Received: 28 October 2020 | Revised: 1 December 2020 | Accepted: 8 December 2020 | Published online: 17 March 2021

DOI: 10.1002/trc2.12135

Alzheimer's & Dementia
Translational Research & Clinical Interventions

RESEARCH ARTICLE

Predicting amyloid risk by machine learning algorithms based on the A4 screen data: Application to the Japanese Trial-Ready Cohort study

Kenichiro Sato[1] | Ryoko Ihara[2,3] | Kazushi Suzuki[3] | Yoshiki Niimi[3] | Tatsushi Toda[1] | Gustavo Jimenez-Maggiora[4] | Oliver Langford[4] | Michael C. Donohue[4] | Rema Raman[4] | Paul S. Aisen[4] | Reisa A. Sperling[5] | Atsushi Iwata[1,6] | Takeshi Iwatsubo[2,3]

[1] Department of Neurology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

[2] Department of Neuropathology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

[3] Unit for Early and Exploratory Clinical Development, The University of Tokyo Hospital, Tokyo, Japan

[4] Alzheimer's Therapeutic Research Institute, University of Southern California, San Diego, California, USA

[5] Brigham and Women's Hospital, Boston, Massachusetts, USA

[6] Department of Neurology, Tokyo Metropolitan Geriatric Medical Center Hospital, Tokyo, Japan

Correspondence
Takeshi Iwatsubo, 7-3-1 Hongo Bunkyoku, Tokyo 113-0033, Japan.
Email: iwatsubo@m.u-tokyo.ac.jp

Funding information
Japan Agency for Medical Research and Development, Grant/Award Numbers: JP17DK0207028, JP19DK0207048

Abstract

Background: Selecting cognitively normal elderly individuals with higher risk of brain amyloid deposition is critical to the success of prevention trials for Alzheimer's disease (AD).

Methods: Based on the Anti-Amyloid Treatment in Asymptomatic Alzheimer's Disease study data, we built machine-learning models and applied them to our ongoing Japanese Trial-Ready Cohort (J-TRC) webstudy participants registered within the first 9 months ($n = 3081$) of launch to predict standard uptake value ratio (SUVr) of amyloid positron emission tomography.

Results: Age, family history, online Cognitive Function Instrument and CogState scores were important predictors. In a subgroup of J-TRC webstudy participants with known amyloid status ($n = 37$), the predicted SUVr corresponded well with the self-reported amyloid test results (area under the curve = 0.806 [0.619–0.992]).

Discussion: Our algorithms may be usable for automatic prioritization of candidate participants with higher amyloid risks to be preferentially recruited from the J-TRC webstudy to in-person study, maximizing efficiency for the identification of preclinical AD participants.

KEYWORDS
J-TRC, machine learning, online recruitment, preclinical Alzheimer's disease, Trial-Ready Cohort, webstudy

1 | INTRODUCTION

Recent advances in the development of mechanism-based therapeutics (i.e., disease-modifying therapies) against Alzheimer's disease (AD), for example, anti-amyloid beta (Aβ) drugs, have illuminated preclinical AD, the earliest clinical stage of AD that precedes AD dementia or mild cognitive impairment (i.e., prodromal AD).[1] Preclinical AD is defined as an asymptomatic stage with evidence of the earliest pathological changes of AD in the brain, that is, positive biomarker signatures of amyloid deposition.[1–4] Because ≈30% of cognitive normal elderly individuals in the Western population[5,6] and ≈24% in the Japanese population[7] have been estimated to have elevated Aβ as confirmed

by amyloid positron emission tomography (PET) or lowered levels of $A\beta_{42}$ in the cerebrospinal fluid (CSF),[8] there has been an increasing concern about the labor and cost of eligibility screening for amyloid status.[9,10] Importantly, individuals with preclinical AD cannot be identified through memory clinics because of the lack of symptoms and motivation to visit hospitals. Thus, there has been a compelling need for a sustainable system that facilitates efficient recruitment of eligible asymptomatic amyloid-positive participants, who are willing to be enrolled in AD clinical trials.[11]

Recently, there have been a couple of worldwide movements to build cohorts of preclinical AD individuals who are eligible for clinical trials of disease-modifying therapy (DMTs) for AD. Among these projects, the Trial-Ready Cohort for Preclinical/Prodromal Alzheimer's Disease (TRC-PAD) in the United States has applied an innovative, two-layered structure consisted of a web-based feeder registry (APT webstudy), from which eligible individuals are referred for in-person, clinical, PET, and biomarker assessments, to systematically screen participants who have high risks for elevated brain amyloid deposition and construct a trial-ready cohort (TRC-PAD) for prevention trials.[12–14] In Japan, we have started a close collaboration with the TRC-PAD team and adopted the basic framework of the webstudy and TRC-PAD. In the Japanese Trial-Ready Cohort (J-TRC) for preclinical and prodromal AD launched in October 2019, cognitive normal elderly volunteers are at first invited to register to the J-TRC webstudy at home by themselves and provide basic demographics, to be monitored for their web-based cognitive performance every 3 months. Among the J-TRC webstudy population, those who may have a higher probability for brain amyloid deposition are further referred to the in-person, J-TRC on-site study.

As of summer 2020, more than 3000 elderly volunteers have registered in the J-TRC webstudy within the first 9 months since its launch, despite the global impact of COVID-19 outbreak, and ≈50% of the registrants have been repeating the scheduled remote cognitive tests. To recruit eligible individuals for the J-TRC onsite study, we need to identify J-TRC webstudy participants with a higher likelihood to have elevated brain amyloid deposition or a higher risk of cognitive decline (Figure 1), in reference to the risk factors clarified by earlier preclinical AD studies such as having *APOE* ε4 allele(s),[6,15–17] older age,[6,15,17] family history of dementia,[6] worse Preclinical Alzheimer Cognitive Composite (PACC) score at screening,[6] or worse serial change in Cognitive Function Instrument (CFI),[6] along with the results of online cognitive tests. Importantly, establishing machine learning–based algorithms by incorporating these potential risk factors[9,10,14] will greatly help us to determine at which priority we should invite the individual webstudy participants.

