Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Association of educational attainment with incident orientation impairment among Chinese older adults with and without stroke: evidence from a longitudinal study with external validation

Ce Shi
Anhui University of Traditional Chinese Medicine
Lihua Wu
Henan University of Traditional Chinese Medicine
Qiqi Yang
Anhui University of Traditional Chinese Medicine
Fei Wang
Anhui University of Traditional Chinese Medicine
Xiang Shang
Anhui University of Traditional Chinese Medicine
Tianxin Jiang
Anhui University of Traditional Chinese Medicine
Baoguo Wang
Anhui University of Traditional Chinese Medicine
Jianhong Gao
Anhui University of Traditional Chinese Medicine
Weiran Li
Anhui University of Traditional Chinese Medicine
Ziyu Ye
Anhui University of Traditional Chinese Medicine
Fei Li
leagcen@163.com

Anhui University of Traditional Chinese Medicine

Research Article

Keywords: Orientation impairment, Educational attainment, Stroke history, Cognitive reserve, External validation

Posted Date: April 14th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9092847/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Orientation impairment is an early and clinically meaningful marker of cognitive vulnerability in older adults and is associated with subsequent functional decline, dementia, and mortality. Although stroke is a well-established risk factor for late-life cognitive impairment, its association with incident orientation impairment in community-dwelling populations remains unclear. Educational attainment, a key indicator of cognitive reserve, may modify the cognitive consequences of stroke, but longitudinal evidence addressing this hypothesis is limited.

Methods

We conducted a population-based longitudinal study using data from the China Health and Retirement Longitudinal Study (CHARLS). Participants aged ≥ 60 years who were free of orientation impairment at baseline (2018) were followed for incident orientation impairment in 2020. Stroke history was defined by self-reported physician diagnosis at baseline. Educational attainment was primarily modeled as a continuous variable (per three-year increase) and additionally evaluated using categorical specifications in sensitivity analyses. Modified Poisson regression with robust variance estimation was used to estimate relative risks (RRs). Effect modification by education was assessed using interaction terms and marginal effect estimation. Sensitivity analyses included inverse probability weighting for loss to follow-up, competing-risk models accounting for death, alternative outcome definitions, and complementary machine-learning analyses. External validation was conducted to examine the transportability of relative risk patterns rather than to formally validate a clinical prediction model.

Results

Among 9,080 participants free of orientation impairment at baseline, 5,289 participants with observed follow-up orientation assessments were included (complete-case with respect to the outcome), of whom 744 developed incident orientation impairment over two years. Stroke history was not independently associated with incident orientation impairment after multivariable adjustment (RR 0.88, 95% CI 0.63–1.24). In contrast, higher educational attainment showed a strong and consistent protective association: each additional three years of education was associated with a 25.6% lower risk of incident orientation impairment (RR 0.74, 95% CI 0.65–0.86). Formal tests for stroke–education interaction were not statistically significant on multiplicative or additive scales; however, marginal estimates were directionally consistent with attenuation of stroke-associated risk at higher education levels, but estimates were imprecise and confidence intervals crossed the null. Results were robust across sensitivity analyses. Machine-learning models demonstrated modest discrimination and did not materially outperform traditional regression models. External validation showed moderate discrimination

but indicated miscalibration of absolute risk estimates, suggesting the need for recalibration in different population settings.

Conclusions

In this nationally representative cohort of older adults, educational attainment emerged as a consistent and robust protective factor against incident orientation impairment, whereas stroke history showed no independent association after adjustment. Evidence for education modifying the stroke–orientation relationship was suggestive but inconclusive. These findings highlight the central role of educational gradients in late-life cognitive vulnerability and underscore the challenges of detecting domain-specific cognitive consequences of stroke in population-based studies.

Background

Population ageing has substantially increased the burden of cognitive impairment and related functional decline worldwide, posing major challenges to health and social care systems for older adults[1–3]. Among different cognitive domains, orientation represents a core component of cognitive functioning and is essential for independent living and everyday decision-making in older adults[4, 5]. Impairment in orientation has been associated with an increased risk of subsequent dementia, functional dependence, and mortality, making it a clinically relevant and epidemiologically meaningful early marker of cognitive vulnerability in later life[6, 7].

