


RESEARCH ARTICLE

# External validation of dementia prediction models in Black or African American and White older adults: A longitudinal population-based study in the United States

Klodian Dhana<sup>1,2</sup>  | Lisa L. Barnes<sup>3,4</sup> | Todd Beck<sup>1,2</sup> | Anisa Dhana<sup>1,2</sup> | Xiaoran Liu<sup>1,2</sup> | Pankaja Desai<sup>1,2</sup> | Ted K. S. Ng<sup>1,2</sup> | Denis A. Evans<sup>1,2</sup> | Kumar B. Rajan<sup>1,2</sup>

<sup>1</sup>Rush Institute for Healthy Aging, Rush University Medical Center, Chicago, Illinois, USA

<sup>2</sup>Department of Internal Medicine, Rush University Medical Center, Chicago, Illinois, USA

<sup>3</sup>Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois, USA

<sup>4</sup>Department of Neurology, Rush University Medical Center, Chicago, Illinois, USA

## Correspondence

Klodian Dhana, Rush University Medical Center 1700 W Van Buren, Suite 245; Chicago, IL 60612, USA.

Email: [klodian\\_dhana@rush.edu](mailto:klodian_dhana@rush.edu)

## Funding information

Alzheimer's Association, Grant/Award Number: AARG-22-974154; National Institute on Aging; National Institutes of Health, Grant/Award Numbers: R01AG073627, R01AG058679, UH2AG083289

## Abstract

**INTRODUCTION:** Identifying people at high risk of Alzheimer's disease (AD) dementia allows for timely intervention, which, if successful, will result in preventing or delaying the onset of the disease.

**METHODS:** Utilizing data from the Chicago Health and Aging Project (CHAP;  $n = 2130$ ), we externally evaluated four risk-prediction models for AD dementia, including Cardiovascular Risk Factors, Aging, and Dementia (CAIDE), Australian National University Alzheimer's Disease Risk Index (ANU-ADRI), Brief Dementia Screening Indicator (BDSI), and Dementia Risk Score (DRS), in Black or African American and White adults.

**RESULTS:** BDSI had the highest discriminate abilities for AD dementia (c-statistics of 0.79 in Black and 0.77 in White adults), followed by ANU-ADRI, within the age range and follow-up period of the original development cohort. CAIDE had the lowest discriminating power (c-statistic  $\leq 0.55$ ). With increasing follow-up periods (i.e., 10–15 years), the discrimination abilities for all models declined.

**DISCUSSION:** Because of racial disparities in AD dementia and longer preclinical and prodromal stages of disease development, race-specific models are needed to predict AD risk over 10 years.

## KEYWORDS

Alzheimer's disease, Black or African American, dementia, risk assessment, validation, White

## Highlights

- Utilizing risk-prediction models to identify individuals at higher risk of Alzheimer's disease (AD) dementia could benefit clinicians, patients, and policymakers. Clinicians

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Author(s). *Alzheimer's & Dementia* published by Wiley Periodicals LLC on behalf of Alzheimer's Association.

could enroll high-risk individuals in clinical trials to test new risk-modifiable treatments or initiate lifestyle modifications, which, if successful, would slow cognitive decline and delay the onset of the disease.

- Current risk-prediction models had good discriminative power during the first 6 years of follow-up but decreased with longer follow-up time.
- Acknowledging the longer preclinical phase of AD dementia development and racial differences in dementia risk, there is a need to develop race-specific risk-prediction models that can predict 10 or 20 years of risk for AD and related dementias.

## 1 | INTRODUCTION

Alzheimer's disease (AD) dementia poses a social, medical, and financial challenge for the United States, impacting individuals, caregivers, families, and the health care system.<sup>1-3</sup> As a result, numerous research studies have been conducted to determine risk factors associated with AD dementia.<sup>4-9</sup> In addition, these risk factors for AD have been utilized to create risk-prediction models, also referred to as risk stratification, which aim to identify individuals at higher risk of AD during the early stages of the disease.<sup>10</sup> Early identification of individuals at higher risk of AD dementia through risk stratification is driven by the potential benefit that those individuals might gain from lifestyle interventions. In fact, utilizing risk-stratification models to screen individuals for enrollment in clinical trials has proven valuable.<sup>11</sup>

Numerous prediction models have emerged in recent years, predominantly developed using data from White individuals, and only a few of these models have been evaluated outside their original study populations, a process known as external validation.<sup>12</sup> External validation evaluates the extent to which the findings of a study, for example, the ability to predict dementia, can be applied to other settings and populations—prediction models that lack external validation may not be applicable in clinical practice. A comparative analysis from the Rotterdam Study, a population-based study in The Netherlands, focused on four commonly used risk-prediction models and demonstrated substantial variability in their capacity to discriminate and predict dementia among older White adults.<sup>13-17</sup> Ultimately, the study suggested the need to develop new risk-prediction models.<sup>17</sup> However, the extent to which the validity of these risk-prediction models may extend to older Black or African American adults, a population with a twofold higher likelihood of developing AD dementia than their counterparts, remains unknown.<sup>18</sup> This study aims to externally evaluate these four commonly used risk predictions for AD dementia in both Black and White older adults residing in the South Side neighborhoods of Chicago.

## 2 | METHODS

### 2.1 | Description of the prediction models included in the analysis

#### 2.1.1 | Cardiovascular Risk Factors, Aging, and Dementia (CAIDE)

The CAIDE risk-prediction model was developed using population-based random samples of the North Karelia Project and the FINMONICA study in Finland, which enrolled individuals 39 to 65 years of age. The scoring included age, sex, education, body mass index (BMI), systolic blood pressure, total cholesterol, and physical activity level.<sup>13</sup> An additional risk score was developed, including the genetic information on whether an individual was an apolipoprotein (APOE)  $\epsilon$ 4 carrier. Diagnosis of dementia was based on criteria of the Fourth Edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). Study participants were followed for 20 years and evaluated for the incidence of dementia.<sup>13</sup>