At the phase of J-TRC webstudy that is conducted totally online without in-person visits, however, we cannot use some of the important associated factors (e.g., *APOE* genotype or PACC scores), and have to rely solely on the demographic data of age, sex, family history of dementia or AD, education years, current employment status, degree of alcohol intake, degree of regular exercise, as well as the online cognitive scores of CFI and CogState. The limitations in the kinds of available data, as well as the lack of reference amyloid results from the J-TRC onsite study, will inevitably lessen the predictive performance.

RESEARCH IN CONTEXT

1. Systematic review: The PubMed database was searched to identify large-scale web-based clinical study trying to identify those with preclinical Alzheimer's disease (AD) to enroll them into the Trial-Ready Cohort for future AD prevention clinical trials. The present study was identified as the first attempt conducted in the Japanese population using the data from the Japanese Trial-Ready Cohort (J-TRC) webstudy.
2. Interpretation: Our models using Anti-Amyloid Treatment in Asymptomatic Alzheimer's Disease trial data predicted amyloid burden in each J-TRC webstudy participant, and the predicted amyloid accumulation corresponded well with the self-reported prior amyloid status in a small subgroup of J-TRC webstudy participants.
3. Future directions: Our prediction algorithms may be usable for automatic prioritization of candidate participants with higher amyloid risks to be preferentially recruited from the J-TRC webstudy to the in-person study, and maximize the efficiency for the identification of preclinical AD participants.

The Anti-Amyloid Treatment in Asymptomatic Alzheimer's Disease (A4) study[18] is a phase 3 randomized, double-blind, placebo-controlled secondary prevention trial of solanezumab versus placebo in clinically normal older individuals with evidence of elevated $A\beta$ on screening PET being conducted at 67 sites in the United States, Canada, Australia, and Japan. The initial screening data of the A4 study were recently made publicly available for AD studies,[6] which encompass most of the corresponding variables as the J-TRC webstudy, including two Cogstate tests performed 2 to 3 months apart prior to randomization and CFI, as well as the standard uptake value ratio (SUVr) by ^{18}F-florbetapir amyloid PET, providing us with the ideal training reference for developing algorithms to predict the amyloid risks in asymptomatic elderly individuals. In this study, we describe our attempts to establish machine-learning algorithms based on the A4 screening data and apply the predicted SUVr calculated from the variables available in the J-TRC webstudy, to the efficient recruitment of the participants to the J-TRC onsite study by prioritizing the invitation to those who potentially have the highest risks for elevated amyloid in brain.

2 | MATERIALS AND METHODS

2.1 | Data acquisition and preprocessing: the J-TRC webstudy

The following data handling and analyses were performed using R 3.5.1 (R Foundation for Statistical Computing). The J-TRC study for

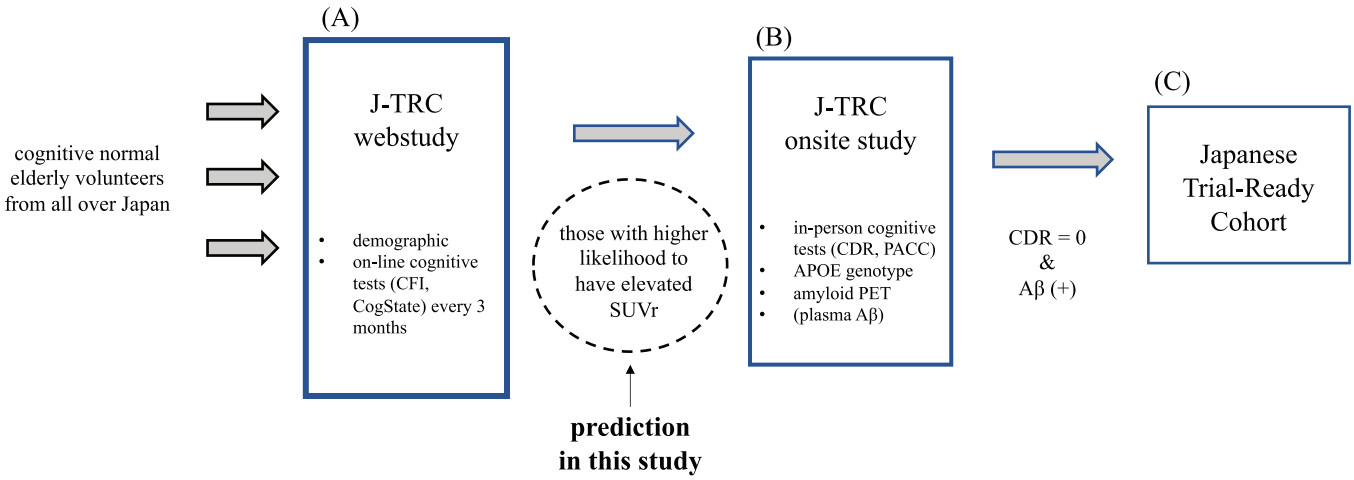

FIGURE 1 Schematic outline of the J-TRC study. Schematic outline of the J-TRC study. Cognitively normal volunteers of 50 to 85 years participate in the J-TRC webstudy by web-based remote cognitive assessment of CFI and CogState every 3 months (A). Those who may have an increased risk of elevated amyloid are further referred to the J-TRC onsite study (B), to conduct detailed assessment including cognitive functions and amyloid status. The J-TRC onsite study eventually aims to build a large (e.g., $n > 300$) Japanese cohort of asymptomatic, amyloid-positive (i.e., preclinical AD) cases being ready for clinical trials of disease modifying drugs in Japan (C). AD, Alzheimer's disease; CDR, Clinical Dementia Rating; CFI, Cognitive Function Instrument; J-TRC, Japanese Trial-Ready Cohort for preclinical and prodromal AD

