Received: 28 August 2024 | Revised: 31 December 2024 | Accepted: 12 January 2025

DOI: 10.1002/alz.14617

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

RESEARCH ARTICLE

Development and internal validation of a risk prediction model for dementia in a rural older population in China

Keke Liu[1,2] | **Tingting Hou**[1,2] | **Yuqi Li**[1] | **Na Tian**[1,2] | **Yifei Ren**[2] | **Cuicui Liu**[1,2] | **Yi Dong**[1] | **Lin Song**[1,2] | **Shi Tang**[1] | **Lin Cong**[1,2] | **Yongxiang Wang**[1,2,3,4] | **Wei Xiao**[1,2] | **Yifeng Du**[1,2,3] | **Chengxuan Qiu**[1,3,4]

[1]Department of Neurology, Shandong Provincial Hospital affiliated to Shandong First Medical University, Jinan, Shandong, P.R. China

[2]Department of Neurology, Shandong Provincial Hospital, Shandong University, Jinan, Shandong, P.R. China

[3]Institute of Brain Science and Brain-Inspired Research, Shandong First Medical University & Shandong Academy of Medical Sciences, Jinan, Shandong, P.R. China

[4]Aging Research Center, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet-Stockholm University, Stockholm, Sweden

Correspondence
Yifeng Du and Chengxuan Qiu, Department of Neurology, Shandong Provincial Hospital affiliated to Shandong University, No. 324 Jingwuweiqi Road, Jinan, Shandong 250021, China.
Email: du-yifeng@hotmail.com and chengxuan.qiu@ki.se

Funding information
Science and Technology Program for Public Wellbeing of Shandong Province, China, Grant/Award Number: 2013kjhm180405; the National Key R&D Program of China, Grant/Award Numbers: 2017YFC1310100, 2022YFC3501404; the National Nature Science Foundation of China, Grant/Award Numbers: 81861138008, 8191101618; Academic Promotion Program of Shandong First Medical University, Grant/Award Number: 2019QL020; Taishan Scholar Program of Shandong Province, China, Grant/Award Number: tsqn202312347; Natural Science Foundation of Shandong Province, Grant/Award Numbers: ZR2022QH106, ZR2021MH392; Swedish Research Council, Grant/Award Numbers: 2017-05819, 2020-01574; Swedish Research Council for Health, Working Life and Welfare, Grant/Award Number: 2023-01125; Swedish Foundation for International Cooperation in Research and Higher Education, Grant/Award Number: CH2019-8320; Brain Science and

Abstract

INTRODUCTION: We sought to develop a practical tool for predicting dementia risk among rural-dwelling Chinese older adults.

METHODS: This cohort study included 2220 rural older adults (age ≥ 65 years) who were dementia-free at baseline (2014) and were followed in 2018. Dementia was diagnosed following the DSM-IV criteria. The prediction model was constructed using Cox models. We used C-index and calibration plots to assess model performance, and the decision curve analysis (DCA) to assess clinical usefulness.

RESULTS: During the 4-year follow-up, 134 individuals were diagnosed with dementia. We identified age, education, self-rated AD8 score, marital status, and stroke for the prediction model, with the C-index being 0.79 (95% confidence interval = 0.75–0.83) and the corrected C-index for internal validation being 0.79. Calibration plots showed good performance in predicting up to 4-year dementia risk and DCA indicated good clinical usefulness.

DISCUSSION: The 4-year dementia risk can be accurately predicted using five easily available predictors in a rural Chinese older population.

KEYWORDS
calibration plot, C-index, dementia, prediction tool, rural population, validation

Highlights

- We developed and internally validated a practical tool for dementia risk prediction among a rural older population in China.

Keke Liu and Tingting Hou contributed equally to this work.

Brain-Like Intelligence Technology Research Projects of China, Grant/Award Numbers: 2021ZD0201801, 2021ZD0201808

- The prediction tool showed good discrimination and excellent calibration for predicting up to 4-year risk of dementia.
- The prediction tool can be used to identify individuals at a high risk for dementia for early preventive interventions.

1 | BACKGROUD

Dementia has been recognized as a global public health priority. There is currently no cure for dementia, and identifying individuals at an increased risk for dementia may provide a time window for early preventive interventions to reduce the risk or delay the onset of dementia.[1,2] Indeed, previous research has shown that preventive interventions targeting modifiable lifestyle factors in an asymptomatic population may represent a cost-effective approach to reducing dementia risk and the disease burden.[3,4] Hence, a practical tool that could identify individuals at an increased risk for dementia will be crucial for effective interventions to delay onset of the dementia syndrome.

Numerous tools for predicting dementia risk have been developed in the past 2–3 decades, with the predictive accuracy, measured using the concordance (C)-statistic, ranging from 0.64 to 0.91.[1,5–10] However, most of the current dementia risk prediction models have been developed among high-income countries (HICs), especially among European and North American populations.[5,7,9] Of note, an external validation study found that the dementia risk prediction models derived from HICs (e.g., Australia and the Netherlands) were indeed not be applicable to populations from low- and middle-income countries (LMICs) such as China.[11,12] This is partly because the risk prediction models for dementia may vary depending on the ethnic, socioeconomic, and cultural backgrounds of the target populations where genetic and environmental risk factors for dementia may differ.[13] Developing the tailored dementia risk prediction models in LMICs for early preventive interventions is also particularly relevant given the facts that worldwide, around two-thirds of people with dementia are living in LMICs and that the number of people with dementia is projected to increase more rapidly in LMICs than in HICs because of a faster increase in the absolute number of older people in LMICs.[14] China, which has the world largest dementia population, has experienced a steady increase in the prevalence and burden of dementia over the past three decades[15]; furthermore, rural older people are at a heightened risk for dementia and are disproportionally affected by the devastating disorder in China.[16] However, preventive interventions to effectively delay dementia onset among individuals at a high risk for dementia have been hampered by the fact that, currently, there is lack of a practical tool for predicting dementia risk in China. Thus, developing a simple, inexpensive, and easily implementable tool to identify individuals at risk for dementia is crucial to facilitate early preventive interventions.[5,17]

Therefore, in this population-based cohort study, we sought to develop a practical tool for predicting dementia risk among a rural-dwelling older population in China. To facilitate implementation, we developed the tool based primarily on easily available predictors and further transformed it into a user-friendly Web-based dementia risk calculator.

2 | METHODS

This study was conducted and reported following the updated guidance for reporting clinical prediction models that use regression or machine learning methods (TRIPOD+AI).[18]

2.1 | Study populations

This was a population-based prospective cohort study. We used data that were collected at two waves of examination in the same geographical rural region of Yanlou Town, Yanggu County, western Shandong province, China, that is, the Shandong Yanggu Study of Aging and Dementia (SYS-AD) in August–December 2014 and the Multimodal Interventions to Delay Dementia and Disability in Rural China (MIND-China) in March–September 2018, as previously reported.[19–22] In brief, in August–December 2014 (baseline), a total number of 3193 participants (age ≥65 years) were examined for SYS-AD; of these, 973 were excluded due to prevalent dementia that was diagnosed at baseline according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria (n = 201), major psychiatric disorders or incomplete data for diagnosis of dementia status (n = 109), death prior to the follow-up examination (n = 215), and loss to the follow-up examination in 2018 (n = 448). Thus, a total of 2220 participants were included in the current analysis. Figure 1 shows the flowchart of the study participants.

2.2 | Data collection and assessments

At baseline (2014), data were collected by trained medical staff following a structured questionnaire through face-to-face interviews, clinical examinations, cognitive testing, and laboratory tests.[19,21] We collected data on sociodemographic factors (e.g., age, sex, education, marital status, and occupation), lifestyle factors (e.g., smoking, alcohol drinking, oral hygiene, and physical activity), cardiometabolic factors (e.g., hypertension, diabetes, and hyperlipidemia), clinical conditions (e.g., cardiovascular disease, stroke, cancer, traumatic brain injury, and thyroid disorders), and apolipoprotein E (*APOE*) genotype.

