PH model. The Xgboost model amounts to a weighted sum of regression trees, which are its non-linear building blocks. Therefore, the model is capable of capturing non-linear predictor-risk relationships. Regression trees are split at every node based on binary/binarized predictor variables (sex, *APOE4*, age, etc.). Therefore, the model also suitably captures interaction effects.

Model training and evaluation

We used the Harrell's concordance index (C-index) [42] for comparing Cox PH and Xgboost PH models' performance as well as selecting the hyperparameters of the Xgboost algorithm. This criterion is one of the most commonly used accuracy indexes in the SA domain and can handle censored data. The C-index has a value between 0 and 1 and can be considered a measure that indicates how well a model can correctly rank the survival data. Briefly, all possible pairs of subjects for whom we can ascertain the ranking of their time-to-event are presented to the algorithm which estimates their proportional instantaneous hazard given x at baseline. The C-index is the proportion of subject pairs in which the model rankings matches the known time-to-event rankings.

For the MCI data, model performance was estimated using 10-fold cross-validation, dividing the survival data into 90% training and 10% testing sets. In Xgboost, the hyperparameters were estimated by further dividing the training set into 75% for learning and 25% for validation. We used the model trained on MCI data to stratify risk in the 100 CN subjects and computed the C-index for this cohort.

In the Xgboost implementation, it is possible to start with a predefined model and revise the model (adjust tree parameters) to adapt to a new set of training data. This operation is akin to "transfer learning" in machine learning. It can serve as a backward transition from the MCI-to-AD model to an CN-to-MCI model without requiring a large number of training data. We used this method to adapt the MCI-to-AD model for CN-to-MCI risk stratification and evaluated its performance using 5-fold cross validation on the 100 CN subject.

Variable importance

For the Xgboost PH model, the importance of each input feature in predicting the final results was computed based on the Xgboost gain parameter. For each predictor, the gain is an indication of how much splitting on that feature has improved class discrimination in the overall model. Variable ranking was performed by applying the gain calculation algorithm and averaging the values over replications of all the folds. Finally, the relative importance of each predictor variable was computed by dividing their gain value by the highest gain (i.e., the gain related to the most effective feature).

For the Cox PH model, variable importance was obtained by constructing models on each input feature separately and subtracting the resulting C-indices of each from the C-index value of the original model (which included all variables). The relative importance of each predictor was calculated by scaling the differences so that the most informative feature (with the lowest difference) receives the relative importance of 1.

RESULTS

Predictor variables in progressive versus right-censored groups

In this study 882 amnestic MCI subjects were followed longitudinally, of whom 346 (39%) progressed to AD dementia with mean (SD) survival time of 30.2 (24.6) months. The remaining 536 (61%) subjects did not progress to dementia with a mean (SD) right-censored survival time of 47.8 (33.4) months.

Table 1 shows differences between progressive and right-censored MCI groups in each of the 14

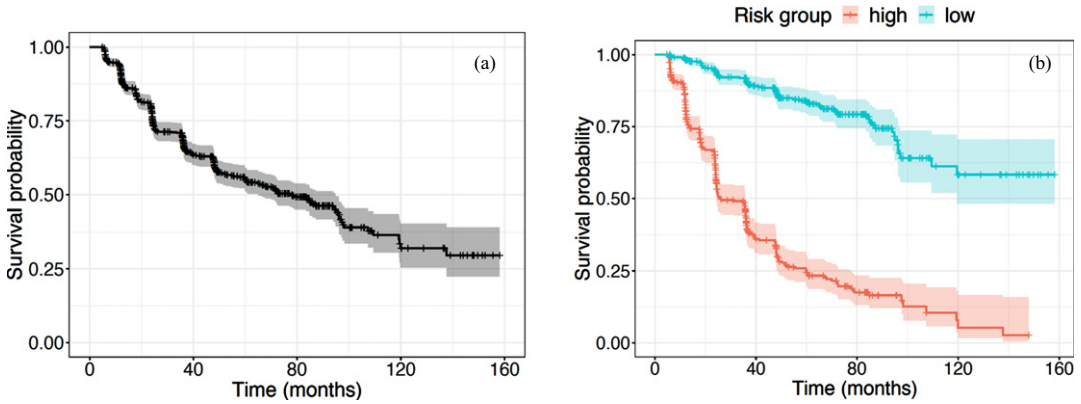

Fig. 2. (a) KM survival curve for the entire MCI cohort. (b) KM survival curves for MCI groups stratified by risk using the Xgboost PH model. The threshold between high and low risk groups was set to the median hazard level.

predictors that comprise x at baseline. The progressive group was older than the right-censored group at baseline: 73.9 ± 7.14 versus 72.5 ± 7.67 years ($p < 0.01$). The groups did not differ significantly by sex or years of education. The percentage of *APOE* ε4+ subjects in the progressive group (64.7%) was significantly higher than the censored group (59.0%) ($p < 0.001$). The progressive group had lower HPF bilaterally ($p < 0.001$). A higher proportion of the progressive group had been scanned at 1.5 T ($p < 0.05$). The progressive group had higher (worse) CDR-SB, FAQ, and ADAS-13 scores ($p < 0.001$), and their performance was significantly lower than the right-censored group in all four composite cognitive measures (ADNI-MEM, ADNI-EF, ADNI-VS, ADNI-LAN) ($p < 0.001$).