Stroke is a leading cause of long-term disability and cognitive impairment in older adults[8]. Beyond its acute neurological consequences, stroke may initiate or accelerate vascular and neurodegenerative processes that adversely affect cognitive functioning over time[9]. Previous studies have consistently demonstrated an increased risk of global cognitive decline and dementia among stroke survivors[10]. However, evidence regarding domain-specific cognitive outcomes—particularly orientation impairment as an early and clinically interpretable marker—remains limited in population-based longitudinal studies.

Importantly, the cognitive consequences of stroke are unlikely to be uniform across individuals and may vary according to markers of cognitive reserve[11, 12]. Educational attainment, a core indicator of cognitive reserve and a key social determinant of health, has been widely recognised as an important determinant of cognitive ageing trajectories[13]. Higher educational attainment may buffer the cognitive consequences of brain pathology through enhanced neural efficiency and compensatory mechanisms[14]. Nevertheless, whether educational attainment modifies the association between stroke history and subsequent orientation impairment remains uncertain, particularly in longitudinal, population-based settings.

Most existing studies have treated education primarily as a confounder or have focused on global cognitive outcomes, with limited attention to potential effect modification across specific cognitive domains. From a public health perspective, understanding heterogeneity in stroke-related cognitive risk

is essential for identifying vulnerable subgroups and developing targeted prevention strategies. This question is particularly relevant in rapidly ageing populations such as China, where substantial educational disparities persist among older adults as a result of historical and socioeconomic factors[15–17].

Using data from a nationally representative longitudinal cohort of Chinese older adults, this study aimed to examine the association between stroke history and incident orientation impairment and to explore whether this association differs according to educational attainment. We further evaluated the robustness of our findings using multiple analytical approaches, including inverse probability weighting, competing risk models accounting for mortality, and complementary machine learning analyses.

Methods

Study design and data source

This study was a population-based longitudinal analysis using data from the China Health and Retirement Longitudinal Study (CHARLS), a nationally representative survey of middle-aged and older adults in China. CHARLS employs a multistage, stratified probability sampling design covering both urban and rural areas and collects comprehensive information on sociodemographic characteristics, health conditions, and cognitive functioning among community-dwelling adults.

Baseline data were obtained from the 2018 wave, with follow-up assessments conducted in 2020. The CHARLS study protocol was approved by the Institutional Review Board of Peking University, and all participants provided written informed consent. The present analysis was based on publicly available, de-identified data.

Study population

The analytic sample was restricted to participants aged 60 years or older at baseline (2018). Among the 5,289 participants included in the analytic sample, the mean age was 61.44 years (SD 10.02). To ensure a clear temporal ordering between exposure and outcome, analyses were further restricted to individuals who were free of orientation impairment at baseline.

Participants were included if they had available baseline information on stroke history and educational attainment and completed the orientation assessment at the 2020 follow-up. Individuals without observed orientation assessments at follow-up were excluded from the primary analysis (complete-case with respect to the outcome). Baseline covariates were retained as observed, with missing categories incorporated for categorical variables in regression models. Potential bias arising from loss to follow-up was examined using inverse probability weighting in sensitivity analyses.

Exposure assessment: stroke history

Baseline stroke history was assessed using self-reported physician diagnosis. Participants were asked whether they had ever been diagnosed with stroke by a physician, and responses were coded as a binary variable (yes or no). Self-reported physician-diagnosed stroke has been widely used in population-based studies and has demonstrated acceptable validity in the CHARLS cohort.

Outcome assessment: orientation impairment

Cognitive performance in CHARLS is evaluated using a standardized methodological framework consistent with that adopted in the U.S. Health and Retirement Study (HRS)[18]. As described by Crimmins et al., this framework conceptualizes cognition across four domains: orientation, calculation, memory, and visuospatial ability. These domains are assessed using structured cognitive tests, yielding a total cognitive score ranging from 0 to 31 points.

Orientation ability is measured using items derived from the Cognitive Status Telephone Interview, including questions on date and season, with a maximum score of 5 points. Calculation ability is assessed through a serial subtraction task (subtracting 7 from 100 consecutively), with a maximum score of 5 points. Memory is evaluated using immediate and delayed word recall tasks, with a maximum score of 20 points. Visuospatial ability is assessed through a figure-copying task, contributing 1 point to the total score.