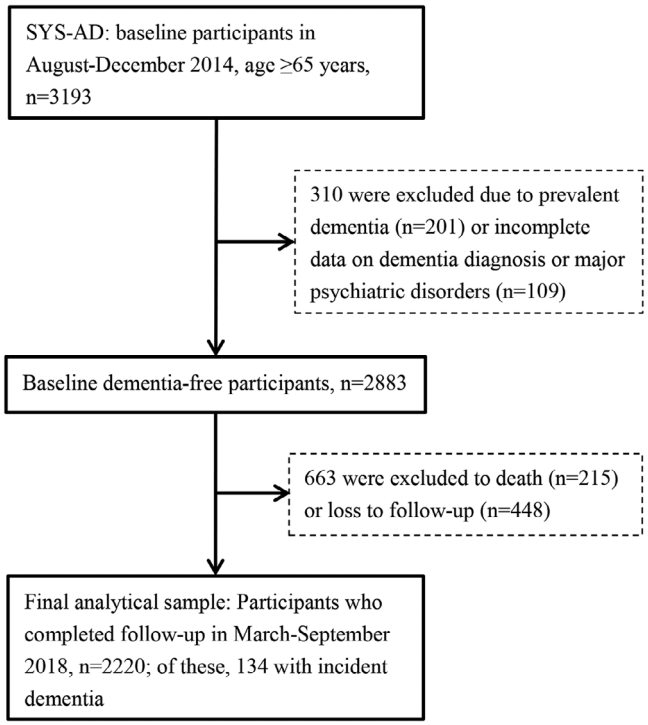

FIGURE 1 Flowchart of study participants, Shandong Yanggu Study on Aging and Dementia, 2014–2018. SYS-AD, the Shandong Yanggu Study on Aging and Dementia.

Arterial blood pressure was measured on the right arm in a sitting position after a 5-min rest using an electronic blood pressure monitor (Omron HEM-7127J; Omron Corporation, Kyoto, Japan), with the cuff maintained at the heart level. Current use of medications (e.g., blood pressure-lowering drugs, blood glucose-lowering drugs, and lipid-lowering drugs) was recorded based on self-reported information, and whenever possible, vials were checked to verify the report. The Ascertain Dementia 8-item Questionnaire (AD8) was administered to participants to assess changes in memory, problem-solving abilities, orientation, and daily activities.[23] After an overnight fast, peripheral blood samples were taken and fasting blood glucose and total cholesterol were measured at the clinical laboratory of local town hospital (DIRUI CS-600B; DIRUI Industrial Co., Ltd., Changchun, China). The *APOE* genotypes were determined using the multiple polymerase chain reaction methods.

We pre-specified 27 potential predictors for dementia according to current literature.[2,24] We dichotomized educational level as "illiterate (no schooling)" and "literacy (primary school or above)," occupation as "farmer" and "non-farmer," and marital status as "married" and "others (e.g., single/widowed/divorced)." The *APOE* genotype was dichotomized into carriers versus non-carriers of the ε4 allele. Tooth brushing habit was categorized as yes versus no. Smoking status was categorized as never/former and current smoking. Alcohol drinking was classified as not drinking alcohol and regular alcohol consumption (i.e., drinking alcohol at least once a week during the past 12 months). Physical inactivity at leisure time was defined as participating in any physical activities less than once a week during the leisure time.

RESEARCH IN CONTEXT

1. **Systematic review**: We searched PubMed for literature on dementia risk prediction models. The large majority of the current prediction models for dementia risk have been developed among high-income countries, especially among European and North American populations. There is a lack of dementia prediction models in the general population in China. The prediction models for dementia risk vary depending on the ethnic, socioeconomic, and cultural backgrounds of the target populations where genetic and environmental risk factors for dementia differ. Thus, the prediction models for dementia risk from high-income countries are not applicable to low- and middle-income countries or ethnically diverse populations.
2. **Interpretations**: We developed a practical tool for predicting dementia risk among a rural-dwelling older population in China. The tool showed good discrimination and excellent calibration for predicting a short-term (up to 4 years) risk of dementia. The prediction tool can be used to identify individuals at a high risk for dementia in the rural older population in China for early interventions. This is highly relevant because evidence has shown that preventive and therapeutic interventions are more effective for delaying cognitive decline and the onset of dementia when being implemented in the early stage (e.g., preclinical or prodromal phase) of the disease. We further created a user-friendly Web-based calculator to facilitate the implementation of this practical tool by health professionals and the public.
3. **Future directions**: Our cohort of Chinese older adults can be a valuable data source for external validation of dementia risk prediction models currently available in the literature. In addition, future research may assess whether the effectiveness of multimodal interventions to delay cognitive decline and the onset of dementia varies by risk levels stratified using this risk prediction model.

Hypertension was defined as systolic blood pressure ≥140 mmHg or diastolic blood pressure ≥90 mmHg or current use of antihypertensive drugs. Diabetes mellitus was defined as having a self-reported history of diabetes, use of glucose-lowering drugs, insulin injection, or fasting blood glucose level ≥7.0 mmol/L. Stroke was ascertained according to self-reported history of stroke and neurological examination.

2.3 | Clinical diagnosis of dementia

The neurologists who specialized in dementia diagnosis and care made the clinical diagnosis of dementia according to the guidelines

outlined in the DSM-IV criteria,[25] in which a three-step diagnostic procedure was followed, as previously described.[23] Briefly, the trained interviewers first made a preliminary judgment for participants who were suspected of having cognitive impairment or dementia based on their performances during the interviews and neuropsychological assessments. Then, the senior neurologists make a preliminary diagnosis of dementia by reviewing all records of functional, medical, neurological, psychiatric, and neuropsychological data, and preliminary judgments made by interviewers. Finally, the neurologists conducted the second face-to-face interviews with participants who received a preliminary diagnosis of dementia or who had insufficient information for the diagnosis of dementia and informants (e.g., family members, neighbors, or village doctors who provided routine primary health care services to local residents), and reassessed their medical history, the date of onset of signs or symptoms of dementia, cognitive status, ADLs, and whenever available, neuroimaging data. Following all the interviews and assessments, a panel of neurologists specializing in dementia diagnosis and care made the final clinical diagnosis of dementia.

2.4 | Statistical analysis

2.4.1 | Model development

The model was developed following the steps below. First, we conducted the Wilcoxon rank-sum test or Student *t*-test for continuous variables, and the chi-square test for categorical variables to screen all factors that were potentially associated with dementia at $p < 0.10$. Then, univariable Cox proportional hazards models were used to examine the association of the predictors selected in the previous step with incident dementia and selected those predictors at $p < 0.10$ for further model development, a cut-off of *p*-value for screening predictors in the development of risk prediction models used in previous studies.[26,27] In the Cox models, we used follow-up time as the time scale, which was estimated from the date of baseline examination to the date of dementia onset for individuals who developed dementia and the date of follow-up examination for those who did not. Third, we used the least absolute shrinkage and selection operator (LASSO) technique and the stepwise multivariable Cox regression analysis to select predictors at $p < 0.05$ for the prediction model from all factors that were identified to be associated with incident dementia in the aforementioned univariable Cox regression analysis. The LASSO method is adapted to penalize the model's regression coefficients, which is particularly useful for preventing model overfitting and model misspecification.[28,29] A penalty parameter λ is used to shrink regression coefficients of predictors, eventually removing predictor variables from the model by setting their respective regression coefficients to zero in LASSO regression using the 10-fold cross-validation. We selected the optimal values for parameters λ based on one standard error away from the minimal error (lambda.1se), aiming to make the model more stable, prevent overfitting, and enhance its generalization to new data. Finally, we used Cox proportional hazard models to develop the final model for predicting dementia risk for up to 4 years among a rural older population in China, which included five predictors, that is, age, education, marital status, self-rated AD8 score, and stroke history.