#### 2.1.2 | Australian National University Alzheimer's Disease Risk Index (ANU-ADRI)

The ANU-ADRI was developed using self-reported risk factors for Alzheimer's and dementia. These risk factors were identified from a systematic literature review and included age, sex, education, diabetes, depressive symptoms, traumatic brain injury, smoking, alcohol consumption, social engagement, physical activity, cognitively stimulating activities, fish intake, and pesticide exposure.<sup>19</sup> For people younger than 60 years, the prediction model also includes BMI and cholesterol. The model was evaluated in three cohorts, including the Rush Memory and Aging Project and Cardiovascular Health Cognition Study in the United States and the Kungsholmen Project in Sweden. Diagnosis of dementia differed between studies, with the Rush Memory

and Aging Project based on the National Institute of Neurological and Communicative Diseases and Stroke–Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria. In the Cardiovascular Health Cognition Study, the diagnosis of dementia was based on a progressive or static cognitive deficit of sufficient severity to affect the participant's activities of daily living and impairment in at least two cognitive domains. In the Kungsholmen Project, physicians diagnose dementia according to the DSM-III-R criteria using a validated three-step diagnostic procedure. In the three cohorts study, participants were followed for an average of 6 years to document the incidence of dementia.<sup>19</sup>

### 2.1.3 | Brief Dementia Screening Indicator (BDSI)

The BDSI was developed among individuals 65–79 years of age across four community- or population-based cohort studies in the United States, including the Cardiovascular Health Study, Framingham Heart Study, Health and Retirement Study, and Sacramento Area Latino Study on Aging.<sup>14</sup> The model included age, education, BMI, presence of diabetes, history of stroke, assistance needed with finances or medications, and depressive symptoms. The diagnosis of dementia differed between studies, and the investigators' consensus on the incidence of dementia across studies was based on cognitive impairment in at least two domains, reflecting a decline from prior levels, that affected daily function.<sup>14</sup> Study participants were followed for 6 years for the incidence of dementia.

### 2.1.4 | Dementia Risk Score (DRS)

The DRS utilized data from primary care in the United Kingdom (UK) and the Health Improvement Network (THIN).<sup>15</sup> The THIN database is representative of the UK population, and individuals 60 to 95 years of age were included in the development of the risk-prediction model. The risk-prediction models were developed separately for individuals 60 to 79 years of age and 80 to 95 years of age. The model included age; sex; social deprivation; calendar year at baseline; smoking status; BMI; total cholesterol/high-density lipoprotein cholesterol ratio; systolic blood pressure; history of heavy alcohol use; history of diabetes, coronary heart disease, stroke, or atrial fibrillation, depression, anxiety; and use of anti-hypertensive drugs, statins, or aspirin and other non-steroidal anti-inflammatory drugs (NSAIDs). The dementia diagnoses included AD, vascular dementia, and unspecified or mixed dementia during 5 years of follow-up.<sup>15</sup>

## 2.2 | Study population for external validation of risk-prediction scores

All four risk-prediction scores were validated within the Chicago Health and Aging Project (CHAP), a population-based cohort study focused on risk factors for AD.<sup>20</sup> Established in 1993, CHAP enrolled 10,802 participants 65 years of age or older residing in the South side

### RESEARCH IN CONTEXT

- 1. Systematic review:** Identifying people at high risk of Alzheimer's dementia (AD) dementia has many potential benefits for clinicians, patients, and policymakers. For instance, clinicians could enroll high-risk individuals in clinical trials or initiate risk-modifiable treatment, which, if successful, would slow cognitive decline and delay the onset of the disease.
- 2. Interpretation:** We utilized demographic, genetic, lifestyle, and clinical data from a population-based study, the Chicago Health and Aging Project (CHAP), to externally validate four risk-prediction models for AD dementia, including Australian National University Alzheimer's Disease Risk Index (ANU-ADRI), Brief Dementia Screening Indicator (BDSI), Cardiovascular Risk Factors, Aging, and Dementia (CAIDE), and Dementia Risk Score (DRS), in Black or African American and White older adults. Of all risk-prediction models, BDSI, followed by ANU-ADRI, had good discriminative power during 6 years but decreased with longer follow-up time (i.e., 10–15 years).
- 3. Future directions:** Acknowledging the longer preclinical and prodromal phases of disease development and racial differences in AD dementia risk, there is a need to develop race-specific risk-prediction models that can predict 10 or 20 years of risk for AD dementia.

of Chicago until 2012. Data acquisition included in-home assessments through structured interviewer-administered questionnaires covering various social, lifestyle, and clinical phenotypes.

Of the 10,802 individuals enrolled in the study (Figure S1), a stratified random sample of 2794 participants was further evaluated for clinical diagnosis of AD by detailed clinical examination. A total of 2130 participants were followed for incident AD and comprised the study sample.<sup>4,21</sup> To account for probability of selection for clinical diagnosis, all analyses were sample weight adjusted for stratified random sampling.

## 2.3 | Assessment of predictor variables in CHAP

Race, sex, and education (years of formal schooling) were determined using the 1990 US census questions. Apolipoprotein E (APOE)  $\epsilon$ 4 genotyping was measured on the single nucleotide polymorphisms of the rs7412 and rs429358 by the Broad Institute Center for Genotyping (Cambridge, Massachusetts) using the hME Sequenom MassARRAY platform.<sup>22</sup> BMI was computed by dividing participants' weight (kg) by height (m<sup>2</sup>). Physical activity was assessed through the 1985 US Health Interview Survey, where participants reported time spent in six

activities, including walking for exercise, gardening or yard work, calisthenics or general exercise, bicycle riding, and swimming.<sup>23</sup> Smoking status was self-reported, where participants were categorized as never smokers, current smokers, and former smokers.<sup>24</sup> Alcohol consumption was self-reported in units of alcohol consumed per day. Cognitive activities were assessed using a structured questionnaire that measured participation in the following cognitively stimulating activities during the past year, including reading newspapers, magazines, and books; visiting a museum; playing games like cards, checkers, crosswords, or puzzles; and listening to the radio or watching television. The variable ranges from 1 to 5 and is calculated by averaging the individual item scores.<sup>25</sup> Fish intake was obtained by a validated food frequency questionnaire estimating how often, on average, a participant had consumed a specified amount of foods during the previous year.<sup>26</sup>

Social engagement was determined by responses to questions on attending religious services, visiting a museum, participating in activities or groups outside the home, or working part time or full time.<sup>27</sup> The neighborhood deprivation index was created for each US Census tract based on income, wealth, education, employment/occupation, and housing conditions.<sup>28</sup>

A trained or certified research assistant measured systolic blood pressure. Dyslipidemia, characterized by elevated cholesterol levels, was assessed by examining whether the study participant had been prescribed statins.<sup>29</sup> History of heart disease, stroke, diabetes, and atrial fibrillation was determined by self-report questions from the Established Populations for the Epidemiologic Study of the Elderly. A 10-item depressive symptoms scale, the Center for Epidemiologic Studies Depression was used to quantify symptoms of depression.<sup>30</sup> Traumatic brain injury with loss of consciousness was self-reported. Information on whether a study participant needed help with taking medications and managing money was part of the instrumental activities of daily living assessment.