Table 2 shows differences between converted and right-censored CN groups in each of the 14 predictors that comprise x at baseline. The groups did not differ significantly in age, sex, years of education, percentage of *APOE* ε4 positivity, CDR-SB, ADNI-EF, ADNI-VS, or ADNI-LAN. The converted group had significantly lower HPF bilaterally ($p < 0.01$ for either side). The converted group had significantly higher (worse) FAQ ($p < 0.01$) and ADAS-13 scores ($p < 0.01$), and their performance was significantly lower than the right-censored group in ADNI-MEM ($p < 0.01$).

Kaplan-Meyer (KM) survival curves

For the entire cohort ($n = 882$), the KM estimated survival curve is shown in Fig. 2a. The estimated median survival time was 77.2 (95% CI: 65.9–95.9) months, while the 1st quartile (i.e., time for survival probability of 0.75) was estimated to be 24.3 (95% CI: 23.9–25.7) months.

Figure 3 shows the KM survival curves for the MCI cohort dichotomized with respect to each of the 14 predictors in x. Dichotomized by median age (73 years) at baseline, the younger group had significantly higher survival probabilities ($p < 0.001$) (Fig. 3a). There was no difference in survival probabilities between males and females (Fig. 3b). There was trend level ($p = 0.11$) higher survival rates in individuals with more than 16 years of education (Fig. 3c). *APOE* ε4 + individuals had significantly lower survival rates compared to the *APOE* ε4– group ($p < 0.001$) (Fig. 3d). Dichotomized by thresholding the RHPF at 0.72 and the LHPF at 0.71 (2 standard deviations below mean values estimated in cognitively normal elderly [28]), groups showed significantly different survival rates based on both left and right HPF values ($p < 0.001$) (Fig. 3e, f). Individuals scanned at 1.5 T showed trend level ($p = 0.095$) lower survival probabilities (Fig. 3g). Dichotomizing by median values of both CDR-SB and FAQ variables resulted in groups with significantly different survival rates ($p < 0.001$) (Fig. 3h, i). Thresholding ADAS-13 at 16.7 (approximately 2 standard deviations below mean values found in [28] for cognitively normal elderly) also resulted in groups with significantly different survival rates ($p < 0.001$) (Fig. 3j). Finally, grouping by positive versus negative values of each of the four composite cognitive scores (ADNI-MEM, ADNI-EF, ADNI-VS, ADNI-LAN) resulted in statistically significant differences in survival curves ($p < 0.001$) (Fig. 3k–n). Table 3 shows the median survival times for each of the 28 subgroups in Fig. 3.

Model performance

Both Cox PH and Xgboost PH models use the entire predictor array obtained from subjects at baseline to stratify their risk of AD dementia (for the MCI cohort; $n = 882$) or risk of conversion to MCI (for the CN cohort; $n = 100$). In MCI by 10-fold cross-validation, the average C-index of the Xgboost PH model 84.5% was significantly higher than that of the Cox PH model 83.2% ($p = 0.014$). Removing the