The present study focuses specifically on the orientation domain as the primary outcome. The decision to concentrate on orientation rather than global cognitive score was based on several considerations. First, orientation reflects a core component of early cognitive decline and demonstrates structural stability across survey waves. Second, orientation items exhibit strong measurement consistency over time, facilitating longitudinal comparisons. Third, impairment in orientation has well-established neurobiological relevance in post-stroke cognitive dysfunction.

Incident orientation impairment was defined as the occurrence of impaired orientation at the 2020 follow-up among participants who were free of orientation impairment at baseline in 2018. By restricting the analysis to incident cases, this study minimizes potential reverse causation and strengthens temporal inference between baseline stroke history and subsequent cognitive decline.

Educational attainment

Educational attainment was measured at baseline. In the primary analysis, education was modeled as years of schooling and scaled per three-year increase to facilitate interpretability. Categorical specifications reflecting the Chinese education system were examined in supplementary analyses: no formal education, primary education, middle school education, and high school or higher education. Educational attainment was conceptualised as a stable indicator of early-life socioeconomic position and cognitive reserve and was prespecified as a potential effect modifier of the association between stroke history and incident orientation impairment.

Covariates

Baseline covariates were selected a priori based on previous literature and conceptual considerations. These included demographic characteristics (age and sex), socioeconomic factors (marital status and residential area), and major baseline health-related variables. All covariates were measured at baseline to minimise temporal ambiguity and reduce the risk of overadjustment.

Statistical analysis

The primary objective was to examine the association between baseline stroke history and incident orientation impairment and to explore whether this association differed according to educational attainment. Modified Poisson regression with cluster-robust standard errors was used to estimate relative risks (RRs) and 95% confidence intervals (CIs). Participants with missing orientation assessment at follow-up were excluded from the primary analysis (complete-case with respect to the outcome only). Baseline covariates were retained as observed. For categorical variables, missing values were included as separate indicator categories in regression models rather than excluding participants with incomplete covariate data. This approach is well suited to cohort studies with binary outcomes and provides directly interpretable risk estimates.

Educational attainment was first included as an adjustment variable in multivariable models. Potential effect modification by educational attainment was evaluated by introducing an interaction term between stroke history and education. Statistical evidence for interaction was assessed using Wald tests. To facilitate interpretation, marginal and education-specific associations between stroke history and incident orientation impairment were derived from interaction models.

Several sensitivity analyses were conducted to assess the robustness of the findings. Inverse probability weighting was applied to address potential selection bias due to loss to follow-up. Competing-risk analyses were performed to account for death during follow-up as a competing event for incident orientation impairment. Additional analyses included alternative definitions of orientation impairment and complementary machine-learning approaches as robustness checks rather than as primary inferential tools.

Complementary machine-learning analyses

To further assess the robustness of the observed associations across analytical frameworks, complementary machine-learning analyses were conducted. Multiple supervised learning algorithms were trained using baseline variables to predict incident orientation impairment over the follow-up period. Model performance was evaluated using cross-validated out-of-fold procedures, with discrimination and calibration assessed as supportive metrics.

These analyses were not intended to develop or optimise a predictive model for clinical use. Rather, they served as robustness checks to examine whether the relative importance and direction of associations

for stroke history and educational attainment were consistent with findings from the primary regression-based analyses.

External validation

External validation was performed using data from the Chinese Longitudinal Healthy Longevity Survey (CLHLS), an independent population-based cohort of older adults in China[19]. Inclusion criteria and outcome definitions were aligned as closely as possible with those used in the primary analysis, with an emphasis on conceptual rather than exact measurement equivalence.

The purpose of external validation was to assess the transportability of the observed associations and overall risk patterns, rather than to formally validate a prediction model[20]. Modified Poisson regression models were re-estimated in the CLHLS cohort to assess consistency of relative associations. Cross-validated discrimination and calibration metrics were examined descriptively to evaluate consistency with findings from CHARLS[21].

Statistical software

All statistical analyses were conducted using R software (version 4.5.2). Statistical significance was defined as a two-sided p-value < 0.05.