preclinical and prodromal AD, launched in Japan in 2019 under a research license agreement with the Alzheimer's Therapeutic Research Institute (ATRI), consists of two main study components (Figure 1). First, in the J-TRC webstudy (https://www.j-trc.org/), which is designed based on the Alzheimer's Prevention Trials (APT) Webstudy (https://www.aptwebstudy.org/), in Japanese, cognitively normal elderly volunteer participants at ages 50 to 85 are monitored by web-based remote cognitive assessment of CFI[19] and CogState[20] every 3 months (Figure 1A), and those who are predicted to have an increased risk of elevated brain amyloid or cognitive decline will further be referred to the J-TRC onsite study (Figure 1B), to conduct detailed in-person cognitive assessments, *APOE* genotyping, blood biomarker testing, and determination of brain amyloid status by amyloid PET. The J-TRC onsite study, which is designed based on the TRC-PAD in-person study in the United States, eventually aims to build a large (e.g., $n > 300$) Japanese cohort of preclinical AD individuals being ready for clinical trials in Japan (Figure 1C).

We reviewed the datasets of the J-TRC webstudy participants who registered from October 31, 2019 to June 17, 2020, comprising 4429 registered in total (whether eligible or not). General inclusion criteria in this analysis were defined as follows: participants who completed the registration and demographics input, gave informed consent for study participation, have no prior history of being diagnosed with dementia or AD, and are between 50 and 85 years at the time of registration.

We used the following clinical and cognitive features from the J-TRC webstudy data, which are available in common in the A4 screening and J-TRC webstudy datasets, to include in the predictive models: age, sex (male or female: binary), education years, with/without family (either parents or siblings) history of AD or dementia (binary), online CFI score completed by study participants under an unsupervised condition at screening, and online CogState total score completed up to two times (second at 3 months after initial CogState). We included serial CogState scores because of the potential usefulness of "loss of practice effect" in the cognitive scores of amyloid-positive participants.[16] We converted the final education of each participant to numerical education years as follows: graduated from high school = 12 years, graduated from university/college = 16 years, and graduated from postgraduate school = 18 years. We eventually included $n = 3081$ unique eligible cases from the J-TRC webstudy cohort.

2.2 | Data acquisition and preprocessing: the A4 study

We used the screening datasets of the A4 study obtained from the Laboratory of Neuro Imaging (LONI) (https://ida.loni.usc.edu) in October 2019 with the approval of the data access committee.

As a target to predict, the degree of amyloid accumulation in the A4 study cohort, as represented by the SUVr (value corresponding to the "Composite_Summary" in the "A4_PETSUVR.csv" file) was used: the threshold ≥ 1.15 was used to define elevated brain amyloid[6] (visual evaluation was not taken into account).

We used the clinical and cognitive features at the screening stage of the A4 study as obtained from the J-TRC webstudy data. The CogState score was obtained from the two Computerized Cognitive Composite tests conducted during screening (first time at screening visit 1, the second at screening visit 3 prior to amyloid disclosure). The Z score of each of the following items in CogState normalized within the eligible A4 cases, that is, log response time in *Detection*, log response time in *Identification*, accuracy in *One Card Learning*, and accuracy in *One Back*,[20] was calculated, and the four Z scores were summed to obtain the total CogState score. The intervals between the two CogState

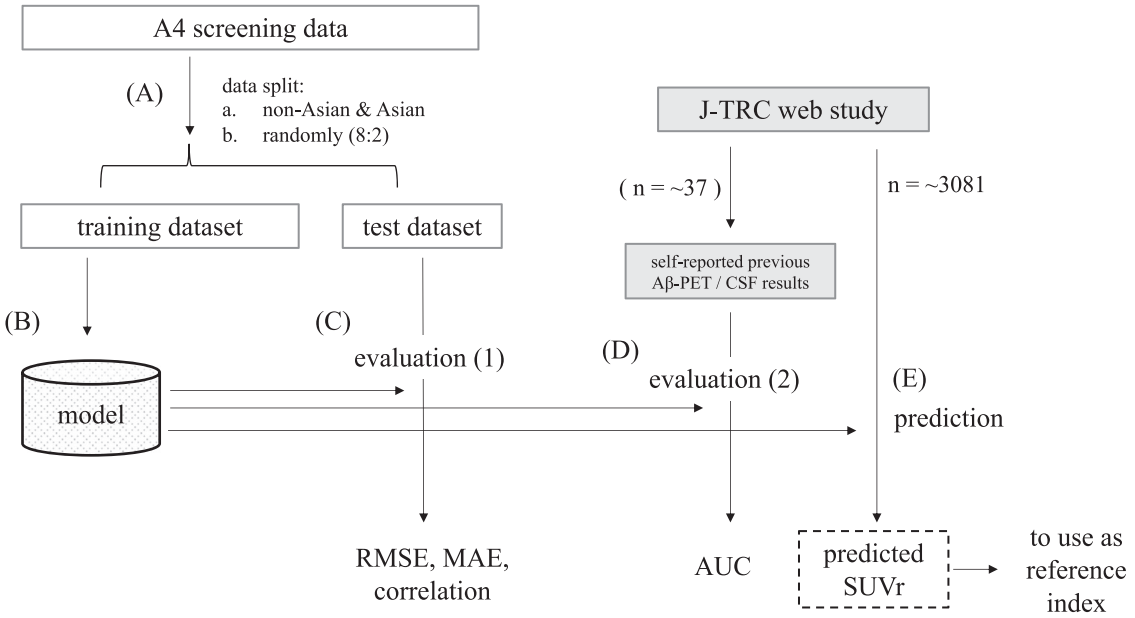

FIGURE 2 Processing workflow of our study. Because the A4 study is mostly composed of participants with non-Asian race, while the participants in J-TRC webstudy are Asian (Japanese), we at first built a model fitted to either the A4 non-Asian training subgroup (A[a]) or the A4 random-split training subgroup (A[b]), then evaluated its performance on the A4 test subgroup (B), and applied the model to the J-TRC webstudy participants (C & D). To evaluate the predictive performance on the A4 test subgroup (B), we calculated MAE and RMSE. The consistency of the informed previous amyloid results with the predicted SUVr (C) was tested by AUC in a subset of J-TRC webstudy participants who reported the results. A4, Anti-Amyloid Treatment in Asymptomatic Alzheimer's Disease; AUC, area under the curve; J-TRC, Japanese Trial-Ready Cohort; MAE, mean absolute error; RMSE, root mean squared error; SUVr, standard uptake value ratio

screenings at visit 1 and visit 3 in the A4 study were estimated to be ≈3 months, which were close to those in the J-TRC webstudy, although those in the A4 study might be slightly shorter by study protocol.