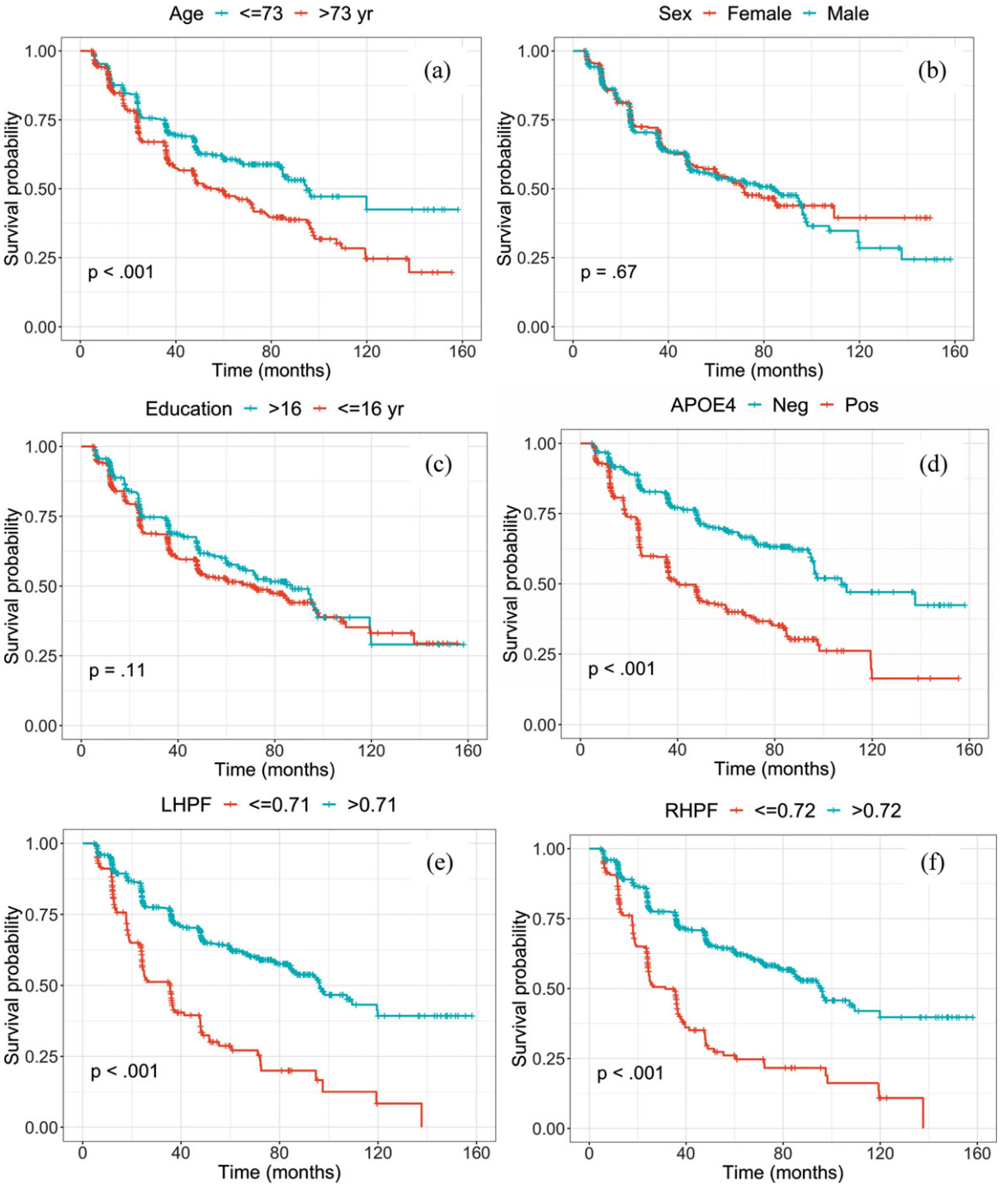

Fig. 3. *(Continued)*

left and right HPF from the dataset, the Xgboost PH model's average C-index decreased significantly to 83.7% ($p = 0.02$).

Figure 4a shows the relative importance of the 14 predictors according to their predictive role in stratifying dementia risk in the Xgboost PH model. ADNI-MEM, the composite score related to memory, was the most informative predictor, followed by FAQ, ADAS-13, and RHPF, respectively. Figure 4b shows the relative importance of the 14 predictors in the