2.4.2 | Internal validation

The performance of the risk prediction model was evaluated using two parameters: discrimination and calibration. Discrimination refers to the capability of a model to correctly differentiate individuals who will develop dementia during the follow-up period and those who will not.[30] We used the C-index to assess the discriminative ability, which indicates the overall proportion of all individuals that can be classified such that people who will develop dementia during the follow-up period indeed have a higher predicted risk. The C-index ranges from 0.5 to 1.0, with 0.5 indicating random concordance and 1.0 indicating the perfect discrimination of the model.[31] Calibration refers to the agreement between the frequencies of outcomes predicted by the model and the observed frequencies of the outcomes under study, which can be evaluated using calibration plots.[30] We assessed the performance of the risk prediction model by applying the enhanced bootstrap methods with 1000 bootstrap samples for each model to evaluating the robustness of the model.[30] The C-index from the final prediction model was corrected for optimism using the enhanced bootstrap method. The Brier score that indicates the overall performance for a model ranges from 0 for a perfect model to 0.25 for a non-informative model with a 50% incidence of the outcome.[32] In addition, decision curve analysis (DCA) was performed to assess the clinical usefulness of the model regarding whether use of a prediction model in the clinic to inform decision-making would do more good than harm.[33]

2.4.3 | A Web-based nomogram and calculator for predicting dementia risk

Nomogram is a reliable tool to quantify the survival likelihood of an individual by incorporating the important factors that affect occurrence of the disease. This approach has been widely used to evaluate prognostic factors of a disease in the general population. A score is given for each of the contributing factors. Then, the total summary score of all factors is calculated and used to predict the survival probability of each individual.[34,35] In addition, we developed a user-friendly Web-based tool for predicting dementia risk using the R packages, which allows to predict dementia risk of individuals by adding predictors to input fields.

2.4.4 | Power estimation and descriptive statistics

We supposed that Cox-Snell $R^2 = 0.05$, the cumulative incidence of dementia was 0.06, and the follow-up time was 4 years. Using the "pmsampsize" package in R, we estimated that the sample size was 1750 with 105 outcomes.

We presented mean (standard deviation) for continuous variables and frequencies (%) for categorical variables. Baseline characteristics of study participants by dementia status determined at follow-up were compared using Kruskal–Wallis tests for continuous variables and chi-squared test or Fisher's exact test for categorical variables.

Out of the 2220 participants in the analytical sample, the proportion of participants with missing values ranged from < 1% (e.g., diabetes, coronary heart disease, stroke, hypertension, occupation, hyperlipidemia, marital status, and self-rated AD8 score) through 1%–3% (e.g., alcohol drinking, tooth brushing, smoking, chronic kidney disease, visual impairment, hearing problems, physical exercise, chronic obstructive pulmonary disease, and traumatic brain injury) to up to ~5% (e.g., *APOE* genotype, diet, and BMI). To assess the extent to whether the main results might be biased due to missing data on various covariables, we performed sensitivity analyses by imputing missing data with multiple imputation (MI) approach using random forest by missForest package in R.

All analyses were performed using R 4.1.0 packages (R Core Team, 2018; R Foundation for Statistical Computing, Vienna, Austria). A two-tailed $p < 0.05$ was considered statistically significant.

3 | RESULTS

3.1 | Characteristics of study participants

At baseline, the mean age of the 2220 participants was 70.40 years (SD = 4.77), 59.41% were females, and 38.78% were illiterate (Table 1). Compared with individuals without dementia, participants who were diagnosed with dementia at the follow-up examination were older at baseline and less educated, more likely to be females and farmers, less likely to smoke, and more likely to live alone and have worse oral health, unbalanced eating habits, a higher self-rated AD8 score, hearing problems, and a history of stroke ($p < 0.05$), but the two groups did not differ significantly in the proportion of *APOE* ε4 allele (Table 1). In addition, compared with individuals who were examined at follow-up ($n = 2220$), those who were lost to follow-up ($n = 663$) were older, more likely to be males, less educated, less likely to be farmers, live alone, currently smoke, drink alcohol, and have hypertension, hyperlipidemia, and diabetes, and more likely to have coronary heart disease and a history of stroke ($p < 0.05$) (Table S1).

During the mean follow-up of 3.99 (SD = 0.39) years, 134 out of the 2220 participants were diagnosed with incident dementia. The crude cumulative incidence of dementia was 6.04% (95% confidence interval [CI]: 5.10–7.13) in the total sample, 4.11% (2.95%–5.67%) in males, and 7.35% (6.03%–8.93%) in females (for sex difference, $p < 0.05$).

3.2 | Selection of variables for the prediction model

Univariable Cox regression analysis suggested that age, sex, education, marital status, self-rated AD8 score, current smoking, tooth brushing, balanced diet, vegetable-based diet, stroke, hearing problem ($p < 0.05$), alcohol consumption ($p = 0.054$), and regular physical exercise ($p = 0.072$) were significantly or marginally associated with the risk of incident dementia. The multivariable Cox regression analysis showed that age, education, marital status, self-rated AD8 score, and stroke history were significantly associated with incident dementia at $p < 0.05$ (Table 2). The optimal parameter λ in LASSO regression was based on lambda.1se ($\lambda = 0.017$) (Figure S1A). The multivariable Cox regression and the optimal λ in LASSO regression analysis suggested that age, education, marital status, self-rated AD8 score, and stroke history were finally selected in the best fit model for predicting dementia risk.

3.3 | Development and performance of the prediction models

Older age was considered the strongest risk factor and educational attainment was a strong protective factor for dementia. Thus, to assess the extent to which the final prediction model could perform better than the basic models that included age and education, we constructed three models, which included only age in model 1; age and education in model 2; and age, education, self-rated AD8 score, marital status, and stroke history in model 3 (Table 3). The C-index in model 3 (final risk prediction model) was significantly different from those from model 1 ($p < 0.001$) and model 2 ($p = 0.004$). The model 3 showed the best predictive performance in different predicted time periods than the other two models, with the C-index (discriminative ability) being 0.79 (95% CI: 0.75–0.83) for predicting the 4-year risk of dementia (Table 3 and Figure 2). The optimism-adjusted Harrell's C-index of 4-year follow-up was 0.79 and brier score was 0.05 for the final risk prediction model, suggesting the high stability with the best prediction model. An overall good calibration was observed based on the comparable intercepts for models across derivation samples at 2, 3, and 4 years of follow-up (Figure 3). As demonstrated by the favorable probability, the DCA analysis showed the better net benefit of the final risk prediction model than the intervention for all and no intervention (Figure 4).

3.4 | Model application: nomogram and Web-based calculator

A nomogram of the weights and points of the dementia risk prediction score allowed to estimate an individual's probability to develop dementia. We developed a nomogram to visualize the 2-, 3-, and 4-year risk of dementia for individuals with different combinations of five predictors (Figure 5). The use of the nomogram involves three steps. First, on the scale for each variable, the value corresponding to a specific individual is read and the point scale is used to calculate the points for all variable values. Second, the points of all variables obtained in the previous step were added up to have the total point scale. Third, the probability of an event corresponding to the total point of the individual at risk is represented on the risk scale. Finally, we developed a Web-based calculator for estimating the predicted dementia risk of individuals based on the final risk prediction model (Figure S2).

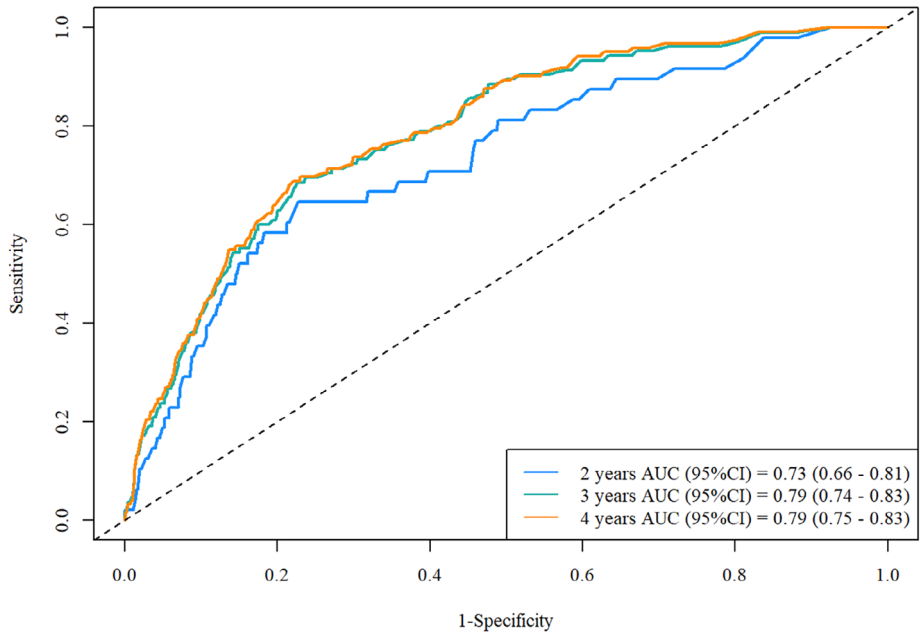

FIGURE 2 The AUC curves of the best model for predicting dementia risk over 2, 3, and 4 years, 2014–2018. AUC, the area under a receiver operating characteristic (ROC) curve; CI, confidence interval.