Information on medications, including the use of anti-hypertensives, anxiolytics, aspirin and other NSAIDs, was obtained during home visits through a direct inspection of the medication that study participants were using.

CHAP did not collect information on pesticide exposure, since the population was urban and had little risk of exposure.

The variables described were assessed at baseline, except for the food frequency questionnaire, which was administered a few years after baseline. The food frequency questionnaire questions study participants about how much food they consumed on average the previous year, and often dietary patterns remain unchanged in older adults. We utilized the food frequency questionnaire to determine fish intake in our study participants.

## 2.4 | Clinical evaluation and diagnosis of Alzheimer's disease dementia

The clinical diagnosis of AD dementia was established following a uniform clinical evaluation, as described previously.<sup>4,20,21</sup> In brief, an exper-

rienced clinician utilized information derived from a comprehensive neurological examination, medical background, and cognitive performance assessments and, with the aid of an algorithmically generated cognitive impairment rating, determined the diagnosis of AD dementia according to the criteria outlined by the collaborative working group of the NINCDS-ADRDA.<sup>31</sup>

## 2.5 | Statistical analysis

We computed the risk scores of every participant of the CHAP study based on published points from each prediction model of the original study. For prediction models that provided the regression coefficients and survival function or baseline risk (i.e., CAIDE<sup>13</sup> and DRS<sup>15</sup>), we created the linear predictor and estimated the risk of AD dementia for each participant of the CHAP study (see [Appendix in the Supplementary Material](#)).

The study evaluated the predictive performance of four risk-prediction models (CAIDE,<sup>13</sup> BDSI,<sup>14</sup> ANU-ADRI,<sup>19</sup> and DSR<sup>15</sup>) in the CHAP study by assessing their discrimination and calibration. Discrimination is the ability of a prediction model to distinguish between an individual who will develop an event (e.g., AD) and one who will not. We quantified discrimination for four models by calculating the c-statistic. Calibration, on the other hand, assesses the agreement between predicted probabilities of disease from the risk-prediction models and the actual incidence of events in the CHAP population. The calibration of the risk-prediction model (i.e., CAIDE<sup>13</sup> and DRS<sup>15</sup>) was evaluated by comparing the predicted risks from the prediction model with the observed risks (cumulative incidence of the event). The mean predicted probability was then plotted against the observed dementia incidence in each quintile of the predicted risk for Black or African American older adults and their White counterparts. In addition, we performed recalibration of the original logistic (i.e., CAIDE<sup>13</sup>) and Cox regression (i.e., DRS<sup>15</sup>) models, updating the intercept and baseline survival function to align with the characteristics of the CHAP study population.

The comparative approach of risk-prediction models was based on the analysis by Licher et al. from the Rotterdam Study.<sup>17</sup> By following a similar approach, we will be able to compare findings from two population-based studies with different demographic characteristics: the Rotterdam Study in The Netherlands, which consists almost entirely of White adults, and the CHAP study, a biracial cohort in Chicago, Illinois, USA. The approach consisted of three primary analyses. First, we calculated risk scores for CHAP participants within the age ranges specified by the original cohorts used to develop the risk-prediction models. We modified the follow-up time in CHAP for incident AD dementia to align with the original risk-prediction cohorts. Second, we validated the risk-prediction model using our entire study population, considering the full age range and follow-up duration (5, 10, and 15 years) of CHAP participants. Third, given the significant impact of age on AD dementia risk, we assessed the c-statistic of the model using two additional models: one considering age alone and another including all risk factors except age. Often researchers focus on incorporating new biomarkers to improve the discriminative power

(i.e., *c*-statistic) of the model without adequately addressing the influence of age alone; therefore, we believe the evaluation of the prediction ability of a model with and without age is justified. In computing risk scores for the entire population, we made adjustments to the scoring for BDSI,<sup>14</sup> which was designed originally for individuals 65 to 79 years of age. To account for this, we allocated an additional point for each year of age beyond the specified range.<sup>17</sup>

Missing data for most of the variables included in the predictions scores were very low (<2% missing), with the highest missing information on fish intake (9.5%), activities of daily living questionnaire (8.9%), atrial fibrillation (8.3%), BMI (6.6%), and APOE ε4 status (3.2%). These missing values were imputed using multivariate imputations by chained equations (MICE) with the mice package, an approach we

have used in previous studies.<sup>23</sup> We used a single imputation method but maximized the accuracy with a wide range of predictors (i.e., age, sex, race, antidepressant or aspirin use, and prevalence of stroke) fitted in all four risk-prediction models. Analyses were performed using R statistical computing, version 4.3.2 (R Foundation for Statistical Computing, Vienna, Austria).

### 3 | RESULTS

Table 1 shows the baseline characteristics of Black or African American and White older adults in the study sample. The mean age of Black adults was 71.7 years (SD 5.1), and of White individuals, 74.8

**TABLE 1** Baseline characteristics of the study population at the baseline.

	Black or African American	White
<i>N</i>	1159	971
<b>Demographic</b>		
Age at baseline, years, mean (SD)	71.7 (5.1)	74.8 (6.3)
Sex, male, <i>n</i> (%)	425 (36.7)	369 (38.0)
Education, years, mean (SD)	11.9 (3.1)	14.2 (3.1)
<b>Genetic</b>		
APOE ε4 carrier, <i>n</i> (%)	414 (35.7)	260 (26.8)
<b>Lifestyle</b>		
Body mass index, kg/m <sup>2</sup> , mean (SD)	29.0 (5.9)	26.3 (4.8)
Physical activity, h/week, median [IQR]	1.0 [0.0, 3.5]	2.5 [0.5, 5.6]
Current smoker, <i>n</i> (%)	167 (14.4)	90 (9.3)
Alcohol use, g/day, median [IQR]	0.0 [0.0, 0.0]	1.5 [0.0, 10.8]
Cognitive activity, comp. score, mean (SD)	3.1 (0.6)	3.5 (0.5)
Fish intake, serving/week, mean (SD)	1.4 (1.2)	1.2 (1.1)
<b>Social</b>		
Social engagement, comp. score, median [IQR]	4.0 [2.0, 5.0]	4.0 [3.0, 5.0]
Neighborhood socioeconomic status, z-score, median [IQR]	−3.2 [−3.9, −1.2]	4.2 [2.6, 6.4]
<b>Clinical</b>		
Systolic blood pressure, mmHg, mean (SD)	138.5 (19.3)	138.4 (19.0)
Diabetes, <i>n</i> (%)	239 (20.6)	112 (11.5)
Stroke, <i>n</i> (%)	85 (7.3)	59 (6.1)
Atrial fibrillation, <i>n</i> (%)	17 (1.5)	78 (8.0)
Depression, <i>n</i> (%)	23 (2.0)	39 (4.0)
Traumatic brain injury with loss of consciousness, <i>n</i> (%)	39 (3.4)	67 (6.9)
Needs help, money/medications, <i>n</i> (%)	167 (14.4)	147 (15.1)
<b>Medication</b>		
Anti-hypertensive use, <i>n</i> (%)	720 (62.1)	494 (50.9)
Statin use, <i>n</i> (%)	122 (10.5)	93 (9.6)
Anxiolytics use, <i>n</i> (%)	25 (2.2)	32 (3.3)
NSAID use, <i>n</i> (%)	216 (18.6)	160 (16.5)
Aspirin use, <i>n</i> (%)	264 (22.8)	341 (35.1)