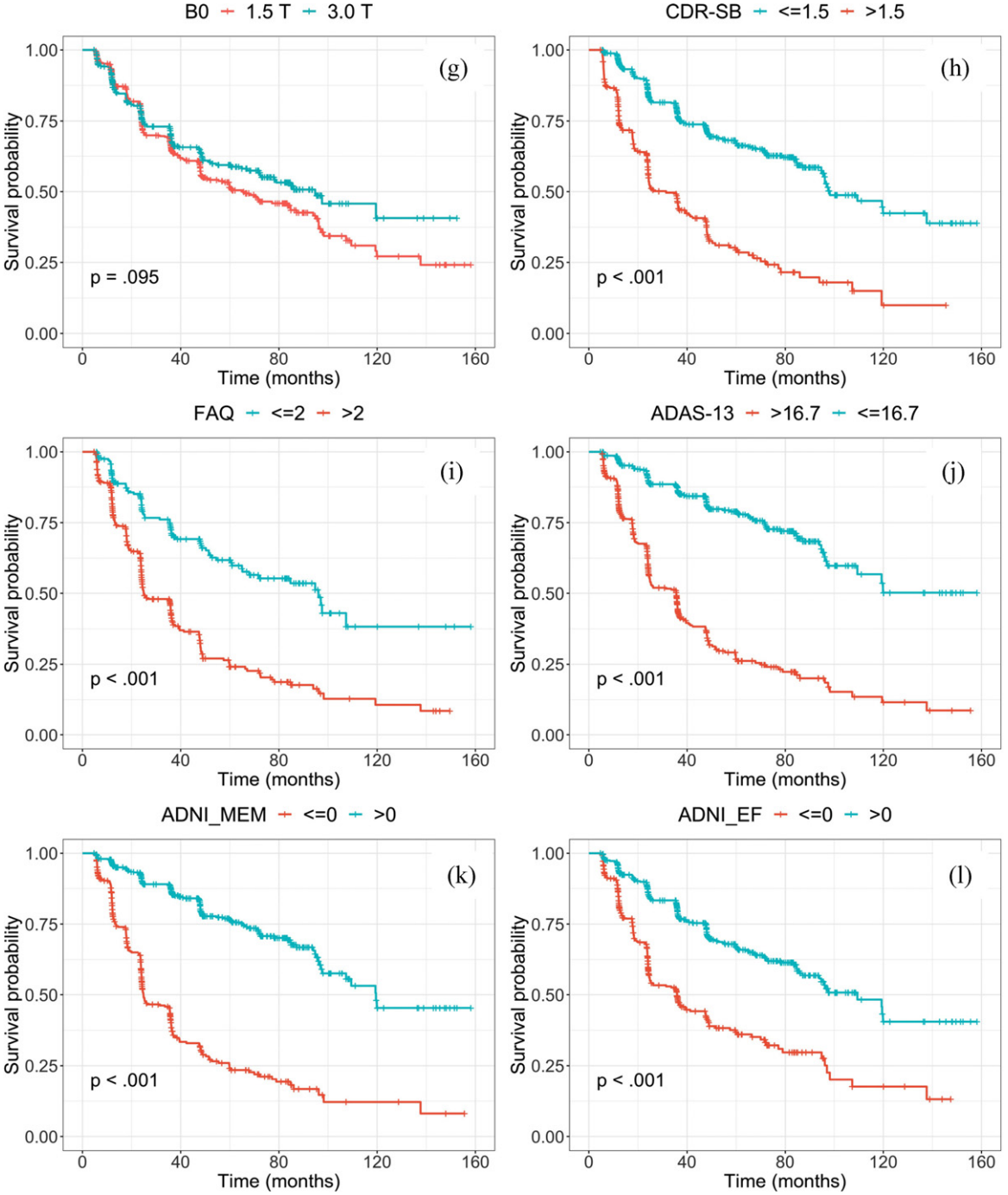

Fig. 3. *(Continued)*

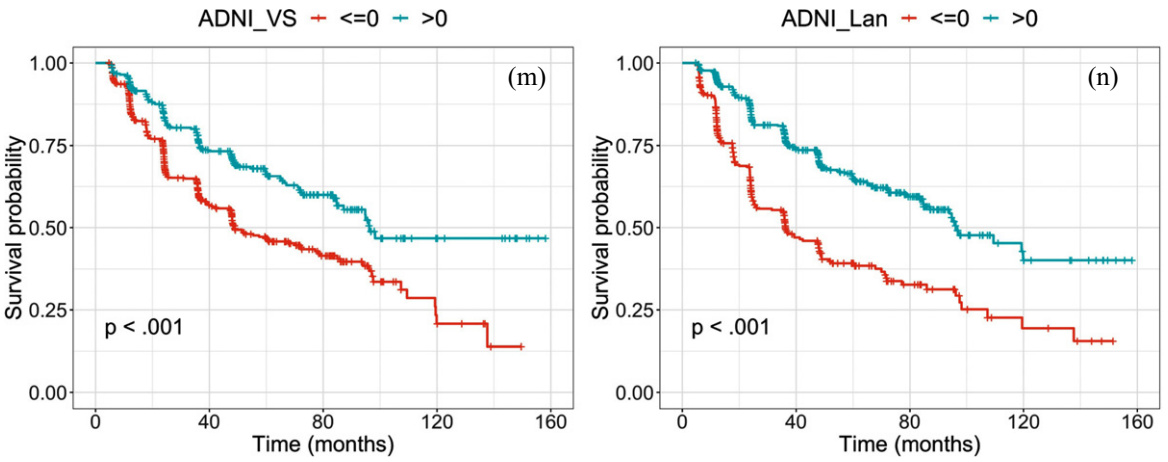

Fig. 3. KM survival curves in MCI groups by (a) age, (b) sex, (c) years of education, (d) *APOE* ε4 status, (e) right HPF, (f) left HPF, (g) scanner field strength, (h) CDR-SB, (i) FAQ, (j) ADAS-13, (k) ADNI-MEM, (l) ADNI-EF, (m) ADNI-VS, and (n) ADNI-LAN. Age, CDR-SB, and FAQ are dichotomized by thresholding at median value. RHPF, LHPF, and ADAS-13 are dichotomized by thresholding at 2 standard deviations below normal values derived in [28]. ADNI-MEM, ADNI-EF, ADNI-VS, and ADNI-LAN are dichotomized as positive versus negative.