Results

Study population and event rates

Among CHARLS participants who were free of orientation impairment at baseline in 2018, 9,080 individuals constituted the risk set. Of these, 5,289 participants had observed orientation assessments at the 2020 follow-up and were included in the primary complete-case analysis. The mean age of the analytic sample was 61.44 years (SD 10.02). Over the two-year follow-up, 744 participants developed incident orientation impairment under the primary definition (orientation score ≤ 3). Using a stricter definition (orientation score ≤ 2), 231 incident cases were identified. The cohort derivation and analytic sample selection are shown in Fig. 1.

Participants were drawn from the China Health and Retirement Longitudinal Study (CHARLS). From the original master dataset (N = 19,395), individuals aged < 60 years or without valid baseline orientation assessment were excluded. The final risk set included 9,080 participants free of orientation impairment at baseline (2018). Among them, 5,289 participants with non-missing orientation assessment at follow-up (2020) were included in the primary complete-case analysis, yielding 744 incident cases of orientation impairment. Primary analyses used modified Poisson regression with cluster-robust standard errors; additional spline, interaction, inverse probability weighting, and machine-learning analyses were conducted as sensitivity and supplementary analyses.

Baseline characteristics stratified by stroke history among participants included in the analytic sample (complete-case with respect to follow-up outcome only; N = 5,289) are shown in Table 1.

Participants with a history of stroke had a substantially higher prevalence of vascular comorbidities, with very large standardized mean differences observed for hypertension and diabetes, indicating pronounced baseline imbalance between groups, whereas age and educational attainment were broadly comparable between groups.

Table 1
Baseline characteristics by stroke history

Characteristic		No stroke	Stroke history	SMD
N		5024	265	
Age, years		61.44 (10.06)	61.53 (9.33)	0.009
Age missing		0 (0.0%)	0 (0.0%)	
Education, years		3.79 (1.94)	3.97 (1.88)	0.090
Education missing		0 (0.0%)	0 (0.0%)	
Sex	Male	2374 (47.3%)	124 (46.8%)	0.013
	Female	2650 (52.7%)	141 (53.2%)	
	Missing	0 (0.0%)	0 (0.0%)	
Hukou	1 Agricultural Hukou	3560 (70.9%)	187 (70.6%)	0.093
	2 Non-agricultural Hukou	1323 (26.3%)	73 (27.5%)	
	3 Unified Residence Hukou	15 (0.3%)	0 (0.0%)	
	Missing	126 (2.5%)	5 (1.9%)	
Hypertension	No	3214 (64.0%)	81 (30.6%)	0.880
	Yes	518 (10.3%)	48 (18.1%)	
	Missing	1292 (25.7%)	136 (51.3%)	
Diabetes	No	3850 (76.6%)	153 (57.7%)	0.541
	Yes	514 (10.2%)	48 (18.1%)	
	Missing	660 (13.1%)	64 (24.2%)	
Myocardial infarction	No	4367 (86.9%)	202 (76.2%)	0.383
	Yes	262 (5.2%)	29 (10.9%)	
	Missing	395 (7.9%)	34 (12.8%)	

Notes: Values are mean (SD) for continuous variables and n (%) for categorical variables. SMD = standardized mean difference.

Association between stroke history, education, and incident orientation impairment

In multivariable modified Poisson regression models with cluster-robust standard errors, baseline stroke history was not independently associated with incident orientation impairment after adjustment for age,

sex, educational attainment, hukou status, and vascular comorbidities (RR 0.88, 95% CI 0.63–1.24; Table 2).

In contrast, educational attainment showed a strong and consistent protective association with incident orientation impairment. When modelled as a continuous variable, each additional three years of education was associated with a 25.6% lower risk of developing orientation impairment over two years (RR 0.74, 95% CI 0.65–0.86). The main multivariable model estimates are shown in Fig. 2.