Note: Depression is defined by the Center for Epidemiologic Studies Depression (CESD)-10 or taking anti-depressive medication.

**TABLE 2** Discriminative ability for predictive models for dementia in the original cohort and Chicago Health and Aging Project (CHAP).

Model	Age range, y	Follow-up, y	c-statistic (95% CI)		
			Original cohorts	CHAP	
				Black or African American	White
CAIDE <sup>a</sup>	39–64	20	0.77 (0.71–0.83)	0.55 (0.51–0.58)	0.53 (0.48–0.57)
+APOE $\epsilon$ 4			0.78 (0.72–0.84)	0.57 (0.54–0.61)	0.59 (0.54–0.63)
ANU-ADRI <sup>b</sup>	55–100	6	0.64 to 0.74	0.78 (0.74–0.83)	0.75 (0.7–0.81)
BDSI <sup>c</sup>	65–79	6	0.68 to 0.78	0.79 (0.74–0.84)	0.77 (0.7–0.85)
DRS <sup>d</sup>	60–79	5	0.84 (0.81–0.87)	0.75 (0.69–0.81)	0.66 (0.58–0.75)
	80–95	5	0.56 (0.55–0.58)	0.61 (0.48–0.73)	0.55 (0.44–0.65)

Notes: Models are ordered in ascending order by publication date. The c-statistic of the original cohorts refers to the data on which the risk-prediction models were first developed or tested, and these estimates were obtained from their respective publications. Studies developing ANU-ADRI and BDSI prediction models utilized several study populations and calculated c-statistic estimates for each study. Specifically, ANU-ADRI was evaluated in three cohorts and provided three c-statistic estimates. BDSI was evaluated in four cohort studies and provided four c-statistic estimates. To simplify the results for comparison with CHAP, we showed only the range for these c-statistic estimates. The 95% CIs of c-statistic estimates of ANU-ADRI and BDSI are shown in the original papers cited above.

Abbreviations: ANU-ADRI, Australian National University Alzheimer's Disease Risk Index; BDSI, Brief Dementia Screening Indicator; CAIDE, Cardiovascular Risk Factors, Aging, and Dementia; DRS, Dementia Risk Score.

<sup>a</sup>Lancet Neurol. 2006 Sep;5(9):735–41. [https://doi.org/10.1016/S1474-4422\(06\)70537-3](https://doi.org/10.1016/S1474-4422(06)70537-3).

<sup>b</sup>PLoS One. 2014 Jan 23;9(1):e86141. <https://doi.org/10.1371/journal.pone.0086141>.

<sup>c</sup>Alzheimers Dement. 2014 Nov;10(6):656–665.e1. <https://doi.org/10.1016/j.jalz.2013.11.006>.

<sup>d</sup>BMC Med. 2016 Jan 21;14:6. <https://doi.org/10.1186/s12916-016-0549-y>.

(SD 6.3). About 63% of the sample were women, similarly distributed among Black and White adults. APOE  $\epsilon$ 4 allele carriership was higher in Black than in White adults: 35.7% and 26.8%, respectively. Compared to White individuals, Black individuals, on average, had higher BMI, lower levels of physical and cognitive activity, and were more likely to be smokers.

Table 2 shows the discriminative ability, as measured by the c-statistic, of the risk-prediction model in the CHAP study, utilizing the age range and follow-up period specific to the original development cohort. Compared to the original cohorts where the risk-prediction scores were first developed or tested, in the CHAP study, CAIDE had the lower c-statistic and BDSI the highest discrimination power. In CHAP, the c-statistics for CAIDE were 0.55 (95% confidence interval [CI] 0.51–0.58) and 0.53 (95% CI 0.48–0.57), respectively, in Black and White individuals. Additional genetic information, the APOE  $\epsilon$ 4, slightly improved the c-statistic in White adults (c-statistic 0.59, 95% CI 0.54–0.63). For BDSI, c-statistics in Black adults were 0.79 (95% CI 0.74–0.84) and in White individuals were 0.77 (95% CI 0.7–0.85), whereas in the original cohorts, c-statistics ranged from 0.68 to 0.78. ANU-ADRI performed slightly better in CHAP than in the original cohorts. In CHAP, the c-statistics for ANU-ADRI were 0.78 (95% CI 0.74–0.83) in Black adults and 0.75 (95% CI 0.7–0.81) in White adults. DRS performed better in the original study population than in the CHAP study.

Table 3 shows the discriminative ability of the risk-prediction models for AD dementia in the entire study sample across 5, 10, and 15 years of follow-up, as well as in the model considering age alone and another model including all risk factors except age. In Black or African American and White older adults, the highest c-statistics were during

5 years of follow-up and decreased with increasing follow-up periods, that is, 10 and 15 years. At 5 years of follow-up for incident AD dementia, ANU-ADRI had the highest discriminative ability with a c-statistics of 0.86 (95% CI 0.81–0.91), which was similar in older Black and White adults. At 10 and 15 years of follow-up, BDSI had the highest c-statistics in both groups. For Black adults, the c-statistics were 0.74 (95% CI 0.71–0.77) at 10 years of follow-up and 0.73 (95% CI 0.70–0.76) at 15 years. In White adults, the c-statistics for BDSI were 0.79 (95% CI 0.75–0.82) and 0.78 (95% CI 0.74–0.81) for 10 and 15 years, respectively. Based on the age variable alone, c-statistics showed lower but comparable discriminative ability for AD dementia in the full model, and the attenuation was slightly higher in Black adults than in White adults. The model without age but with all other risk factors showed the lowest discriminative power across follow-up periods in both groups.