We also used participants' racial data to separate the whole A4 screening data into the non-Asian and Asian subgroup datasets. Samples with missing data in the above modeling features were excluded from the analysis. Eventually, we included $n = 4446$ unique eligible cases from the A4 screening cohort.

2.3 | Model building and evaluation

In this study we intend to build a prediction model for the degree of Aβ deposition by fitting to the A4 screening data as a training reference and to apply the model to the J-TRC webstudy data, to obtain predicted SUVr in each registered J-TRC webstudy participant. Because the A4 study is mostly composed of participants with non-Asian race, whereas the participants in J-TRC webstudy are mostly Japanese, we used two different types of data splitting into the training and test datasets: non-Asian training subgroup and Asian test subgroup (Figure 2A[a]), and randomly split training subgroup and test subgroup (Figure 2A[b]). We at first built a model fitted to the A4 training subgroup (Figure 2A[a, b]), evaluated its performance on the A4 test subgroup (Figure 2B), and then applied the models to the J-TRC webstudy cases (Figure 2C,D).

In the training models, we used several types of algorithms one by one, among the following: generalized linear regression (GLM), penalized GLM (Elastic Net), support vector machine (SVM), random forest (RF), stochastic gradient boosting (GBM), and eXtreme gradient boosting (XGB).[21,22] We used the R package caret[21] for training with 10-fold cross-validation and the automated hyperparameter tuning.

We separately evaluated the performance of several models with different combination of features, as listed in Table 1: model (1), basic demographics (age, sex, education years, family history of dementia or AD) only; model (2), variables in model (1) plus initial CFI score; model (3), variables in model (1) plus initial CogState score; model (4), variables in model (1) plus initial CFI and initial CogState; model (5), variables in model (4) plus the difference in CogState scores between the second and the first assessment (the second score is subtracted by the first score). We have chosen these different combinations because not all J-TRC webstudy participants completed both CFI and CogState, and approximately half of the participants who completed the first-time CogState have not completed the second-time CogState to date. Therefore, model (1) has the largest number of eligible J-TRC cases to impute, while the model (5) has the smallest number of eligible J-TRC cases to impute (Table 2). For evaluating the predictive performance on the A4 test subgroup (Figure 2B), we calculated the mean absolute error (MAE) and root mean squared error (RMSE) by using R package MLmetrics.[23]

The trained model was applied to each case of the J-TRC webstudy cohort, to obtain predicted SUVr. In addition, because a small subgroup of participants ($n = \approx 36$) had prior amyloid PET or CSF Aβ tests elsewhere and have registered the results in the J-TRC webstudy

TABLE 1 Basic demographics of cohorts included in this study

	A4 screening data		J-TRC Webstudy		
	Non-Asian ($n = 4277$)	Asian ($n = 169$)	All ($n = 3081$)	Age > = 65 ($n = 1153$)	Who has previous Aβ -PET/CSF results conducted elsewhere ($n = 37$)
SUVr	1.03 (0.97–1.17)	1.02 (0.96–1.08)	—	—	—
Aβ-positive	1186/4277 (27.7%)????(* SUVr > = 1.15)	29/169 (17.2%)????(* SUVr > = 1.15)	—	—	8/37 (21.6%)????(6/37 in PET, 2/37 in CSF)????(* self-reported previous results)
Age	70.28 (67.49–74.17)	71.56 (68.29–74.32)	61 (55–69)	71 (68– 76)	64.5 (59.25–72)
Sex (female)	2574 4277 (60.2%)	64/169 (37.9%)	1688/3081 (54.8%)	444/1153 (38.5%)	20/37 (54.1%)
Asian race (yes)	0 / 4277 (0%)	169/169 (100%)	3065/3081 (99.5%)	1145/1153 (99.3%)	36/37 (97.3%)
Family history of AD or dementia (yes)	2942/4277 (68.8%)	74/169 (43.8%)	1241/3081 (40.3%)	402/1153 (34.9%)	14/37 (37.8%)
Education (years)	16 (15–18)	16 (16–18)	16 (12–16)	16 (12–16)	16 (16–16)
Retired (yes)	3236/4277 (75.7%)	134/169 (79.3%)	1049/3081 (34%)	755/1153 (65.5%)	19/37 (51.4%)
Having *APOE* ε4 allele(s) (yes)	1506/4277 (35.2%)	37/169 (21.9%)	—	—	—
CFI-self (1st) score	1.5 (0.5–3)	2.5 (1–4.5)	3 (1.5–4.5)	3.5 (2–5)	2.5 (1.125–4.875)
CogState (1st) total score	−0.052 (−1.586–1.525)	−0.22 (−1.944–1.218)	−0.261 (−1.715–1.081)	−0.662 (−2.07–0.736)	0.107 (−1.052–1.129)
CogState (2nd) total score	−0.053 (−1.564–1.537)	−0.229 (−1.775–1.278)	0.41 (−0.961–1.691)	0.174 (−1.298–1.527)	0.231 (−0.331–2.47)

Notes: The number of J-TRC webstudy cases represents that of those registered from October 31, 2019 to July 17, 2020.
Numerical variables are given in median and interquartile range, and categorical variables are shown with frequency and %.
Abbreviations: Aβ, amyloid beta; AD, Alzheimer's disease; APOE, apolipoprotein E; CFI, Cognitive Function Instrument; CSF, cerebrospinal fluid; J-TRC, Japanese Trial-Ready Cohort; PET, positron emission tomography; SUVr, standard uptake value ratio.