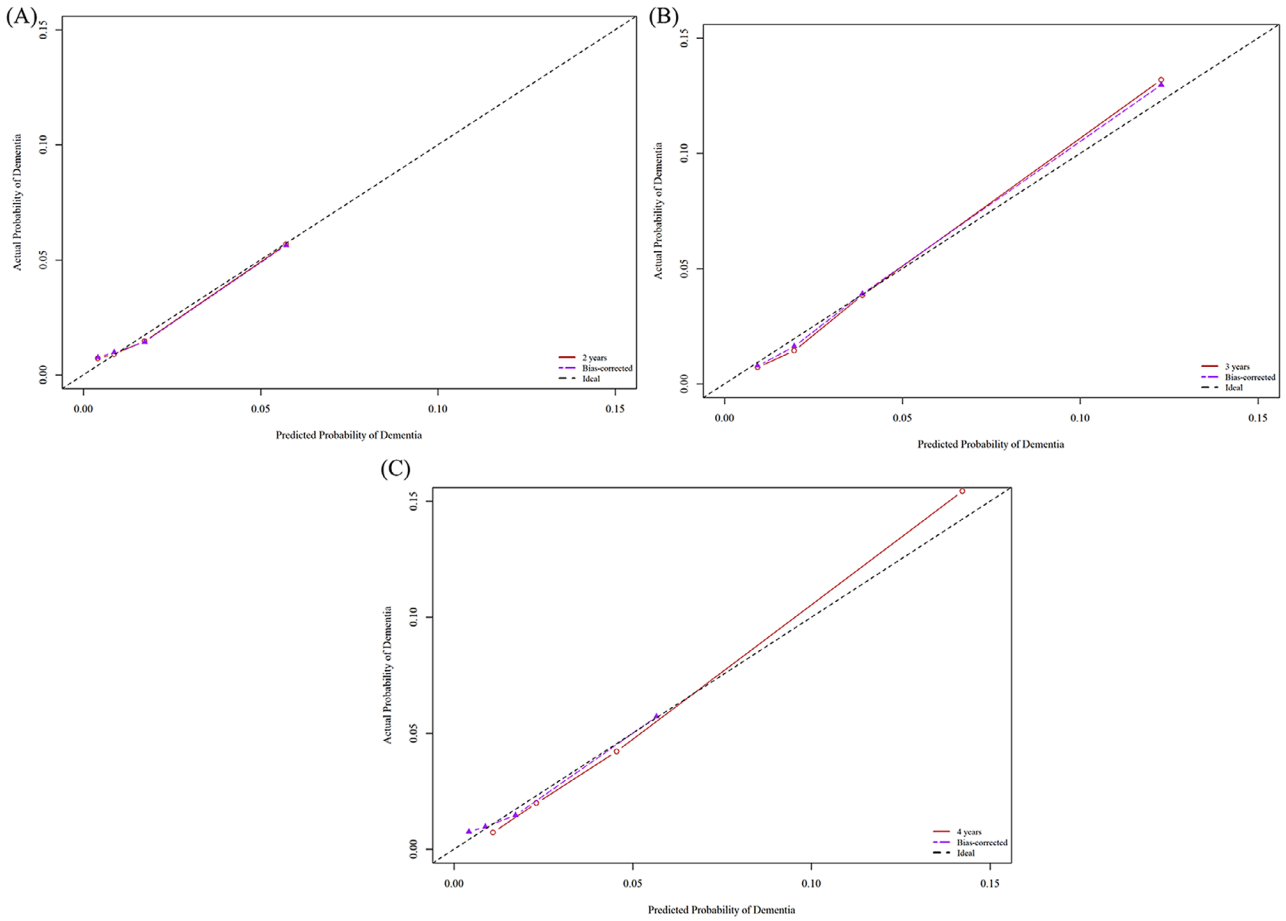

FIGURE 3 Calibration curves of the best model for predicting dementia risk over (A) 2, (B) 3, and (C) 4 years, 2014–2018. The x-axis indicates the probability of predicted incident dementia, and the y-axis indicates the actual probability of occurrence of incident dementia. The black line indicates the agreement between the predicted and actual incident dementia.

TABLE 1 Baseline characteristics of study participants by dementia status at follow-up.

	Total sample	Dementia status at follow-up		
Characteristics[a]	(*n* = 2220)	No (*n* = 2086)	Yes (*n* = 134)	*p*-Value
Age (years)	70.39 (4.77)	70.14(4.59)	74.33 (5.77)	<0.001
Female sex	1319 (59.41)	1222 (58.58)	97 (72.39)	<0.001
Education				
Illiteracy	861 (38.78)	768 (36.82)	93 (69.40)	
Literacy	1359 (61.22)	1318 (63.18)	41 (30.60)	<0.001
Occupation				
Farmers	2005 (90.48)	1877 (90.15)	128 (95.52)	
Non-farmers	211 (9.52)	205 (9.85)	6 (4.48)	0.040
Marital status				
Married	1660 (75.08)	1588 (76.46)	72 (53.73)	
Single/widowed	551 (24.92)	489 (23.54)	62 (46.27)	<0.001
Self-rated AD8 score ≥2	1052 (47.60)	952 (45.86)	100 (74.63)	<0.001
Current smoking	749 (34.11)	716 (34.71)	33 (24.81)	0.019
Alcohol drinking	749 (34.09)	714 (34.59)	35 (26.32)	0.051
Tooth brushing	1048 (47.70)	1001 (48.50)	47 (35.34)	0.003
Balanced diet	388 (18.41)	374 (18.86)	14 (11.20)	0.032
Meat-based diet	105 (5.01)	100 (5.08)	5 (4.00)	0.592
Vegetable-based diet	1675 (77.12)	1563 (76.54)	112 (86.15)	0.011
Regular physical exercise	1254 (58.11)	1172 (57.62)	82 (66.13)	0.062
Traumatic brain injury	133 (6.15)	125 (6.16)	8 (6.11)	0.981
Hypertension	1638 (73.88)	1533 (73.56)	105 (78.95)	0.170
Hyperlipidemia	619 (27.93)	576 (27.65)	43 (32.33)	0.243
Diabetes mellitus	305 (13.74)	289 (13.85)	16 (12.03)	0.553
Thyroid disorders	17 (0.78)	15 (0.73)	2 (1.53)	0.622
Cancer	33 (1.51)	31 (1.51)	2 (1.52)	1.000
Chronic kidney disease	19 (0.87)	17 (0.83)	2 (1.52)	0.733
Visual impairment	289 (13.21)	266 (12.94)	23 (17.42)	0.141
Hearing problem	148 (6.77)	133 (6.48)	15 (11.36)	0.030
Coronary heart disease	407 (18.34)	378 (18.12)	29 (21.80)	0.287
COPD	166 (7.60)	158 (7.69)	8 (6.11)	0.507
Stroke	195 (8.79)	173 (8.30)	22 (16.54)	0.001
Body mass index (kg/m^2)	25.11 (3.75)	25.13 (3.76)	24.69 (3.66)	0.209
APOE ε4 allele	325 (15.37)	304 (15.18)	21 (18.92)	0.287

Note: Data are mean (standard deviation) or *n* (%).