Figure 1 shows the calibration plots for CAIDE and DRS in the CHAP study sample using the original (i.e., development) model age range and follow-up period. The 18-year estimates based on the CAIDE model and the 5-year risk estimate based on the DRS model were, on average, poorly calibrated, underestimating the risk of AD dementia in Black and White adults. Based on the CAIDE model, the average predicted risks were 3.1% for Black adults and 2.0% for their White counterparts, compared to the observed weighted cumulative AD dementia incidence of 14.2% for Black and 9.4% for White adults. Based on the DRS model, the average predicted risk during 5-year follow-up was 0.9% in Black adults and 1.5% in White adults compared to the observed weighted cumulative AD dementia incidence of 4.3% in Black individuals and 4.1% in White individuals. Recalibration of the prediction models improved the calibration, although it underestimated the risk, especially for Black individuals.

**TABLE 3** Discriminative ability for predictive models for dementia in Chicago Health and Aging Project across different follow-up periods.

	c-statistics and 95% CI at different follow-up periods					
	Black or African American			White		
	5 years	10 years	15 years	5 years	10 years	15 years
<b>CAIDE</b>						
CAIDE	0.58 (0.51–0.65)	0.56 (0.52–0.61)	0.56 (0.52–0.59)	0.55 (0.48–0.63)	0.54 (0.48–0.59)	0.53 (0.48–0.57)
+APOE ε4	0.58 (0.52–0.65)	0.58 (0.54–0.63)	0.58 (0.54–0.62)	0.61 (0.53–0.69)	0.6 (0.55–0.65)	0.59 (0.54–0.64)
Age only	NA	NA	NA	NA	NA	NA
Without age	0.58 (0.51–0.65)	0.56 (0.52–0.61)	0.56 (0.52–0.59)	0.55 (0.48–0.63)	0.54 (0.48–0.59)	0.53 (0.48–0.57)
<b>ANU-ADRI</b>						
ANU-ADRI	0.86 (0.81–0.91)	0.71 (0.68–0.75)	0.7 (0.67–0.74)	0.86 (0.81–0.91)	0.74 (0.7–0.79)	0.74 (0.69–0.78)
Age only	0.78 (0.7–0.86)	0.68 (0.64–0.72)	0.67 (0.64–0.71)	0.83 (0.77–0.9)	0.73 (0.69–0.77)	0.72 (0.68–0.76)
Without age	0.78 (0.72–0.84)	0.62 (0.58–0.66)	0.61 (0.57–0.65)	0.69 (0.59–0.79)	0.63 (0.58–0.68)	0.62 (0.58–0.67)
<b>BDSI</b>						
BDSI	0.81 (0.76–0.85)	0.74 (0.71–0.77)	0.73 (0.7–0.76)	0.79 (0.74–0.85)	0.79 (0.75–0.82)	0.78 (0.74–0.81)
Age only	0.74 (0.69–0.8)	0.7 (0.67–0.74)	0.7 (0.66–0.74)	0.73 (0.67–0.79)	0.74 (0.7–0.78)	0.73 (0.69–0.77)
Without age	0.75 (0.69–0.8)	0.68 (0.64–0.72)	0.66 (0.63–0.7)	0.73 (0.67–0.8)	0.71 (0.66–0.76)	0.7 (0.66–0.75)
<b>DRS</b>						
DRS	0.77 (0.72–0.82)	0.72 (0.69–0.76)	0.71 (0.68–0.75)	0.71 (0.65–0.76)	0.72 (0.68–0.76)	0.71 (0.67–0.75)
Age only	0.74 (0.69–0.8)	0.71 (0.67–0.74)	0.7 (0.67–0.74)	0.73 (0.67–0.79)	0.74 (0.7–0.78)	0.73 (0.69–0.77)
Without age	0.7 (0.64–0.77)	0.66 (0.61–0.7)	0.64 (0.6–0.68)	0.67 (0.6–0.75)	0.68 (0.62–0.73)	0.67 (0.62–0.72)

Note: CAIDE assigns a similar score to all individuals aged 53 years and older, whereas all CHAP participants are 65 years and older; therefore, we cannot estimate the c-statistic for age only.

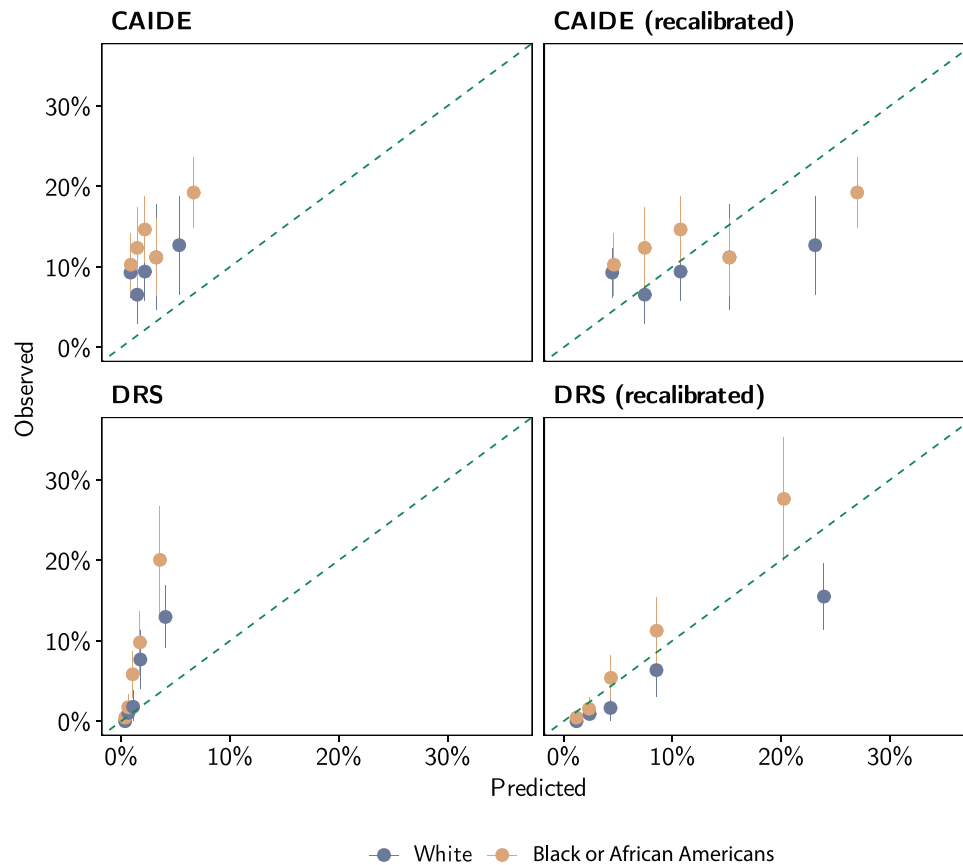
Abbreviations: ANU-ADRI, Australian National University Alzheimer's Disease Risk Index; BDSI, Brief Dementia Screening Indicator; CAIDE, Cardiovascular Risk Factors, Aging, and Dementia; DRS, Dementia Risk Score.