Table 3
Median (95% CI) survival times for MCI groups stratified according to each element of the 14-dimensional predictor space

		Stratification	Higher-risk group	Lower-risk group	*p*
1	Age (y)	>73 versus ≤73	57.5 (47.9–70.1)	94.8 (84.8-NA)	<0.001
2	Sex (M/F)	F versus M	71.6 (59.8-NA)	84.8 (54.8–96.2)	0.67
3	Education (y)	≤16 versus >16	70.0 (49.0–96.1)	87.2 (67.8–119.3)	0.11
4	*APOE* ε4 (+/−)	+versus −	40.7 (36.1–48.9)	107.3 (96.1-NA)	<0.001
5	RHPF	≤0.71 versus >0.71	26.2 (24.1–36.3)	94.9 (84.8–119.9)	<0.001
6	LHPF	≤0.71 versos >0.71	35.6 (24.2–36.9)	96.2 (86.0–119.9)	<0.001
7	B_0 (1.5 T/3 T)	1.5T versus 3T	65.1 (50.3–87.2)	94.8 (72.3-NA)	0.095
8	CDR-SB	>1.5 versus ≤1.5	31.4 (24.3–39.3)	98.2 (96.1-NA)	<0.001
9	FAQ	>2 versus ≤2	25.0 (24.1–36.1)	96.1 (67.8-NA)	<0.001
10	ADAS-13	>16.7 versus ≤16.7	36.6 (24.8–36.2)	120.0 (109.4-NA)	<0.001
11	ADNI-MEM	− versus +	24.9 (24.1–35.6)	119.5 (97.5-NA)	<0.001
12	ADNI-EF	− versus +	35.9 (24.8–47.6)	109.4 (94.0-NA)	<0.001
13	ADNI-VS	− versus +	48.9 (47.8–71.8)	96.2 (84.8-NA)	<0.001
14	ADNI-LAN	− versus +	36.3 (35.4–48.1)	96.2 (87.2-NA)	<0.001

NA indicates that the upper limit could not be determined by the KM method. The *p*-values are obtained using the log-rank test.

Cox PH model. The orders of variables' importance are mostly consistent between the two models.

Figure 2b shows the KM survival curves of high-risk and low-risk MCI groups stratified based on their Xgboost PH risk being above or below the median risk in the cohort. The survival rates were significantly different based on a log-rank test ($p < 0.001$). The median (95% CI) survival times for the high-risk group was 25.3 (24.3–35.9) months.

We also applied the Cox PH and Xgboost PH models, trained by each of the 10 folds of the MCI cohort, to stratify the risk of conversion to MCI in the CN cohort. The Xgboost model performed better than the Cox model in every fold. The average C-index for was 69.6% for the Cox model and 73.3% for the Xgboost model. The different was statistically significant ($p < 10^{-6}$). When we used the trained MCI-to-AD model as an initial model in the Xgboost algorithm and adjusted the model parameters using additional subjects from the CN-to-MCI cohort, the C-index estimated by 5-fold cross-validation increased to 75.1% for the Xgboost model.

DISCUSSION

A key priority in the management of AD is the development of early intervention strategies that can stop or slow disease progression in the pre-dementia phase. Development and application of these preven-

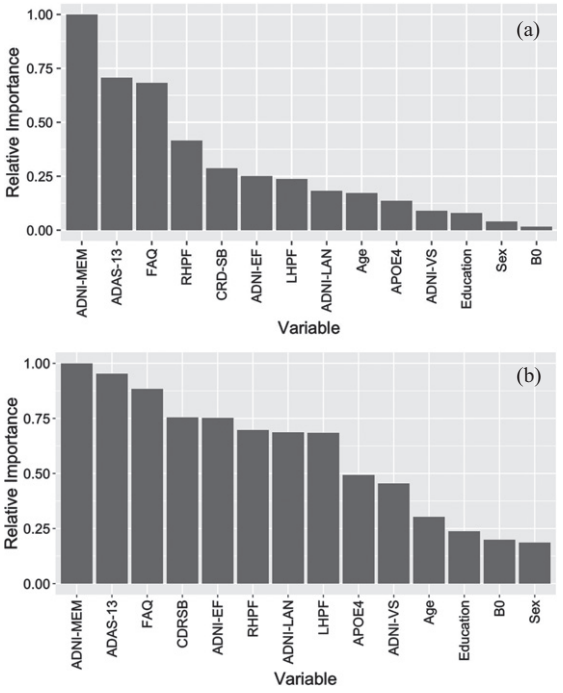

Fig. 4. Variable importance of dementia risk prediction in MCI: (a) Xgboost PH model, (b) Cox PH model.

tive approaches can only be effective if individuals at higher risk of AD dementia can be accurately identified through cost-effective, non-invasive, and widely available tests. Multiple markers have been identified that differentiate AD from normal aging [9–11]. However, each predictor by itself is a relatively weak indicator of AD in the pre-dementia phase. Therefore, the current consensus is to combine multiple concurrently measured predictors to obtain an overall strong disease indicator. Since AD predictors interact [26–29] and may be non-linearly associated with risk of progression, advanced machine learning algorithms are required to assess risk of decline (from MCI to dementia or from CN to MCI) from arrays of concurrently measured predictors. Here, we used the tree-based Xgboost, a relatively new and powerful machine learning method, to build a predictive model to estimate the risk of developing AD dementia for individuals with MCI in a SA framework. The developed model ranks patients based on their likelihood of progression to AD. The Xgboost PH model achieved a C-index of 84.5% which is significantly greater than the C-index of the Cox PH model (83.2%) ($p = 0.014$). This result indicates that there are interactions amongst predictors and non-linear predictor-risk relationship captured by the Xgboost regression algorithm beyond what is explained by the Cox linear regression model.