Table 2
Association of stroke history and educational attainment with incident orientation impairment (CHARLS 2018–2020)

Variable	RR	95% CI	P value
Stroke history (yes vs no)	0.888	0.630–1.251	0.498
Education (per 3-year increase)	0.752	0.653–0.867	< 0.001
Age (per 1-year increase)	0.998	0.992–1.005	0.619
Female (vs male)	0.907	0.789–1.041	0.164
Hukou: Non-agricultural (ref: Agricultural)	0.631	0.509–0.783	< 0.001
Hukou: Missing (ref: Agricultural)	0.847	0.529–1.356	0.490
Hukou: Unified residence (ref: Agricultural)	0.504	0.074–3.431	0.484
Hypertension (yes vs no)	0.897	0.705–1.142	0.377
Diabetes (yes vs no)	0.960	0.762–1.209	0.729
Myocardial infarction (yes vs no)	0.895	0.657–1.218	0.480

Notes: Relative risks (RRs) were estimated using modified Poisson regression with a log link and cluster-robust standard errors (clustered by community/household). The model adjusted for stroke history, education (per 3-year increase), age (linear), sex, hukou, hypertension, diabetes, and myocardial infarction. Missing indicator categories were included for categorical covariates in all regression models. Analytic sample: N = 5289, events = 744.

Relative risks (RRs) and 95% confidence intervals (CIs) were estimated using modified Poisson regression with cluster-robust standard errors. Models were adjusted for age, sex, hukou status, hypertension, diabetes, and myocardial infarction. Relative risks are shown on a logarithmic scale.

To further characterize the education-related gradient, adjusted relative risks across years of education, relative to 9 years of education, are shown in Fig. 3.

Adjusted relative risks of incident orientation impairment across years of education were estimated using a modified Poisson regression model with cluster-robust standard errors. Relative risks are presented relative to 9 years of education (reference). Shaded areas represent 95% confidence intervals.

Education non-linearity check

To evaluate whether the education–risk relationship was non-linear, we compared the linear education specification against a natural spline specification. The spline model yielded a marginally lower AIC than the linear model (AIC 4361.22 vs 4362.01), but the likelihood ratio test did not provide strong evidence of non-linearity (LRT p = 0.091). Therefore, the linear per-3-year education gradient was retained as the primary, interpretable specification, with spline results considered supportive (Fig. 4).

Natural spline models were used to assess potential non-linearity in the association between years of education and incident orientation impairment, with 9 years as the reference. Shaded areas represent 95% confidence intervals. The spline analysis was conducted to evaluate the functional form of the association rather than to support causal interpretation at extreme education levels.

Effect modification by educational attainment

The multiplicative interaction between stroke history and educational attainment (per three-year increase in education) was not statistically significant (interaction term (stroke × education per three-year increase): RR 0.74, 95% CI 0.43–1.26). Although the direction of the interaction suggested a possible attenuation of stroke-associated risk at higher education levels, uncertainty around the estimates was substantial (Fig. 5).

Predicted risks were derived from the fully adjusted modified Poisson model including a stroke-by-education interaction term. Solid lines represent point estimates and shaded areas indicate 95% confidence intervals. Estimates at extreme education levels should be interpreted cautiously due to limited sample size.

To facilitate interpretation, marginal relative risks of stroke history were derived at selected levels of education from the interaction model (Table 3). The estimated relative risk of stroke was highest among participants with no formal education and decreased monotonically with increasing education. However, all confidence intervals crossed the null, indicating insufficient statistical evidence to support definitive effect modification.

Table 3
Marginal relative risks of stroke history across education levels

Education (years)	Stroke RR	95% CI
0	1.252	0.656–2.388
6	0.679	0.362–1.271
9	0.500	0.163–1.530
12	0.368	0.071–1.898

Exploratory analyses assessing additive interaction using a binary education threshold (≥ 9 years vs. <9 years) yielded imprecise and unstable estimates, reflecting sparse data within some strata. Accordingly,

results from additive interaction analyses were considered exploratory.

Sensitivity analyses

Inverse probability weighting was applied to address potential bias due to loss to follow-up. The handling of baseline covariate missingness (via missing indicator categories) remained unchanged in weighted models. After weighting and truncation of extreme weights, effect estimates were highly consistent with those from the primary analysis. Stroke history remained non-significantly associated with incident orientation impairment (RR 0.90, 95% CI 0.64–1.27), while educational attainment retained a strong protective association (per + 3 years: RR 0.73, 95% CI 0.64–0.85).

When a stricter outcome definition was applied, the association between stroke history and incident orientation impairment remained non-significant, although the point estimate shifted above unity (RR 1.19, 95% CI 0.69–2.03). The protective association of education persisted but was attenuated and less precise, consistent with reduced statistical power due to the smaller number of events.