Abbreviations: AD8, The Ascertain Dementia 8-item Questionnaire; *APOE*, apolipoprotein E gene; COPD, chronic obstructive pulmonary disease.

[a]The number of people with missing values was 4 in occupation, 9 in marital status, 10 in self-rated AD8 score, 24 in smoking, 23 in alcohol drinking, 180 in BMI, 23 in tooth brushing, 112 in balanced diet, 126 in meat-based diet, 48 in vegetable-based diet, 62 in exercise, 59 in traumatic brain injury, 3 in hypertension, 4 in hyperlipidemia, 1 in diabetes mellitus, 35 in thyroid disorders, 33 in cancer, 33 in chronic kidney disease, 33 in visual impairment, 34 in hearing problems, 1 in coronary heart disease, 35 in COPD, 2 in stroke history, and 106 in *APOE* genotype.

3.5 | Sensitivity analysis

First, the analyses using the imputed datasets yielded results that were overall similar to those from the analysis of the original complete datasets, as fully reported in Supplementary Tables S2–S4 and Supplementary Figure S1B and Figures S3–S5. Furthermore, the analyses using the competing risk models while taking into account competing risk due to death yielded the results that were overall consistent with those from the original Cox models (Table S5), with the AUC of the final model for predicting 4-year dementia risk being 0.78 (95% CI: 0.74–0.82) as compared with the AUC of 0.79 (95% CI: 0.75–0.83) that was estimated using the original Cox models.

TABLE 2 Hazard ratio (95% CI) of dementia associated with different factors.

	Univariable model			Multivariable model		
Characteristics	***β***	**HR (95% CI)**	***p*-value**	***β***	**HR (95% CI)**	***p*-value**
Age (years)	0.13	1.14 (1.11–1.17)	<0.001	0.10	1.11 (1.07–1.15)	<0.001
Female sex	0.60	1.82 (1.25–2.66)	0.002	—	—	—
Education (literacy vs. illiteracy)	−1.32	0.27 (0.19–0.39)	<0.001	−0.71	0.49 (0.30–0.80)	0.004
Occupation (non-farmers vs. farmers)	−1.15	0.32 (0.12–0.86)	0.023	—	—	
Marital status (single or widowed vs. married)	0.99	2.69 (1.91–3.78)	<0.001	0.53	1.69 (1.13–2.53)	<0.001
AD8 score ≥2	1.21	3.37 (2.28–4.97)	<0.001	0.80	2.21 (1.42–3.45)	<0.001
Current smoking	−0.47	0.63 (0.42–0.93)	0.021	—	—	—
Current alcohol drinking	−0.38	0.68 (0.47–1.01)	0.054	—	—	—
Tooth brushing	−0.53	0.59 (0.42–0.85)	0.004	—	—	—
Balanced diet	−0.60	0.55 (0.32–0.96)	0.036	—	—	—
Vegetable-based diet	0.63	1.88 (1.14–3.09)	0.013	—	—	—
Regular physical exercise	−0.06	0.94 (0.89–1.01)	0.074	—	—	—
Hearing problems	0.58	1.79 (1.05–3.07)	0.033	—	—	—
Stroke	0.75	2.12 (1.34–3.35)	0.001	1.12	3.08 (1.79–5.28)	0.001

Multivariable model included age, sex, occupation, marital status, self-rated AD8 score, smoking, current drinking alcohol, tooth brushing, balanced diet, vegetable-based diet, physical exercise, hearing problem, and stroke.
Abbreviations: AD8, Ascertain Dementia 8-item Questionnaire.; CI, confidence interval; HR, hazard ratio.

TABLE 3 C-index (95% CI) of different models for predicting dementia.

Models	Predictors	C-index	95% CI
Model 1	Age	0.71	0.66–0.76
Model 2	Age, education	0.75	0.71–0.78
Model 3	Age, education, self-rated AD8 score, marital status, stroke	0.79	0.75–0.83

Abbreviations: AD8, Ascertain Dementia 8-item Questionnaire; CI, confidence interval.

4 | DISCUSSION

In this population-based prospective cohort study, we developed a practical tool for predicting dementia risk in a rural older population in China. We used LASSO regression and stepwise Cox proportional hazards to identify five easily available predictors for the risk prediction model that included age, education, marital status, self-rated AD8 score, and stroke. We demonstrated that the practical tool had good performance in predicting up to 4-year risk of dementia, which could identify at-risk individuals for dementia and optimize risk stratification for early preventive interventions as well as for multimodal intervention studies and clinical trials of new treatments. We transformed this model into a user-friendly dementia risk nomogram and a Web-based calculator, which could facilitate communication and implementation of the dementia risk prediction tool in the target populations.

It has been increasingly recognized that preventive or therapeutic interventions for dementia are more effective when targeting individuals who are at an increased risk of dementia.[36,37] Several models for predicting dementia risk that incorporated demographic factors (e.g., age, sex, and education), lifestyles, social factors, and comorbidities with genetic, blood-based, or neuroimaging biomarkers have been reported among North American and European older populations, with the C-statistic ranging from 0.65 to 0.90.[5,38–42] In addition, the Finnish Cardiovascular Risk Factors, Aging and Dementia risk (CAIDE) score was designed to predict late-life dementia risk when people were assessed at midlife of 39–64 years of age.[1] The German Study on Aging, Cognition and Dementia (AgeCoDe) score was developed to predict dementia risk among older adults who were aged ≥75 years and had overall high educational attainment.[43] However, these previous risk prediction models, when applied to Chinese older populations, exhibited poor performance.[11,44] In the present study, we developed a risk prediction model for dementia among a rural Chinese older population (age ≥65 years) with very limited education and the model showed fairly good performance in terms of discriminative ability (area under the curve [AUC] = 0.79). The overall accuracy of our prediction model for dementia risk was comparable to the models previously published from the general older populations and the patient populations in the primary care settings.[40,45]