## 4 | DISCUSSION

In this study, we externally validated four commonly used risk-prediction models for AD dementia in Black or African American and White older adults residing in the South Side of Chicago in Illinois. These models incorporated a range of risk factors, including demographics, genetics, lifestyle, clinical factors, social factors, and medication use. ANU-ADRI and BDSI demonstrated good discrimination power (i.e., the ability to differentiate between individuals who will develop dementia and those who will not) over a 6-year follow-up period in both Black and White older adults, consistent with the c-statistics of their original cohorts. CAIDE demonstrated the poorest discrimination ability. The discrimination power decreased with longer follow-up, a trend observed particularly among Black adults. Given the longer preclinical and prodromal phase of AD, over 10 years, and the presence of racial disparities in disease risk, these findings underscore the importance of developing tailored risk-prediction models for Black and White individuals that can predict the 10- and 20-year risk of AD dementia.

There is a growing clinical interest in risk-prediction models, also known as risk stratification, for identifying individuals at increased risk of AD dementia.<sup>12</sup> This interest originates from the potential advantages of early intervention as a good strategy to slow cognitive decline as older adults age and ultimately prevent or postpone

AD dementia onset. For instance, the CAIDE risk-prediction model<sup>13</sup> has proven valuable in screening individuals for enrollment in the FIN-GER trial, which resulted in a significant beneficial intervention effect on overall cognitive performance.<sup>11</sup> However, as new risk-prediction models emerge, clinicians will confront challenging decisions regarding which model to select, especially in the absence of comparative analyses.<sup>12</sup> Moreover, more information is needed on whether current risk-prediction models perform well across different racial/ethnic groups. In the current analysis, we selected the four most common risk-prediction models and conducted a head-to-head comparative analysis evaluating the performance of models separately in American Black and White adults living in the United States. We selected these models based on an analysis from the Rotterdam Study in The Netherlands,<sup>17</sup> which allows us to compare the performance on external validity between CHAP and Rotterdam Study. CHAP and the Rotterdam Study differ in population characteristics, with CHAP being biracial population-based and the Rotterdam Study comprising almost exclusively White adults. In addition, CHAP White adults, on average, are older than Rotterdam Study participants (75 vs 69 years). Nevertheless, both studies highlighted the need for updated models to predict the risk of AD dementia. In the Rotterdam Study, the DRS provided the highest c-statistic, whereas, in the CHAP White adults, BDSI outperformed other models at the 10-year follow-up. DRS was developed using data from the Health Improvement Network, a



**FIGURE 1** Calibration plot of CAIDE and DRS risk-prediction models for Alzheimer's disease in the Chicago Health and Aging Project (CHAP). CAIDE predicts the risk of dementia during 20 years of follow-up, whereas DRS predicts it during 5 years of follow-up. In CHAP, we censored follow-up data according to the prediction models and data availability. Recalibration consists of updating the intercept for the CAIDE model, whereas for the DRS model, it consists of updating the baseline survival function and the mean predictor values. The dashed diagonal line shows the ideal calibration between predicted and observed values. CAIDE, Cardiovascular Risk Factors, Aging, and Dementia; DRS, Dementia Risk Score.

representative of the United Kingdom population. BDSI was developed from four community- or population-based cohort studies in the United States, including the Cardiovascular Health Study and Health and Retirement Study, which comprised 15% and 7% African American individuals, respectively. In addition, the Rotterdam Study showed that risk factors included in the prediction models have limited added value above and beyond age in AD dementia prediction.<sup>17</sup> In the CHAP study, age alone explained the majority of the predictive power; however, additional risk factors added value for specific models.

The CAIDE<sup>13</sup> model had the lowest performance within our study population. A key characteristic of the CAIDE model is its development for midlife adults (aged 40–64 years), suggesting potential limitations when applied to older adults, that is, 65 years and older. In addition, because the CHAP study comprises older adults (mean age 73 years), it is expected that CHAP participants have more comorbidities and other health conditions than participants of the study used to develop the CAIDE score. This likely contributes to the discrepancies between the studies and explains the poor performance of the CAIDE prediction model in our study population. Nevertheless, our results on

CAIDE model performance aligned well with findings from the previous analysis of the risk-prediction model.<sup>17</sup>

Prediction models, such as ANU-ADRI<sup>19</sup> and BDSI,<sup>14</sup> demonstrated reasonably good performance in our population when it was limited to a follow-up period of 6 years based on the development cohort. However, as we extended the follow-up to 10 and 15 years, we noted a decrease in the discriminative ability of these models, especially for BDSI in Black adults. AD is a progressive neurodegenerative disorder characterized by the accumulation of amyloid plaques and tangles, a process that begins 15 to 20 years before clinical manifestation.<sup>32</sup> Given that these prediction models are intended primarily to identify individuals at high risk for early life interventions,<sup>10</sup> as primary prevention through lifestyle changes is most effective when initiated early, it is crucial to develop models specifically for predicting the 10- and 20-year risk of AD. In addition, these risk-prediction models were developed predominantly using data from White adults, and we applied them to Black individuals despite knowing racial disparities in risk factors and AD dementia prevalence.<sup>1</sup> Therefore, race-specific risk-prediction models for AD dementia are very much needed.