The difference between the models' performances (C-index) was more pronounced when applied to the CN cohort to estimate CN-to-MCI conversion risk (Xgboost: 73.3% versus Cox: 69.6%; $p < 10^{-6}$). This suggests that model advantages may reveal themselves more readily as the diagnosis task becomes more challenging. In general, survival predictions may be improved in one of two ways: 1) by applying more suitable models; and 2) by employing better disease indicators. In both cases, improvements will likely be incremental. Here, we have shown that applying the Xgboost regressor in the PH model results in improvement in risk assessment.

As shown in Fig. 2a, the median (95% CI) survival time for the entire ADNI amnestic MCI cohort was 77.2 (65.9–95.9) months. However, when stratified by high versus low risk using Xgboost PH, the median survival time in the high-risk group reduced to 25.3 (24.3–35.9) months. This result indicates the utility of our risk assessment method for sample enrichment in clinical trials to increase the likelihood of detecting positive therapeutic outcomes [7, 8].

Variable importance analysis on our Xgboost PH model (Fig. 4) indicated that ADNI-MEM is the most predictive feature, confirming the effect of AD on memory deficiency at early stages of the disease. Other high-ranking predictors were ADAS-13, FAQ and RHPF. Table 2 shows that these variables are also early indicators of impending cognitive decline in CN individuals.

The only imaging measure considered in this study was HPF, which can be easily calculated from T1-weighted MRI scans. We investigated the ability of this measure to improve the prediction performance and act as an effective biomarker of neuronal injury in AD. We demonstrated that the HPF had a significant contribution to the model's performance: C-index 84.5% versus 83.7% ($p = 0.02$). It was observed that the prediction power of RHPF is almost twice that of LHPF (Fig. 4). This concurs with the findings reported in [12] where the RHPF had a higher variable importance than the LHPF in a Random Forest model for stable versus progressive MCI classification. It has been shown that in normal aging the left hippocampus declines faster than the right hippocampus [28]. Therefore, the deterioration of the right hippocampus may be a more ominous indicator of underlying AD pathology in MCI.

It should be emphasized that although we utilized low-cost and non-invasive predictors, adding other

risk factors could potentially increase the strength of our proposed model. Until recently, Aβ and tau proteins could only be extracted through invasive procedures. Yet recently, it has been shown that information on these factors may be acquired through simple blood tests [45]. The data obtained from these tests have not been sufficiently collected, but the use and evaluation of their effect in subsequent studies are recommended. Furthermore, the procedure described in the current paper may serve as an initial assessment tool for patients with subjective memory complaints. If results indicate high-risk for conversion to MCI/dementia, then further tests (e.g., CSF analysis and Aβ and tau PET imaging) may be recommended to obtain a more accurate diagnostic picture.

Future strategies such as relaxing the PH assumption, including additional low-cost novel predictors, and using longitudinal data should be investigated to achieve further improvements in model accuracy leading to significant cost savings in clinical trials, long-term disease management, and patient care.

A limitation of this work was the assumption of non-informative censoring which may not be necessarily true. SA leads to reliable and unbiased results only if the non-informative data censoring condition is met (i.e., participants should not have left the study for a reason related to their disease status). This ought to be considered more carefully in future studies. Another limitation of this study was that the model trained for MCI-to-AD progression was applied for CN-to-MCI conversion. Ideally, for the latter application the model should be trained on CN subjects with likely improved results. This would require a larger cohort of longitudinally observed CN.

ACKNOWLEDGMENTS

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 (https://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.

Authors' disclosures available online (https://www.j-alz.com/manuscript-disclosures/21-5266r1).

REFERENCES

[1] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack CR Jr, Kaye J, Montine TJ, Park DC, Reiman EM, Rowe CC, Siemers E, Stern Y, Yaffe K, Carrillo MC, Thies B, Morrison-Bogorad M, Wagster MV, Phelps CH (2011). Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 280-292.

[2] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, Snyder PJ, Carrillo MC, Thies B, Phelps CH (2011) The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging -Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 270-279.

[3] McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, Mohs RC, Morris JC, Rossor MN, Scheltens P, Carrillo MC, Thies B, Weintraub S, Phelps CH (2011) The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 263-269.

[4] Gauthier SG (2005) Alzheimer's disease: The benefits of early treatment. *Eur J Neurol* **12** (Suppl 3), 11-16.