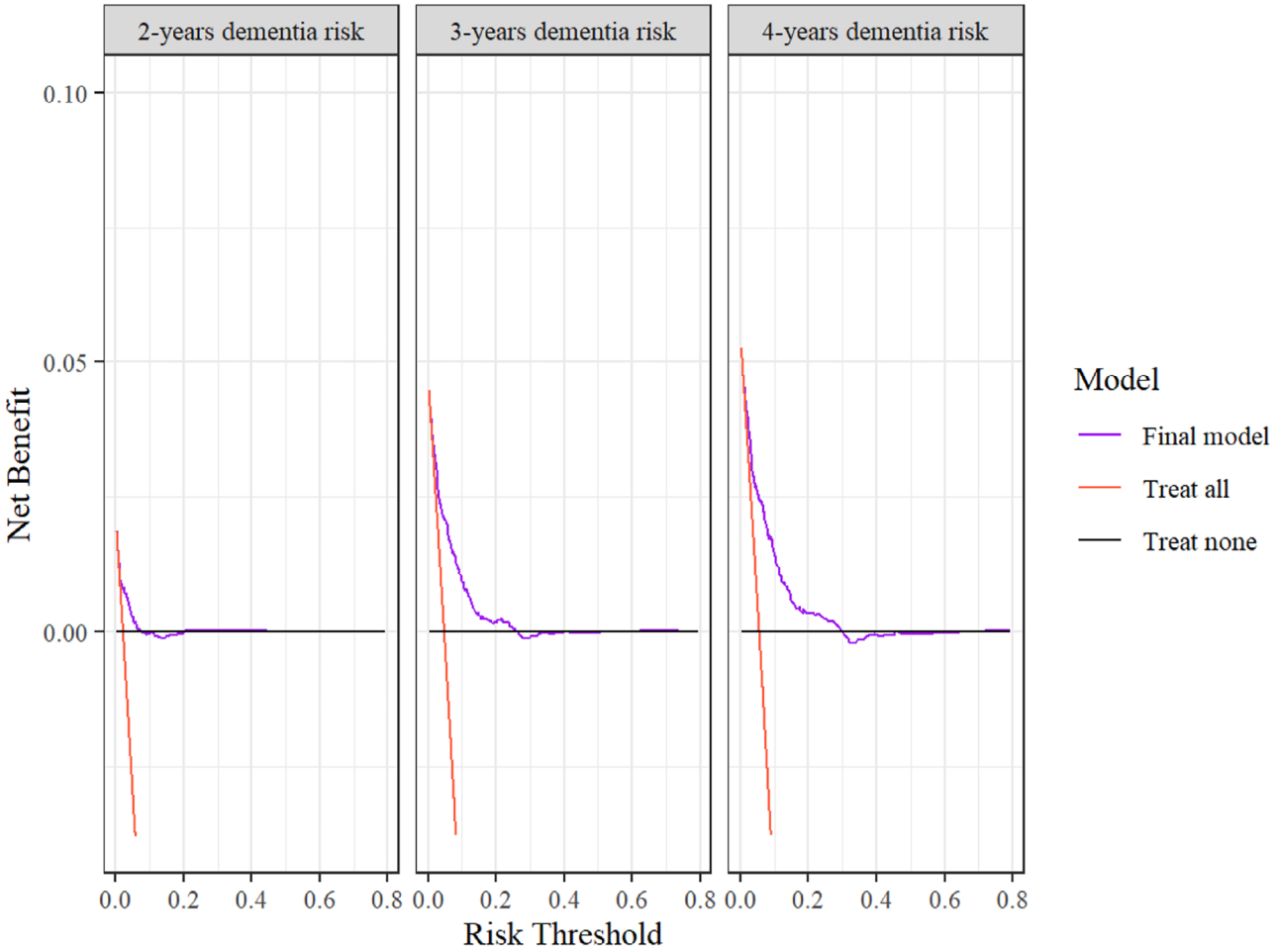

FIGURE 4 Decision curve analysis (DCA) for predicting dementia risk at 2, 3, and 4 years, 2014–2018. The purple line represents the net benefit of the final prediction model across different threshold probabilities; the vertical red line represents the standardized net benefit for the strategy in which all older people underwent the early intervention (treat all); the horizontal solid black line (at 0 on the y-axis) represents the standardized net benefit for the strategy in which none of the older people underwent the early intervention (treat none).

Given the fact that risk and protective factors for dementia vary across ethnically, geographically, and socio-culturally diverse populations, it is imperative to develop tailored risk prediction models in different populations.[46] Older age and low education have been included in almost all the risk prediction models for dementia owing to their strong associations with dementia risk in almost all study populations.[40,44] Thus, we first built the basic risk prediction models by including only age (model 1) and age plus education (model 2). Then, we developed the final risk prediction model (model 3) that included the selected five predictors that aimed to show whether the final prediction model could perform better than the basic models. Of the five predictors in our final prediction model, self-rated AD8 score and marital status were not part of most of the previous dementia risk prediction models developed in European and North American populations. The self-rated AD8 was a brief screening tool for cognitive status sensitive to early cognitive changes, which might be powerful for predicting a short-term risk of dementia. Indeed, some cognitive screening tests (e.g., subjective cognitive decline and MMSE tests) have been included in the prediction models for dementia risk in previous studies,[11,41] but the self-rated AD8 was relatively simple and easy to administer. In addition, our study showed that marital status (e.g., single, widowed, or divorced versus married) was an independent predictor for dementia. This could be due partly to the fact that being married was related to increased access to social engagement, social support, and social integration; these psychosocial factors have been frequently linked to a reduced dementia risk and cognitive health.[47,48]

Our risk prediction model included a history of stroke, instead of various individual cardiovascular risk factors (e.g., hypertension, obesity, diabetes, and hyperlipidemia) that were included in several previous risk prediction models for dementia.[7,49] This was consistent with a dementia prediction model developed in a Dutch population.[40] Stroke, as a major risk factor for dementia, especially for a short-term risk of dementia, could double or triple dementia risk depending on stroke severity, lesion volume, lesion location, and multiplicity.[50] Stroke causes direct injury to the brain structures and functional connections that are critical for cognitive function.[51] Notably, none of the traditional lifestyle and cardiometabolic risk factors was selected for our risk prediction model, due partially to the fact that some of these factors (e.g., hypertension and obesity) could act as risk fac-

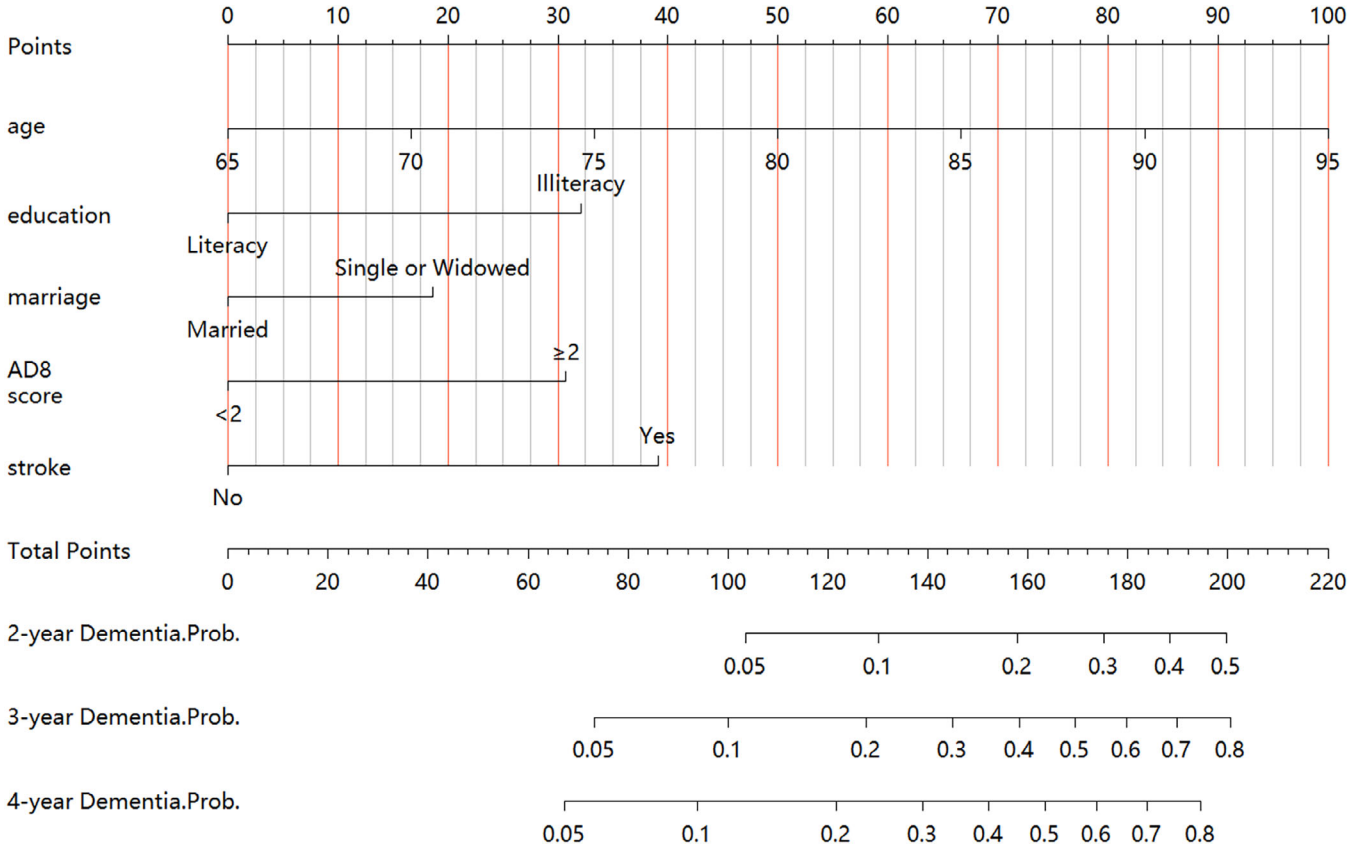

FIGURE 5 The nomogram for the prediction of dementia risk based on five predictors (i.e., age, education, marital status, self-rated AD8 score, and history of stroke). AD8, Ascertain Dementia 8-item Questionnaire.

tors for dementia only when occurring in midlife but not in late life.[52] Of note, the *APOE* ε4 allele was not shown to be an independent risk factor for dementia in our study sample. This was in line with the fact that, compared with European populations, the proportion of carrying the *APOE* ε4 allele was relatively low and the association of *APOE* ε4 allele with AD dementia was weak among Chinese populations.[41,53]

Our risk prediction tool for dementia is based on readily available information without involving complex clinical, biochemical, or genetic biomarkers and could be used in the rural older population to identify individuals at increased risks of dementia for early preventive interventions. In addition, our risk prediction tool can be delivered via the Internet, which provides a simple approach for dementia risk assessment that can be easily used by health professionals (e.g., primary care or memory clinic settings) and the general public as well. This will facilitate implementation of public health policies for early interventions to reduce the risk and delay the onset of dementia, especially regarding multimodal interventions.