Another risk-prediction score, the Rapid Risk Assessment of Dementia (RADaR), was developed within the Rush Memory and Aging Project.<sup>33</sup> RADaR utilizes five self-reported questions, including memory complaints, difficulty with finances, orientation to time and place, and delayed recall of three words. It has shown good discrimination ability for predicting the 3-year incidence of AD dementia in older Black and White individuals. This risk calculation tool is convenient for primary care settings because it requires fewer than 10 min to administer. However, because it can only assess 3-year AD dementia risk, its relevance to primary prevention may be limited, as the process of disease development typically begins at least 10 years before clinical manifestation.<sup>34</sup>

The strengths of our study are a population-based design, long-term follow-up, and accurate diagnosis of AD dementia through structured clinical neurological evaluations with neuropsychological testing. In addition, the availability of a wide range of demographic, genetic, lifestyle, and social risk factors allowed us to compare several prediction models. However, several methodological limitations warrant mentioning. We made several adjustments to risk factors when data were unavailable; for example, we used the status of statin medication to determine whether an individual had dyslipidemia (e.g., higher cholesterol levels). This modification may be acceptable because using a prediction model in clinical practice may require modifications.<sup>17</sup> In addition, we computed the ANU-ADRI score without information on pesticide exposure; regardless, ANU-ADRI performed well in our population. Another limitation is that the results of this study may not be generalized beyond this study population (i.e., CHAP), which consists of Black American and White residents 65 years of age and older from a geographically defined community on the South Side of Chicago. Nevertheless, this is the first study to independently assess the performance of these prediction scores in Black Americans.

In conclusion, most AD dementia prediction models focus on short-term risk estimation and are derived from studies that predominantly include White individuals. Acknowledging the existence of racial disparities in AD dementia and the longer preclinical and prodromal phases of disease development, our study emphasizes the need to identify race-specific predictors for Black and White adults and develop risk-prediction models that can estimate the risk of AD dementia risk over extended periods, such as 10 and 20 years.

## ACKNOWLEDGMENTS

The authors thank research participants and staff of the Chicago Health and Aging Project (CHAP) study. This study was funded by the Alzheimer's Association Research Grant (AARG-22-974154), National Institute on Aging and the National Institutes of Health (R01AG073627, R01AG058679, and UH2AG083289).

## CONFLICT OF INTEREST STATEMENT

Klodian Dhana is funded by the Alzheimer's Association and National Institutes of Health (NIH) research grants and reports no conflicts of interest. Lisa L. Barnes, Todd Beck, Anisa Dhana, Xiaoran Liu, Pankaja Desai, Ted K.S. Ng, Denis A. Evans, and Kumar B. Rajan report no con-

flicts of interest. Author disclosures are available in the [supporting information](#).

## CONSENT STATEMENT

The Rush University Medical Center Institutional Review Board approved population interviews, clinical evaluations, and DNA extraction for this study before participant enrollment. Written informed consent was obtained from all participants in the study.

## ORCID

Klodian Dhana  <https://orcid.org/0000-0002-6397-7009>

## REFERENCES

1. Dhana K, Beck T, Desai P, Wilson RS, Evans DA, Rajan KB. Prevalence of Alzheimer's disease dementia in the 50 US states and 3142 counties: a population estimate using the 2020 bridged-race postcensal from the national center for health statistics. *Alzheimers Dementia*. 2023;19(10):4388-4395. doi: [10.1002/alz.13081](https://doi.org/10.1002/alz.13081)
2. Cummings JL, Goldman DP, Simmons-Stern NR, Ponton E. The costs of developing treatments for Alzheimer's disease: a retrospective exploration. *Alzheimers Dementia*. 2022;18(3):469-477. doi: [10.1002/alz.12450](https://doi.org/10.1002/alz.12450)
3. GBD 2016 Neurology Collaborators. Global, regional, and national burden of neurological disorders, 1990-2016: a systematic analysis for the global burden of disease study 2016. *Lancet Neurol*. 2019;18(5):459-480. doi: [10.1016/S1474-4422\(18\)30499-X](https://doi.org/10.1016/S1474-4422(18)30499-X)
4. Dhana K, Franco OH, Ritz EM, et al. Healthy lifestyle and life expectancy with and without Alzheimer's dementia: population based cohort study. *BMJ*. 2022:e068390. Published online April 13. doi: [10.1136/bmj-2021-068390](https://doi.org/10.1136/bmj-2021-068390)
5. Dhana K, Evans DA, Rajan KB, Bennett DA, Morris MC. Healthy lifestyle and the risk of Alzheimer dementia. *Neurology*. 2020;95(4):e374-e383. doi: [10.1212/wnl.0000000000009816](https://doi.org/10.1212/wnl.0000000000009816)
6. Dhana K, Barnes LL, Liu X, et al. Genetic risk, adherence to a healthy lifestyle, and cognitive decline in African Americans and European Americans. *Alzheimers Dementia*. 2021;18(4):572-580. doi: [10.1002/alz.12435](https://doi.org/10.1002/alz.12435)
7. Lourida I, Hannon E, Littlejohns TJ, et al. Association of lifestyle and genetic risk with incidence of dementia. *JAMA*. 2019;322(5):430. doi: [10.1001/jama.2019.9879](https://doi.org/10.1001/jama.2019.9879)
8. Licher S, Ahmad S, Karamujić-Čomić H, et al. Genetic predisposition, modifiable-risk-factor profile and long-term dementia risk in the general population. *Nat Med*. 2019;25(9):1364-1369. doi: [10.1038/s41591-019-0547-7](https://doi.org/10.1038/s41591-019-0547-7)
9. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet North Am Ed*. 2020;396(10248):413-446. doi: [10.1016/S0140-6736\(20\)30367-6](https://doi.org/10.1016/S0140-6736(20)30367-6)
10. D'Agostino RB, Grundy S, Sullivan LM, Wilson P. Validation of the framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. *JAMA*. 2001;286(2):180-187. doi: [10.1001/jama.286.2.180](https://doi.org/10.1001/jama.286.2.180)
11. Ngandu T, Lehtisalo J, Solomon A, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. *Lancet (London, England)*. 2015;385(9984):2255-2263. doi: [10.1016/S0140-6736\(15\)60461-5](https://doi.org/10.1016/S0140-6736(15)60461-5)
12. Tang EYH, Harrison SL, Errington L, et al. Current developments in dementia risk prediction modelling: an updated systematic review. *PLoS One*. 2015;10(9):e0136181. doi: [10.1371/journal.pone.0136181](https://doi.org/10.1371/journal.pone.0136181)
13. Kivipelto M, Ngandu T, Laatikainen T, Winblad B, Soininen H, Tuomilehto J. Risk score for the prediction of dementia risk in 20 years among middle aged people: a longitudinal, population-based