TABLE 2 Models with different combinations of variables

	Combination of variables	Number of eligible participants in the J-TRC Webstudy
Model 1	Demographics (= age, sex, education, family history, retired or not, Asian or not*) only	$n = 2511$
Model 2	demographics + CFI	$n = 2498$
Model 3	demographics + CogState (1st)	$n = 1692$
Model 4	demographics + CFI + CogState (1st)	$n = 1692$
Model 5	demographics + CFI + CogState (1^{st}) + difference of CogState (2^{nd} – 1^{st})	$n = 849$

Noets: The number of the J-TRC webstudy participants represents that of those registered between October 31, 2019 and July 17, 2020.
*The variable of race (i.e., Asian or not here) was not included in case of the race-based data splitting (Figure 2A[a]).
Abbreviations: CFI, Cognitive Function Instrument; J-TRC, Japanese Trial-Ready Cohort.

demographics (as binary: amyloid-positive or not), we evaluated whether, in this small subset, the self-reported previous results are consistent with the predicted SUVr (Figure 2C) by area under the curve (AUC), as calculated by the R package MLmetrics. Although some of the participants who reported their own previous amyloid results may have participated in the A4 study screening conducted in Tokyo, Japan, we could not confirm how they actually knew their own amyloid status, due to the webstudy data specifications.

2.4 | Ethics

The J-TRC webstudy has been approved by the University of Tokyo Graduate School of Medicine institutional ethics committee (ID: 2019132NI-[3]), and online informed consent was obtained from each participant upon registration. Using the A4 study data in this research has been approved by the University of Tokyo Graduate School of Medicine institutional ethics committee (ID: 11628-[3]).

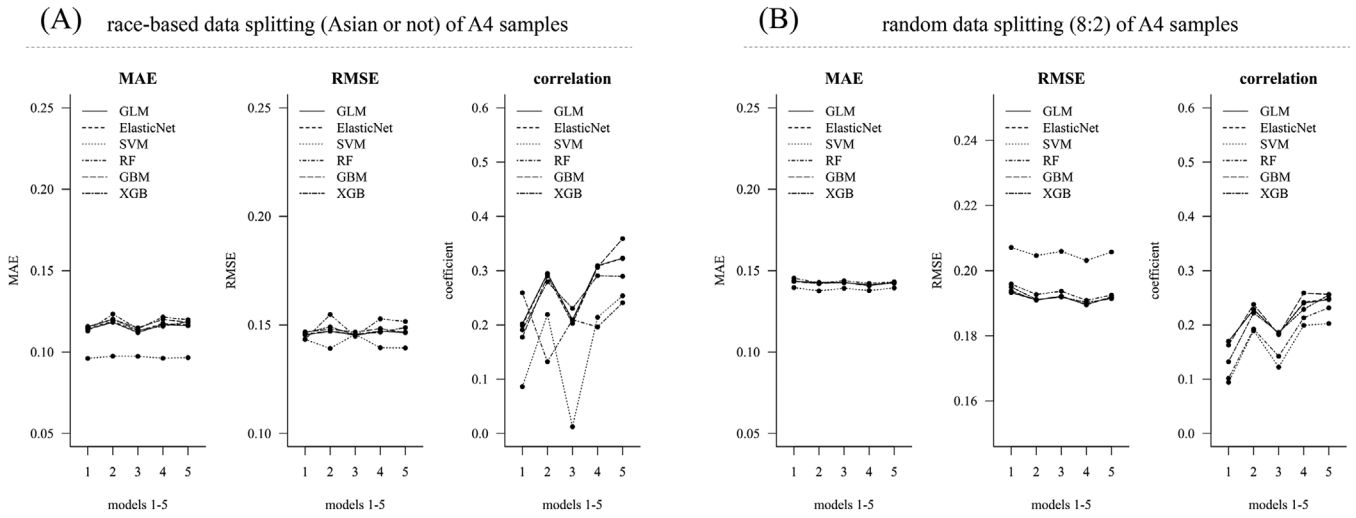

FIGURE 3 Changes in performance by different models/algorithms. The predictive performance (i.e., MAE, RMSE, and correlation coefficient here) of the models on the A4 Asian test subgroup (A) or on the A4 random-split test subgroup (B). Note that each plotted line shows performance change of different algorithms as the model used is varied (model 1–5, in x axis), and when focusing on the correlation coefficient (C), models including CFI (i.e., models 2, 4, and 5) showed better correlation than other models (i.e., models 1 and 3) especially in the algorithms of GLM, ElasticNet, GBM, and XGB (C). A4, Anti-Amyloid Treatment in Asymptomatic Alzheimer's Disease; CFI, Cognitive Function Instrument; ElasticNet, penalized GLM; GBM, stochastic gradient boosting; GLM, generalized linear regression; MAE, mean absolute error; RF, random forest; RMSE, root mean squared error; SVM, support vector machine; XGB, eXtreme gradient boosting

3 | RESULTS

3.1 | Basic demographics

Basic characteristics are shown in Table 1, revealing some differences among the included three cohorts (A4 non-Asian subgroup, A4 Asian subgroup, and the whole J-TRC webstudy). We also listed the two subgroups from the J-TRC webstudy cohort, that is, a subgroup with older participants and a small subgroup of participants reporting their own previous amyloid test results. The proportion of amyloid-positive cases (SUVr ≥ 1.15) was 27.7% (1186/4277) in the non-Asian A4 cohort, 17.2% in the A4 Asian subgroup (29/169 in total). More than half of the A4 Asian subgroup included here should be composed of Japanese cases, as there were 20 amyloid-positive cases among the screened 100 participants from Japan in A4 study (unpublished data). The positive rate in the previous amyloid tests (self-reported) was 21.6% (8/37) in a subset of J-TRC webstudy participants with prior results.