Some limitations of our study also deserve mentioning. First, we used a regularization method (LASSO) that automatically selected and subsequently shrank effect sizes of important predictors for the model, which might underestimate the effects of some predictors in the development sample but increase the likelihood of replication in the validation setting. Furthermore, our cohort study engaged older adults who were living in rural communities in China. On one hand, this sociodemographic group has been substantially underrepresented in research of dementia risk prediction, and findings from this study may bridge the relevant knowledge gaps. On the other hand, we do not have appropriate populations for external validation of our risk prediction tool, which deserves further investigation in the future. This should be kept in mind when applying our risk prediction model for dementia to populations with different ethnic, socioeconomic, and sociocultural backgrounds.

In conclusion, we developed and internally validated a practical tool for predicting dementia risk among older adults who were living in the rural communities in China. The tool showed good discrimination and excellent calibration for predicting a short-term (up to 4 years) risk of dementia. The prediction tool can be used to identify individuals at a high risk for dementia for early preventive interventions in the rural population. We further created a user-friendly Web-based calculator to facilitate the implementation of this practical tool by health professionals and the public. This is highly relevant given that evidence has accumulated that preventive and therapeutic interventions are more effective for delaying cognitive decline and clinical onset of the dementia syndrome when being implemented in the early stage (e.g., preclinical or prodromal phase) of the disease. The dataset for our study could be used in the future to externally validate the dementia risk prediction models available in the literature among a rural Chinese older population.

ACKNOWLEDGMENTS

The authors thank all the participants of the SYS-AD Study and the MIND-China Project as well as staff at the Yanlou Town Hospital and the MIND-China Research Group at the Department of Neurology in Shandong Provincial Hospital who were involved in the data collection and management.

The SYS-AD Study was financially supported by the Science and Technology Program for Public Wellbeing of Shandong Province, China (grant no. 2013kjhm180405). The MIND-China Project was financially supported in part by grants from the Brain Science and Brain-Like Intelligence Technology Research Projects of China (grants no.: 2021ZD0201801 and 2021ZD0201808), the National Key R&D Program of China (grants no.: 2017YFC1310100 and 2022YFC3501404), the National Nature Science Foundation of China (grants no.: 81861138008 and 8191101618), the Academic Promotion Program of Shandong First Medical University (grant no.: 2019QL020), and the Taishan Scholar Program of Shandong Province, China (grant no.: tsqn202312347). This work was further funded by the Natural Science Foundation of Shandong Province (grants no.: ZR2022QH106 and ZR2021MH392). C. Qiu received grants from the Swedish Research Council (grants no.: 2017-05819 and 2020-01574) for the Sino-Sweden Joint Research Projects, the Swedish Research Council for Health, Working Life and Welfare (program grant no.: 2023-01125 M. Kivipelto as program PI; C. Qiu as the work-package leader), and the Swedish Foundation for International Cooperation in Research and Higher Education (STINT, grant no.: CH2019-8320) for the Joint China-Sweden Mobility program, Stockholm, Sweden. The funding agency had no role in the study design, data collection and analysis, writing of this manuscript, and in the decision to submit the work for publication.

The SYS-AD and the MIND-China projects were approved by the Ethics Committee at Shandong Provincial Hospital affiliated to Shandong University in Jinan, Shandong. All participants provided written informed consent to participate in the study. Research has been conducted in accordance with the ethical principles for medical research involving human subjects expressed in the Declaration of Helsinki.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest. Author disclosures are available in the Supporting Information.

DATA AVAILABILITY STATEMENT

The study protocol and the original data and codes used in preparation of this article are available from the corresponding author upon reasonable request.

REFERENCES

1. 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
2. Livingston G, Huntley J, Liu KY, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. *Lancet*. 2024;404(10452):572-628. doi:10.1016/s0140-6736(24)01296-0
3. Kivipelto M, Solomon A, Ahtiluoto S, et al. The Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER): study design and progress. *Alzheimers Dement*. 2013;9(6):657-665. doi:10.1016/j.jalz.2012.09.012
4. Ranson JM, Rittman T, Hayat S, et al. Modifiable risk factors for dementia and dementia risk profiling. A user manual for Brain Health Services-part 2 of 6. *Alzheimers Res Ther*. 2021;13(1):169. doi:10.1186/s13195-021-00895-4
5. 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
6. Rawtaer I, Feng L, Yuen VH, et al. A risk score for the prediction of neurocognitive disorders among community-dwelling Chinese older adults. *Dement Geriatr Cogn Disord*. 2016;41(5-6):348-358. doi:10.1159/000447448
7. Exalto LG, Quesenberry CP, Barnes D, Kivipelto M, Biessels GJ, Whitmer RA. Midlife risk score for the prediction of dementia four decades later. *Alzheimers Dement*. 2014;10(5):562-570. doi:10.1016/j.jalz.2013.05.1772
8. Hou XH, Feng L, Zhang C, Cao XP, Tan L, Yu JT. Models for predicting risk of dementia: a systematic review. *J Neurol Neurosurg Psychiatry*. 2019;90(4):373-379. doi:10.1136/jnnp-2018-318212
9. Palmqvist S, Tideman P, Cullen N, et al. Prediction of future Alzheimer's disease dementia using plasma phospho-tau combined with other accessible measures. *Nat Med*. 2021;27(6):1034-1042. doi:10.1038/s41591-021-01348-z
10. Brain J, Kafadar AH, Errington L, et al. What's new in dementia risk prediction modelling? An updated systematic review. *Dement Geriatr Cogn Dis Extra*. 2024;14(1):49-74. doi:10.1159/000539744
11. Stephan BCM, Pakpahan E, Siervo M, et al. Prediction of dementia risk in low-income and middle-income countries (the 10/66 Study): an independent external validation of existing models. *Lancet Glob Health*. 2020;8(4):e524-e535. doi:10.1016/s2214-109x(20)30062-0
12. Alshahrani M, Sabatini S, Mohan D, et al. Dementia risk prediction modelling in low- and middle-income countries: current state of evidence. *Front Epidemiol*. 2024;4:1397754. doi:10.3389/fepid.2024.1397754
13. Andrews SJ, Boeriu AI, Belloy ME, et al. Dementia risk scores, apolipoprotein E, and risk of Alzheimer's disease: one size does not fit all. *Alzheimers Dement*. 2024;20(12):8595-8604. doi:10.1002/alz.14300
14. Wimo A, Seeher K, Cataldi R, et al. The worldwide costs of dementia in 2019. *Alzheimers Dement*. 2023;19(7):2865-2873. doi:10.1002/alz.12901
15. Ding D, Zhao Q, Wu W, et al. Prevalence and incidence of dementia in an older Chinese population over two decades: the role of education. *Alzheimers Dement*. 2020;16(12):1650-1662. doi:10.1002/alz.12159
16. Jia L, Quan M, Fu Y, et al. Dementia in China: epidemiology, clinical management, and research advances. *Lancet Neurol*. 2020;19(1):81-92. doi:10.1016/s1474-4422(19)30290-x
17. 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
18. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024;385:e078378. doi:10.1136/bmj-2023-078378
19. Ren Y, Dong Y, Hou T, et al. Prevalence, incidence, and progression of cognitive impairment, no dementia among rural-dwelling Chinese