- study. *Lancet Neurol.* 2006;5(9):735-741. doi: [10.1016/S1474-4422\(06\)70537-3](https://doi.org/10.1016/S1474-4422(06)70537-3)
14. Barnes DE, Beiser AS, Lee A, et al. Development and validation of a brief dementia screening indicator for primary care. *Alzheimers Dementia.* 2014;10(6):656-665.e1. doi: [10.1016/j.jalz.2013.11.006](https://doi.org/10.1016/j.jalz.2013.11.006)
  15. Anstey KJ, Cherbuin N, Herath PM, et al. A self-report risk index to predict occurrence of dementia in three independent cohorts of older adults: the ANU-ADRI. *PLoS One.* 2014;9(1):e86141. doi: [10.1371/journal.pone.0086141](https://doi.org/10.1371/journal.pone.0086141)
  16. Walters K, Hardoon S, Petersen I, et al. Predicting dementia risk in primary care: development and validation of the dementia risk score using routinely collected data. *BMC Med.* 2016;14:6. doi: [10.1186/s12916-016-0549-y](https://doi.org/10.1186/s12916-016-0549-y)
  17. Licher S, Yilmaz P, Leening MJG, et al. External validation of four dementia prediction models for use in the general community-dwelling population: a comparative analysis from the rotterdam study. *Eur J Epidemiol.* 2018;33(7):645-655. doi: [10.1007/s10654-018-0403-y](https://doi.org/10.1007/s10654-018-0403-y)
  18. Barnes LL, Bennett DA. Alzheimer's disease in african americans: risk factors and challenges for the future. *Health Aff (Millwood).* 2014;33(4):580-586. doi: [10.1377/hlthaff.2013.1353](https://doi.org/10.1377/hlthaff.2013.1353)
  19. Anstey KJ, Cherbuin N, Herath PM. Development of a new method for assessing global risk of Alzheimer's disease for use in population health approaches to prevention. *Prev Sci.* 2013;14(4):411-421. doi: [10.1007/s11121-012-0313-2](https://doi.org/10.1007/s11121-012-0313-2)
  20. Evans DA, Bennett DA, Wilson RS, et al. Incidence of alzheimer disease in a biracial urban community: relation to apolipoprotein e allele status. *Arch Neurol.* 2003;60(2):185-189. doi: [10.1001/archneur.60.2.185](https://doi.org/10.1001/archneur.60.2.185)
  21. Rajan KB, Wilson RS, Weuve J, Barnes LL, Evans Denis A. Cognitive impairment 18 years before clinical diagnosis of alzheimer disease dementia. *Neurology.* 2015;85(10):898-904. doi: [10.1212/WNL.0000000000001774](https://doi.org/10.1212/WNL.0000000000001774)
  22. Wenham PR, Price WH, Blandell G. Apolipoprotein e genotyping by one-stage PCR. *Lancet (London, England).* 1991;337(8750):1158-1159. doi: [10.1016/0140-6736\(91\)92823-k](https://doi.org/10.1016/0140-6736(91)92823-k)
  23. Dhana K, Aggarwal NT, Rajan KB, Barnes LL, Evans Denis A, Morris MC. Impact of the apolipoprotein e  $\epsilon 4$  allele on the relationship between healthy lifestyle and cognitive decline: a population-based study. *Am J Epidemiol.* 2021;190(7):1225-1233. doi: [10.1093/aje/kwab033](https://doi.org/10.1093/aje/kwab033)
  24. Aggarwal NT, Bienias JL, Bennett DA, et al. The relation of cigarette smoking to incident Alzheimer's disease in a biracial urban community population. *Neuroepidemiology.* 2006;26(3):140-146. doi: [10.1159/000091654](https://doi.org/10.1159/000091654)
  25. Wilson RS, Scherr PA, Schneider JA, Tang Y, Bennett DA. Relation of cognitive activity to risk of developing alzheimer disease. *Neurology.* 2007;69(20):1911-1920. doi: [10.1212/01.wnl.0000271087.67782.cb](https://doi.org/10.1212/01.wnl.0000271087.67782.cb)
  26. Morris MC, Tangney CC, Wang Y, Sacks FM, Bennett DA, Aggarwal NT. MIND diet associated with reduced incidence of Alzheimer's disease. *Alzheimers Dementia.* 2015;11(9):1007-1014. doi: [10.1016/j.jalz.2014.11.009](https://doi.org/10.1016/j.jalz.2014.11.009)
  27. Barnes LL, de Leon CFM, Bienias JL, Evans DA. A longitudinal study of black-white differences in social resources. *J Gerontol B Psychol Sci Soc Sci.* 2004;59(3):S146-53. doi: [10.1093/geronb/59.3.s146](https://doi.org/10.1093/geronb/59.3.s146)
  28. Aggarwal NT, Everson-Rose SA, Evans DA. Social determinants, race, and brain health outcomes: findings from the chicago health and aging project. *Curr Alzheimer Res.* 2015;12(7):622-631. doi: [10.2174/1567205012666150701102606](https://doi.org/10.2174/1567205012666150701102606)
  29. Rajan KB, Barnes LL, Wilson RS, Weuve J, McAninch Elizabeth A, Evans DA. Blood pressure and risk of incident Alzheimer's disease dementia by antihypertensive medications and APOE  $\epsilon 4$  allele. *Ann Neurol.* 2018;83(5):935-944. doi: [10.1002/ana.25228](https://doi.org/10.1002/ana.25228)
  30. Radloff LS. A self-report depression scale for research in the general population. *Appl Psychol Meas.* 1977;1:385-401.
  31. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of department of health and human services task force on Alzheimer's disease. *Neurology.* 1984;34(7):939-944. doi: [10.1212/wnl.34.7.939](https://doi.org/10.1212/wnl.34.7.939)
  32. Dhana K, Aggarwal P, James BD, et al. Healthy lifestyle and cognition in older adults with common neuropathologies of dementia. *JAMA Neurol.* 2024;81(3):233-239. Published online February 2024. doi: [10.1001/jamaneuro.2023.5491](https://doi.org/10.1001/jamaneuro.2023.5491)
  33. Capuano AW, Shah RC, Blanche P, et al. Derivation and validation of the rapid assessment of dementia risk (RADaR) for older adults. *PLoS One.* 2022;17(3):e0265379. doi: [10.1371/journal.pone.0265379](https://doi.org/10.1371/journal.pone.0265379)
  34. Villemagne VL, Burnham S, Bourgeat P, et al. Amyloid  $\beta$  deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurol.* 2013;12(4):357-367. doi: [10.1016/S1474-4422\(13\)70044-9](https://doi.org/10.1016/S1474-4422(13)70044-9)

## SUPPORTING INFORMATION

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

**How to cite this article:** Dhana K, Barnes LL, Beck T, et al. External validation of dementia prediction models in Black or African American and White older adults: A longitudinal population-based study in the United States. *Alzheimer's Dement.* 2024;20:7913–7922. <https://doi.org/10.1002/alz.14280>