J-TRC webstudy participants were significantly younger (median 61 years old) than the A4 screening participants (median 70.3 years old in the non-Asian subgroup), because of the differences in the inclusion criteria. When comparing the A4 Asian subgroup and the older (≥65 years old) J-TRC webstudy participants, age, sex, education years, CFI, and CogState showed similar distributions.

3.2 | Model performance within the A4 population

As plotted in Figure 3, the predictive performance of the models evaluated on the A4 non-Asian test subgroup was generally limited: the MAE was ≈0.10-0.125 (Figure 3A) and the RMSE was ≈0.15 (Figure 3A), regardless of the type of models (x-axis in Figure 3) or the algorithms (in drawn lines). The correlation coefficients (Figure 3A) in the models including CFI (i.e., models 2, 4, and 5) were higher than those in other models (i.e., models 1 and 3), especially in the algorithms of GLM, ElasticNet, GBM, and XGB. The predictive performance evaluated on the A4 random-split test subgroup also showed similar performance distribution (Figure 3B). Therefore, we mainly used the algorithm of GLM for the following calculations, because it is conventional and simple to calculate.

Figure S1 in supporting information shows an example of the Y-Y plot between the predicted SUVr versus true SUVr in the race-based data splitting (i.e., non-Asian and Asian here; Figure S1A) or in the random-splitting (Figure S1B), showing a significant but low level of association in either cases. In addition, being consistent with such slight differences depending on the type of model, age, CFI scores, and family history were the top three important variables in the A4-fitted, GLM-based model (Figure S2 in supporting information). The variable importance revealed that the CogState score (score on the first time, and the difference in scores between the second and the first) also is valid as a predictive variable, although its significance is lower than that of the above three variables.

3.3 | Predicting SUVr for the J-TRC webstudy participants

We then obtained the predicted SUVr on each of the J-TRC webstudy participants (Figure 2E), and examined the correlation between the predicted SUVr and each important continuous variable. Age, CFI, and CogState score had a moderate level of correlation with the predicted

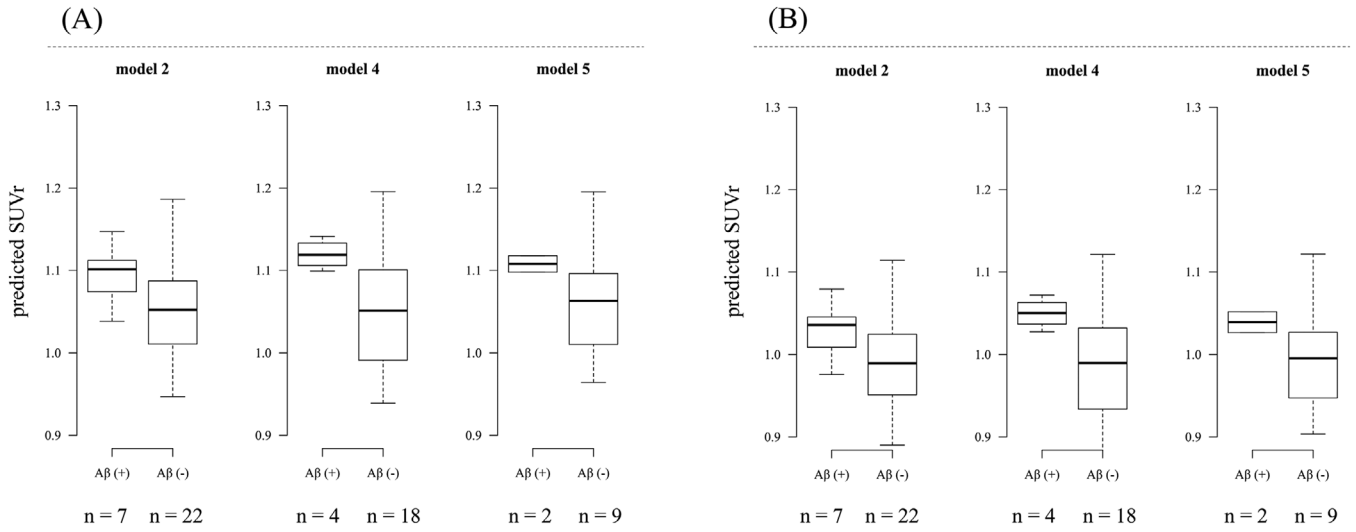

FIGURE 4 Performance on a subgroup of J-TRC webstudy participants who reported previous amyloid test results. The predictive performance (i.e., AUC here) of the models based on the A4 non-Asian training subgroup (A) or on the A4 random-split training subgroup (B), onto a small subgroup of J-TRC webstudy participants. Predicted SUVr had a fair correspondence with the self-reported previous amyloid test results (as binary: positive or not in either amyloid PET or CSF-Aβ42) in a subgroup of J-TRC webstudy participants. A4, Anti-Amyloid Treatment in Asymptomatic Alzheimer's Disease; Aβ, amyloid beta; AUC, area under curve; CSF, cerebrospinal fluid; J-TRC, Japanese Trial-Ready Cohort; PET, positron emission tomography; SUVr, standard uptake value ratio

SUVr (Figure S3A-C in supporting information, for the predicted value based on the A4 non-Asian training dataset).