older adults. *J Alzheimers Dis*. 2022;85(4):1583-1592. doi:10.3233/JAD-215236
20. Wang Y, Han X, Zhang X, et al. Health status and risk profiles for brain aging of rural-dwelling older adults: data from the interdisciplinary baseline assessments in MIND-China. *Alzheimers Dement (N Y)*. 2022;8(1):e12254. doi:10.1002/trc2.12254
21. Liu R, Ren Y, Hou T, et al. Associations of sleep timing and time in bed with dementia and cognitive decline among Chinese older adults: a cohort study. *J Am Geriatr Soc*. 2022;70(11):3138-3151. doi:10.1111/jgs.18042
22. Hou T, Liu K, Fa W, et al. Association of polygenic risk scores with Alzheimer's disease and plasma biomarkers among Chinese older adults: a community-based study. *Alzheimers Dement*. 2024;20(10):6669-6681. doi:10.1002/alz.13924
23. Dong Y, Wang Y, Liu K, et al. Dementia screening in rural-dwelling Chinese older adults: the utility of a smell test and the self-rated AD8. *J Am Geriatr Soc*. 2022;70(4):1106-1116. doi:10.1111/jgs.17586
24. Ren L, Liang J, Wan F, Wang Y, Dai XJ. Development of a Clinical Risk Score Prediction Tool for 5-, 9-, and 13-year risk of dementia. *JAMA Netw Open*. 2022;5(11):e2242596. doi:10.1001/jamanetworkopen.2022.42596
25. American Psychiatric Association (APA). *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed. (DSM-IV). APA; 1994.
26. Sorror ML, Storer BE, Fathi AT, et al. Development and validation of a novel acute myeloid leukemia-composite model to estimate risks of mortality. *JAMA Oncol*. 2017;3(12):1675. doi:10.1001/jamaoncol.2017.2714
27. Chowdhury MZI, Turin TC. Variable selection strategies and its importance in clinical prediction modelling. *Fam Med Community Health*. 2020;8(1):e000262. doi:10.1136/fmch-2019-000262
28. Pavlou M, Ambler G, Seaman SR, et al. How to develop a more accurate risk prediction model when there are few events. *BMJ*. 2015;351:h3868. doi:10.1136/bmj.h3868
29. Tibshirani R. Regression shrinkage and selection via the LASSO. *J R Stat Soc Series B Stat Methodol*. 1996;58:267-288. doi:10.1111/j.2517-6161.1996.tb02080.x
30. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. Springer; 2008.
31. Wang Y, Li J, Xia Y, et al. Prognostic nomogram for intrahepatic cholangiocarcinoma after partial hepatectomy. *J Clin Oncol*. 2013;31(9):1188-1195. doi:10.1200/jco.2012.41.5984
32. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models. *Epidemiology*. 2010;21(1):128-138. doi:10.1097/EDE.0b013e3181c30fb2
33. Vickers AJ, Holland F. Decision curve analysis to evaluate the clinical benefit of prediction models. *Spine J*. 2021;21(10):1643-1648. doi:10.1016/j.spinee.2021.02.024
34. Zhang Z, Kattan MW. Drawing nomograms with R: applications to categorical outcome and survival data. *Ann Transl Med*. 2017;5(10):211. doi:10.21037/atm.2017.04.01
35. Hartaigh BO, Gransar H, Callister T, et al. Development and validation of a simple-to-use nomogram for predicting 5-, 10-, and 15-year survival in asymptomatic adults undergoing coronary artery calcium scoring. *JACC Cardiovasc Imaging*. 2018;11(3):450-458. doi:10.1016/j.jcmg.2017.03.018
36. Kivipelto M, Mangialasche F, Ngandu T. Can lifestyle changes prevent cognitive impairment?. *Lancet Neurol*. 2017;16(5):338-339. doi:10.1016/s1474-4422(17)30080-7
37. Sommerlad A, Livingston G. Preventing Alzheimer's dementia. *BMJ*. 2017;359:j5667. doi:10.1136/bmj.j5667
38. Tang EY, 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
39. 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
40. Licher S, Leening MJG, Yilmaz P, et al. Development and validation of a Dementia Risk Prediction Model in the general population: an analysis of three longitudinal studies. *Am J Psychiatry*. 2019;176(7):543-551. doi:10.1176/appi.ajp.2018.18050566
41. Hall A, Pekkala T, Polvikoski T, et al. Prediction models for dementia and neuropathology in the oldest old: the Vantaa 85+ cohort study. *Alzheimers Res Ther*. 2019;11(1):11. doi:10.1186/s13195-018-0450-3
42. Bermudez C, Graff-Radford J, Syrjanen JA, et al. Plasma biomarkers for prediction of Alzheimer's disease neuropathologic change. *Acta Neuropathol*. 2023;146(1):13-29. doi:10.1007/s00401-023-02594-w
43. Heser K, Tebarth F, Wiese B, et al. Age of major depression onset, depressive symptoms, and risk for subsequent dementia: results of the German Study on Ageing, Cognition, and Dementia in Primary Care Patients (AgeCoDe). *Psychol Med*. 2012;43(8):1597-1610. doi:10.1017/s0033291712002449
44. Sindi S, Calov E, Fokkens J, et al. The CAIDE Dementia Risk Score App: the development of an evidence-based mobile application to predict the risk of dementia. *Alzheimers Dement (Amst)*. 2015;1(3):328-333. doi:10.1016/j.dadm.2015.06.005
45. Geethadevi GM, Peel R, Bell JS, et al. Validity of three risk prediction models for dementia or cognitive impairment in Australia. *Age Ageing*. 2022;51(12):afac307. doi:10.1093/ageing/afac307
46. Anstey KJ, Peters R, Zheng L, et al. Future directions for dementia risk reduction and prevention research: an International Research Network on Dementia Prevention Consensus. *J Alzheimers Dis*. 2020;78(1):3-12. doi:10.3233/jad-200674
47. Liu H, Zhang Z, Choi S-W, Langa KM, Carr D. Marital status and dementia: evidence from the Health and Retirement Study. *J Gerontol B Psychol Sci Soc Sci*. 2020;75(8):1783-1795. doi:10.1093/geronb/gbz087
48. Fratiglioni L, Marseglia A, Dekhtyar S. Ageing without dementia: can stimulating psychosocial and lifestyle experiences make a difference? *Lancet Neurol*. 2020;19(6):533-543. doi:10.1016/s1474-4422(20)30039-9
49. Chosy EJ, Edland SD, Gross N, et al. The CAIDE Dementia Risk Score and the Honolulu-Asia Aging Study. *Dement Geriatr Cogn Disord*. 2019;48(3-4):164-171. doi:10.1159/000504801
50. Kuźma E, Lourida I, Moore SF, Levine DA, Ukoumunne OC, Llewellyn DJ. Stroke and dementia risk: a systematic review and meta-analysis. *Alzheimers Dement*. 2018;14(11):1416-1426. doi:10.1016/j.jalz.2018.06.3061
51. Rost NS, Brodtmann A, Pase MP, et al. Post-stroke cognitive impairment and dementia. *Circ Res*. 2022;130(8):1252-1271. doi:10.1161/circresaha.122.319951
52. Qiu C, Fratiglioni L. A major role for cardiovascular burden in age-related cognitive decline. *Nat Rev Cardiol*. 2015;12(5):267-277. doi:10.1038/nrcardio.2014.223
53. Corrada MM, Paganini-Hill A, Berlau DJ, Kawas CH. Apolipoprotein E genotype, dementia, and mortality in the oldest old: the 90+ Study. *Alzheimers Dement*. 2013;9(1):12-18. doi:10.1016/j.jalz.2011.12.004

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: Liu K, Hou T, Li Y, et al. Development and internal validation of a risk prediction model for dementia in a rural older population in China. *Alzheimer's Dement*. 2025;21:e14617. https://doi.org/10.1002/alz.14617