[5] Budd D, Burns LC, Guo Z, L'italien G, Lapuerta P (2011) Impact of early intervention and disease modification in patients with predementia Alzheimer's disease: A Markov model simulation. *Clinicoecon Outcomes Res* **3**, 189-195.

[6] Folch J, Busquets O, Ettcheto M, Sánchez-López E, Castro-Torres RD, Verdaguer E, Garcia ML, Olloquequi J, Casadesús G, Beas-Zarate C, Pelegri C, Vilaplana J, Auladell C, Camins A (2018) Memantine for the treatment of dementia: A review on its current and future applications. *J Alzheimers Dis* **62**, 1223-1240.

[7] Grill JD, Monsell SE (2014) Choosing Alzheimer's disease prevention clinical trial populations. *Neurobiol Aging* **35**, 466-471.

[8] Spencer BE, Digma LA, Jennings RG, Brewer JB; Alzheimer's Disease Neuroimaging Initiative and the A4 Study Team (2021) Gene- and age-informed screening for preclinical Alzheimer's disease trials. *Alzheimers Dement* **17**, 457-465.

[9] Craig-Schapiro R, Fagan AM, Holtzman DM (2009) Biomarkers of Alzheimer's disease. *Neurobiol Dis* **35**, 128-140.

[10] Tan CC, Yu JT, Tan L (2014) Biomarkers for preclinical Alzheimer's disease. *J Alzheimers Dis* **42**,1051-1069.

[11] Counts SE, Ikonomovic MD, Mercado N, Vega IE, Mufson EJ (2017) Biomarkers for the early detection and progression of Alzheimer's disease. *Neurotherapeutics* **14**, 35-53.

[12] Ardekani BA, Bermudez E, Mubeen AM, Bachman AH (2017) Prediction of incipient Alzheimer's disease dementia in patients with mild cognitive impairment. *J Alzheimers Dis* **55**, 269-281.

[13] Beheshti I, Demirel H, Matsuda H (2017) Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm. *Comput Biol Med* **83**,109-119.

[14] Hojjati SH, Ebrahimzadeh A, Khazaee A, Babajani-Feremi A (2018) Predicting conversion from MCI to AD by integrating rs-fMRI and structural MRI. *Comput Biol Med* **102**, 30-39.

[15] Facal D, Valladares-Rodriguez S, Lojo-Seoane C, Pereiro AX, Anido-Rifon L, Juncos-Rabadán O (2019) Machine learning approaches to studying the role of cognitive reserve in conversion from mild cognitive impairment to dementia. *Int J Geriatr Psychiatry* **34**, 941-949.

[16] Mubeen AM, Asaei A, Bachman AH, Sidtis JJ, Ardekani BA (2017) A six-month longitudinal evaluation significantly improves accuracy of predicting incipient Alzheimer's disease in mild cognitive impairment. *J Neuroradiol* **44**, 381-387.

[17] Li K, O'Brien R, Lutz M, Luo S; Alzheimer's Disease Neuroimaging Initiative (2018) A prognostic model of Alzheimer's disease relying on multiple longitudinal measures and time-to-event data. *Alzheimers Dement* **14**, 644-651.

[18] Spencer BE, Jennings RG, Brewer JB (2019) Combined biomarker prognosis of mild cognitive impairment: An 11-year follow-up study in the Alzheimer's Disease Neuroimaging Initiative. *J Alzheimers Dis* **68**, 1549-1559.

[19] Platero C, Tobar MC (2020) Longitudinal survival analysis and two-group comparison for predicting the progression of mild cognitive impairment to Alzheimer's disease. *J Neurosci Methods* **341**, 108698.

[20] Spooner A, Chen E, Sowmya A, Sachdev P, Kochan NA, Trollor J, Brodaty H (2020) A comparison of machine learning methods for survival analysis of high-dimensional clinical data for dementia prediction. *Sci Rep* **10**, 20410.

[21] Li H, Habes M, Wolk DA, Fan Y (2019) A deep learning model for early prediction of Alzheimer's disease dementia based on hippocampal magnetic resonance imaging data. *Alzheimers Dement* **15**, 1059-1070.

[22] Kryscio RJ, Schmitt FA, Salazar JC, Mendiondo MS, Markesbery WR (2006) Risk factors for transitions from normal to mild cognitive impairment and dementia. *Neurology* **66**, 828-832.

[23] Gallagher D, Kiss A, Lanctot KL, Herrmann N (2018) Toward prevention of mild cognitive impairment in older adults with depression: An observational study of potentially modifiable risk factors. *J Clin Psychiatry* **80**,18m12331.

[24] Cox DR (1972) Regression models and life-tables. *J Royal Stat Soc B* **34**,187-202.

[25] Chen T, Guestrin C (2016) Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785-794.

[26] Altmann A, Tian L, Henderson VW, Greicius MD (2014) Sex modifies the APOE-related risk of developing Alzheimer disease. *Ann Neurol* **75**, 563-573.