Next, we evaluated the predictive performance for the self-reported previous amyloid PET or CSF-Aβ results in a subset (Figure 2D) of J-TRC webstudy participants. The predicted value based on the A4 non-Asian training dataset corresponded well with the participants' previous results (i.e., amyloid-positive or not), with AUC = 0.727 (95% confidence interval [CI]: 0.533–0.922) in model (2), AUC = 0.806 (95% CI: 0.619–0.992) in model (4), and AUC = 0.833 (95% CI: 0.578–1.000) in model (5; see Figure 4A). Furthermore, the predicted value based on the random-split training dataset also showed a good correspondence with the participants' previous results: AUC = 0.727 (95% CI: 0.522–0.933) in model (2), AUC = 0.806 (95% CI: 0.619–0.992) in model (4), and AUC = 0.778 (95% CI: 0.456–1.000) in model (5; see Figure 4B).

4 | DISCUSSION

In this study, we built predictive models for the degree of amyloid deposition in amyloid PET by including only variables available in the J-TRC webstudy, based on the A4 screening data as a training reference, thereby aiming to use the predicted SUVr as one of the helpful indexes to automatically extract the list of candidate J-TRC webstudy participants whom we should preferentially invite from the webstudy to the in-person J-TRC onsite study.

Our results showing that the predicted SUVr had a clear correlation with age, family history, and CFI score in the J-TRC webstudy participants (Figure S3) are consistent with those of the earlier studies examining factors associated with positive amyloid in non-demented elderly individuals.[6,15,17] Using the current prediction model as a composite of predictive variables will be superior to using individual variables, as such in automatic ranking the whole webstudy participants as well as in attempting non-linear algorithms. Furthermore, the predicted SUVr showed a fair correspondence with the results of previous amyloid tests in a subset of J-TRC webstudy participants. These results support our current approach to predict the amyloid levels by machine learning algorithm, despite the weakness that we currently do not have amyloid PET data to validate our strategy, and that the J-TRC webstudy cohorts and the reference A4 study might have somewhat different characteristics such as the lower expected prevalence of amyloid positivity in the Japanese population (≈24% in J-ADNI[7] vs. ≈30% in the White population[5,6]).

It should, however, be noted that the model achieved relatively suboptimal performance even upon predicting the SUVr in the A4-Asian subgroup (R = 0.30 at best), which may be inevitable with the small sample size in that subgroup and the current combination of variables lacking *APOE* genotype or PACC scores. Indeed, similarly lower performance is noted by a recent earlier study,[14] in which number needed to screen (NNS) was 2.52 in a Remote model which corresponds to the model (4) in this study, while the NNS improved to 1.78 when incorporating *APOE* genotype and PACC score. To further increase the precision of the predictability of brain amyloid levels, continuous validation and update of the prediction scheme should be essential along with the accumulation of actual data, especially on amyloid PET and *APOE* genotype, in the J-TRC onsite study. Because the composition ratio of *APOE* alleles might be different between non-Asian (mostly White) and Asian populations (Table 1), and the effect of *APOE* may be different depending on the ethnicity, so the A4-based prediction algorithm including *APOE* genotype may show somewhat lower performance when applied to our J-TRC population. Longitudinal changes in the cognitive

outcomes, especially lack of practice or worsening of CogState scores, may also be promising variables that could be tested as soon as we obtain sufficient longitudinal data.

The performance of several models (models 1–5) comprised of different combinations of variables showed slight improvement in the correlation coefficient when incorporating CFI in addition to the basic demographics. We had evaluated these different models separately, because the number of J-TRC participants eligible to impute into these models decreases (Table 2, rightmost column) as we require the participants to complete both CFI and CosState. In line with the larger burden required in completing CogState than CFI, there was approximately one-third (32.3%, 806/2498) of cases who failed to complete CogState among those who completed CFI, while there were few cases (0.5%, 13/2511) who failed to complete CFI among those who registered the demographics (Table 2). Based on the larger importance (Figure S2) and the smaller failure rate of CFI than CogState (Table 2), we suggest that incorporating the CogState (first-time) into a prediction model is not mandatory so as not to lessen the number of recruitable participants, while incorporating CogState would increase the prediction accuracy to a smaller extent.

In the model training, we used different ways to split A4 whole data—race-based splitting (i.e., non-Asian and Asian subgroups; Figure 2A[a]) and random-splitting (Figure 2A[b])—both result in similar distribution in the performance metrics across different models/algorithms (Figure 3). This suggests that the models including CFI (i.e., models 2, 4, and 5 in Table 2) with the GLM algorithm consistently yielded good prediction performance regardless of the variability between the training and test datasets, and these models/algorithm settings might be a rough basis in the future actual training of predictive models within the J-TRC study participants only. Although the prediction performance of the models in the A4 test subgroup was relatively poor ($R \approx 0.2–0.3$), we consider that this does not always matter because the main purpose of our current study was to seek for a good prediction model available for the J-TRC webstudy participants, but not for A4 screening cases.

Our current approach has some limitations. First, the demographic registration to the webstudy is based on self-reporting, so the registered data are not validated, especially on the accuracy of the previous amyloid tests. Second, the age distribution of the A4 screening cohort (65–85 years old as inclusion criteria) is significantly older than that of the J-TRC webstudy (50–85 years old as inclusion criteria), which might lessen the applicability of A4-fitted models to J-TRC cases, especially to younger participants between 50 and 65 years of age. Third, although we excluded those who noted a prior diagnosis of dementia or AD (247 among 3365 [7.4%] of those with eligible age, registration completed, and consent given) from the J-TRC webstudy population, they are not qualified for not having dementia. And fourth, simply prioritizing those with too-high CFI scores (e.g., CFI > 10) for invitation may lead to an increase in the number of individuals with mild cognitive impairment or dementia among the participants invited to J-TRC onsite study; this may facilitate the inclusion of prodromal AD, but might confound the prediction of asymptomatic amyloid-positive (i.e., preclinical AD) participants. We may need to examine whether we can define the eligible range of CFI for more-reliable recruitment of the asymptomatic amyloid-positive participants.

To conclude, we described our current provisional attempts that will automatically extract the list of candidate participants of the J-TRC webstudy to identify who should be preferentially invited to the in-person, J-TRC onsite study, to increase the predictability of amyloid-positive, asymptomatic individuals. To make this invitation process more systematic and efficient, we will continue to update the predicting models along with the progress of the identification of amyloid-positive individuals in the J-TRC onsite study, to confirm and secure the validity of this approach.