Competing-risk analyses accounting for death as a competing event produced results broadly consistent with the primary findings. Higher educational attainment was associated with a markedly lower subdistribution hazard of incident orientation impairment, whereas stroke history showed no clear association.

Complementary machine-learning analyses

Across complementary machine-learning models evaluated using cross-validated out-of-fold procedures, discrimination for incident orientation impairment was modest. Elastic Net regression achieved the highest out-of-fold performance after calibration (AUC 0.71), followed by random forest (AUC 0.70) and ridge logistic regression (AUC 0.69). A stacking meta-learner did not materially improve discrimination.

Overall, machine-learning models did not materially improve discrimination compared with traditional regression approaches (Fig. 6). Importantly, across models, educational attainment consistently ranked among the most influential predictors, whereas stroke history contributed limited incremental predictive information, supporting the primary regression-based findings.

Boxplots display the distribution of cross-validated area under the receiver operating characteristic curve (AUC) for logistic regression (LR), Elastic Net, and random forest (RF) models. Each point represents the AUC from one cross-validation repeat with cluster-level grouping to prevent information leakage. Boxes indicate the interquartile range, horizontal lines denote medians, and whiskers represent observed ranges.

External validation

External validation was conducted using data from the Chinese Longitudinal Healthy Longevity Survey (CLHLS). In this independent cohort, discrimination was moderate, with a true out-of-fold AUC of 0.68.

Across repeated resampling, discrimination remained stable, indicating reasonable rank-order transportability.

Calibration assessment revealed systematic underestimation of absolute risk in the external cohort, as indicated by a positive calibration intercept. A calibration slope greater than 1 suggests that predicted risks were insufficiently dispersed relative to observed risks. These findings suggest that while relative risk patterns were broadly consistent across cohorts, recalibration would be required for accurate absolute risk estimation in different population settings (Fig. 7).

Observed incidence is plotted against mean predicted risk across deciles of predicted probability. Points represent group-wise observed risks with corresponding 95% confidence intervals. The dashed diagonal line indicates perfect calibration. Model discrimination and calibration metrics are shown within the figure, including the true out-of-fold AUC, calibration intercept, and calibration slope.

Discussion

In this nationally representative longitudinal cohort of older adults free of orientation impairment at baseline, we examined the association between stroke history and incident orientation impairment over a two-year follow-up, with particular attention to the potential modifying role of educational attainment. Three main findings emerged. First, educational attainment showed a strong, consistent, and robust protective association with incident orientation impairment across all analytical approaches. In contrast, age was not independently associated with incident orientation impairment after multivariable adjustment. This may reflect restriction of the analytic sample to participants free of baseline orientation impairment, the relatively narrow age distribution, and the short follow-up interval. Second, stroke history was not independently associated with incident orientation impairment after multivariable adjustment. Third, although interaction estimates were directionally consistent with attenuation of stroke-associated risk at higher education levels, confidence intervals were wide and crossed the null, and formal tests did not provide statistical evidence for effect modification.

Education as a strong and consistent protective factor

Across regression-based analyses, sensitivity analyses, competing-risk models, and complementary machine-learning approaches, educational attainment emerged as the most stable determinant of incident orientation impairment. Each additional three years of education was associated with approximately a 25% lower risk of developing orientation impairment over two years, a finding that was remarkably consistent across model specifications and outcome definitions.

These results align with the cognitive reserve framework, which posits that education enhances resilience to age-related and pathological brain changes through more efficient neural processing and compensatory mechanisms[22, 23]. Importantly, our findings extend prior evidence beyond global cognitive outcomes to orientation impairment, a clinically interpretable and functionally relevant cognitive domain that often represents an early manifestation of cognitive vulnerability in older adults.

The absence of strong evidence for non-linearity further supports the interpretation of education as a graded, cumulative protective factor rather than a threshold effect.

Non-agricultural hukou status was also associated with a lower risk of incident orientation impairment. This finding may reflect differential access to healthcare resources, social capital, and lifelong cognitive stimulation between urban and rural populations in China.