[27] Ardekani BA, Convit A, Bachman AH (2016) Analysis of the MIRIAD data shows sex differences in hippocampal atrophy progression. *J Alzheimers Dis* **50**, 847-857.

[28] Ardekani BA, Izadi NO, Hadid SA, Meftah AM, Bachman AH (2020) Effects of sex, age, and apolipoprotein E genotype on hippocampal parenchymal fraction in cognitively normal older adults. *Psychiatry Res Neuroimaging* **301**, 111107.

[29] Duarte-Guterman P, Albert AY, Barha CK, Galea LAM (2021) Sex influences the effects of APOE genotype and Alzheimer's diagnosis on neuropathology and memory. *Psychoneuroendocrinology* **129**, 105248.

[30] Bachman AH, Ardekani BA (2019) Change point analyses in prodromal Alzheimer's disease. *Biomark Neuropsychiatry* **3**, 100028.

[31] Goff DC, Zeng B, Ardekani BA, Diminich ED, Tang Y, Fan X, Galatzer-Levy I, Li C, Troxel AB, Wang J (2018) Association of hippocampal atrophy with duration of untreated psychosis and molecular biomarkers during initial antipsychotic treatment of first-episode psychosis. *JAMA Psychiatry* **75**, 370-378.

[32] O'Bryant SE, Waring SC, Cullum CM, Hall J, Lacritz L, Massman PJ, Lupo PJ, Reisch JS, Doody R (2008) Staging dementia using Clinical Dementia Rating Scale Sum of Boxes scores: A Texas Alzheimer's research consortium study. *Arch Neurol* **65**, 1091-1095.

[33] Pfeffer RI, Kurosaki TT, Harrah Jr CH, Chance JM, Filos S (1982) Measurement of functional activities in older adults in the community. *J Gerontol* **37**, 323-329.

[34] Mohs RC, Knopman D, Petersen RC, Ferris SH, Ernesto C, Grundman M, Sano M, Bieliauskas L, Geldmacher D, Clark C, Thal LJ (1997) Development of cognitive instruments for use in clinical trials of antidementia drugs: Additions to the Alzheimer's Disease Assessment Scale that broaden its scope. *Alzheimer Dis Assoc Disord* **11** (Suppl 2), 13-21.

[35] Crane PK, Carle A, Gibbons LE, Insel P, Mackin RS, Gross A, Jones RN, Mukherjee S, Curtis SM, Harvey D, Weiner M, Mungas D (2012) Development and assessment of a composite score for memory in the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Brain Imaging Behav* **6**, 502-516.

[36] Gibbons LE, Carle AC, Mackin RS, Harvey D, Mukherjee S, Insel P, Curtis SM, Mungas D, Crane PK (2012) A composite score for executive functioning, validated in Alzheimer's Disease Neuroimaging Initiative (ADNI) participants with baseline mild cognitive impairment. *Brain Imaging Behav* **6**, 517-527.

[37] Choi SE, Mukherjee S, Gibbons LE, Sanders RE, Jones RN, Tommet D, Mez J, Trittschuh EH, Saykin A, Lamar M, Rabin L, Foldi NS, Sikkes S, Jutten RJ, Grandoit E, Mac Donald C, Risacher S, Groot C, Ossenkoppele R (2020) Development and validation of language and visuospatial composite scores in ADNI. *Alzheimers Dement* **6**, e12072.

[38] Kline A, Kline T, Hossein Abad ZS, Lee J (2020) Novel feature selection for artificial intelligence using item response theory for mortality prediction. In *Proceedings of the 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society*, pp. 5729-5732.

[39] Kleinbaum DG, Klein M (2010) *Survival Analysis*, Springer, New York.

[40] Kaplan EL, Meier P (1958) Nonparametric estimation from incomplete observations. *J Am Stat Assoc* **53**, 457-481.

[41] Mantel N (1966) Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemother Rep* **50**,163-170.

[42] Harrell FE, Califf RM, Pryor DB, Lee KL, Rosati RA (1982) Evaluating the yield of medical tests. *JAMA* **247**, 2543-2536.

[43] Roberts R, Knopman DS (2013) Classification and epidemiology of MCI. *Clin Geriatr Med* **29**, 753-772.

[44] (2021) 2021 Alzheimer's disease facts and figures. *Alzheimers Dement* **17**, 327-406.

[45] Chong JR, Ashton NJ, Karikari TK, Tanaka T, Saridin FN, Reilhac A, Robins EG, Nai YH, Vrooman H, Hilal S, Zetterberg H, Blennow K, Lai MKP, Chen CP (2021) Plasma P-tau181 to Aβ42 ratio is associated with brain amyloid burden and hippocampal atrophy in an Asian cohort of Alzheimer's disease patients with concomitant cerebrovascular disease. *Alzheimers Dement* **17**, 1649-1662.