ACKNOWLEDGMENT

This study was supported by Japan Agency for Medical Research and Development grants JP17DK0207028 and JP19DK0207048.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to report.

REFERENCES

1. Sperling RA, Aisen PS, Beckett LA, et al. 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*. 2011;7:280-292.
2. Jack CR Jr, Knopman DS, Jagust WJ, et al. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol*. 2013;12:207-216.
3. Donohue MC, Sperling RA, Salmon DP, et al. Australian Imaging, Biomarkers, and Lifestyle Flagship Study of Ageing; Alzheimer's Disease Neuroimaging Initiative; Alzheimer's Disease Cooperative Study. The preclinical Alzheimer cognitive composite: measuring amyloid-related decline. *JAMA Neurol*. 2014;71:961-970.
4. Cummings J. The National Institute on Aging-Alzheimer's Association Framework on Alzheimer's disease: application to clinical trials. *Alzheimers Dement*. 2019;15:172-178.
5. Jansen WJ, Ossenkoppele R, Knol DL, et al. Prevalence of Cerebral Amyloid Pathology in Persons Without Dementia. *JAMA*. 2015;313:1924-1938. Prevalence of cerebral amyloid pathology in persons without dementia: a meta-analysis.
6. Sperling RA, Donohue MC, Raman R, et al. Association of Factors With Elevated Amyloid Burden in Clinically Normal Older Individuals. *JAMA Neurol*. 2020;77:1-11.
7. Iwatsubo T, Iwata A, Suzuki K, et al. Japanese and North American Alzheimer's Disease Neuroimaging Initiative studies: harmonization for international trials. *Alzheimers Dementia*. 2018;14:1077-1087.
8. Jack CR Jr, Bennett DA, Blennow K, et al. A/T/N: an unbiased descriptive classification scheme for Alzheimer disease biomarkers. *Neurology*. 2016;87:539-547.
9. Insel PS, Palmqvist S, Mackin RS, et al. Assessing risk for preclinical β-amyloid pathology with APOE, cognitive, and demographic information. *Alzheimers Dement (Amst)*. 2016;4:76-84.
10. Ansart M, Epelbaum S, Gagliardi G, et al. Alzheimer's Disease Neuroimaging Initiative* and the INSIGHT-preAD study. Reduction of recruitment costs in preclinical AD trials: validation of automatic pre-screening algorithm for brain amyloidosis. *Stat Methods Med Res*. 2020;29:151-164.
11. Sperling R, Cummings J, Donohue M, Aisen P. Global Alzheimer's Platform Trial Ready Cohorts for the Prevention of Alzheimer's Dementia. *J Prev Alzheimers Dis*. 2016;3:185-187.

12. Aisen PS, Sperling RA, Cummings J, et al. The Trial-Ready Cohort for Preclinical/Prodromal Alzheimer's Disease (TRC-PAD) Project: an Overview. *J Prev Alzheimers Dis*. 2020;7(4):208-212.
13. Jimenez-Maggiora GA, Bruschi S, Raman R, et al. TRC-PAD: accelerating Recruitment of AD Clinical Trials through Innovative Information Technology. *J Prev Alzheimers Dis*. 2020;7(4):208-212.
14. Langford O, Raman R, Sperling RA, et al. Predicting Amyloid Burden to Accelerate Recruitment of Secondary Prevention Clinical Trials. *J Prev Alzheimers Dis*. 2020;7(4):213-218.
15. Gottesman RF, Schneider AL, Zhou Y, et al. The ARIC-PET amyloid imaging study: brain amyloid differences by age, race, sex, and APOE. *Neurology*. 2016;87:473-480.
16. Ihara R, Iwata A, Suzuki K, et al. Clinical and cognitive characteristics of preclinical Alzheimer's disease in the Japanese Alzheimer's Disease Neuroimaging Initiative cohort. *Alzheimers Dement (N Y)*. 2018;4:645-651.
17. Vermunt L, Muniz-Terrera G, Ter Meulen L, et al. Prescreening for European Prevention of Alzheimer Dementia (EPAD) trial-ready cohort: impact of AD risk factors and recruitment settings. *Alzheimers Res Ther*. 2020;12:8.
18. Sperling RA, Rentz DM, Johnson KA, et al. The A4 study: stopping AD before symptoms begin?. *Sci Transl Med*. 2014;6:228fs13.
19. Amariglio RE, Donohue MC, Marshall GA, et al. Tracking early decline in cognitive function in older individuals at risk for Alzheimer disease dementia: the Alzheimer's Disease Cooperative Study Cognitive Function Instrument. *JAMA Neurol*. 2015;72:446-454.
20. Mielke MM, Machulda MM, Hagen CE, et al. Performance of the CogState computerized battery in the Mayo Clinic Study on Aging. *Alzheimers Dement*. 2015;11:1367-1376.
21. Kuhn Max. Contributions from Jed Wing, Steve Weston, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, Zachary Mayer, Brenton Kenkel, the R Core Team, Michael Benesty, Reynald Lescarbeau, Andrew Ziem, Luca Scrucca, Yuan Tang, Can Candan and Tyler Hunt. (2018). caret: Classification and Regression Training. R package version 6.0-81. Accessed: January 2019. (https://CRAN.R-project.org/package=caret)
22. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol*. 2019;110:12-22.
23. Yan Y. MLmetrics: Machine Learning Evaluation Metrics. R package version 1.1.1. 2016. (https://CRAN.R-project.org/package=MLmetrics)

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Sato K, Ihara R, Suzuki K, et al. Predicting amyloid risk by machine learning algorithms based on the A4 screen data: Application to the Japanese Trial-Ready Cohort study. *Alzheimer's Dement*. 2021;7:e12135. https://doi.org/10.1002/trc2.12135