Stroke history and orientation impairment

Contrary to expectations based on studies of global cognitive decline and dementia, stroke history was not independently associated with incident orientation impairment in fully adjusted models. The point estimate below unity should not be interpreted as protective; rather, it likely reflects residual confounding, selection of relatively resilient stroke survivors, and limited statistical power. The marked imbalance in vascular comorbidities between stroke and non-stroke groups (e.g., hypertension SMD > 0.8) further underscores the potential for residual confounding. Several factors may explain this finding[24–26]. Stroke history in this study was based on self-reported physician diagnosis and did not capture information on stroke subtype, severity, lesion location, timing, or recurrence. Such heterogeneity may attenuate associations with specific cognitive domains[27–30].

In addition, the analytic cohort was restricted to individuals who were free of orientation impairment at baseline, thereby excluding those with early or persistent post-stroke cognitive deficits and potentially selecting a relatively resilient subgroup of stroke survivors. Adjustment for age, educational attainment, and vascular comorbidities may have further accounted for much of the excess cognitive risk associated with stroke history[31–33]. Notably, when a stricter outcome definition was applied, the point estimate for stroke history shifted above unity, although confidence intervals remained wide, suggesting limited statistical power rather than clear evidence of an effect.

Education as a potential modifier of stroke-related risk

A central aim of this study was to assess whether educational attainment modifies the association between stroke history and incident orientation impairment. Although the direction of interaction estimates and marginal risk patterns were consistent with a buffering effect of higher education, formal tests for multiplicative and additive interaction did not yield statistically conclusive evidence.

These findings should be interpreted cautiously. Detecting interaction effects typically requires substantially larger sample sizes than those needed to detect main effects, particularly when exposures and outcomes are relatively infrequent[34–36]. Moreover, the coarse measurement of stroke history and the short follow-up period may have limited our ability to detect subtle, domain-specific modification effects[37]. Rather than refuting the cognitive reserve hypothesis, our results suggest that any modifying effect of education on stroke-related orientation impairment is likely modest in magnitude or difficult to detect in population-based settings using available measures[38].

Insights from machine-learning and external validation analyses

Complementary machine-learning analyses provided additional context for interpreting the primary findings. Despite the application of multiple algorithms and calibration procedures, predictive discrimination for incident orientation impairment remained modest, and no model materially outperformed traditional regression approaches. Across models, educational attainment consistently ranked among the most influential predictors, whereas stroke history contributed limited incremental predictive information. These results reinforce the conclusion that education plays a central role in shaping short-term cognitive vulnerability at the population level[39–41].

External validation in an independent cohort demonstrated moderate discrimination and broadly consistent risk patterns, supporting the transportability of relative associations. However, miscalibration of absolute risk estimates highlights the need for recalibration when applying risk models across populations with different characteristics and measurement instruments[42]. Together, these analyses support the robustness of the main findings while underscoring the challenges of predicting incident cognitive outcomes in older adults.

Strengths and limitations

Key strengths of this study include its longitudinal design, nationally representative sample, explicit complete-case definition with respect to the outcome, transparent handling of baseline covariate missingness through indicator categories, focus on incident orientation impairment, rigorous handling of loss to follow-up, and the integration of multiple analytical frameworks. The consistent protective association of education across diverse methods strengthens confidence in the central conclusions.

Several limitations warrant consideration. Stroke history was measured crudely and lacked clinical detail, limiting etiological inference. The follow-up period was relatively short, potentially underestimating longer-term cognitive consequences of stroke. Orientation impairment represents only one dimension of cognitive function and may not capture subtler deficits in other domains. Finally, limited statistical power constrained definitive conclusions regarding effect modification.

Missingness for certain vascular comorbidities, particularly hypertension, was substantial and differential by stroke status, which may have introduced residual confounding despite the use of missing indicator categories.

Implications and future directions

In this population-based longitudinal study of older adults, educational attainment emerged as a consistent and robust protective factor against incident orientation impairment, whereas stroke history showed no independent association after adjustment for sociodemographic and vascular factors. Evidence for education modifying the stroke–orientation relationship was suggestive but inconclusive. These findings highlight the central role of educational gradients in late-life cognitive vulnerability and

underscore the complexity of identifying domain-specific cognitive consequences of stroke in community-dwelling populations.

Conclusion

In summary, higher education was strongly and consistently associated with a lower risk of incident orientation impairment over two years, whereas stroke history showed no independent association after adjustment. Evidence for education modifying the stroke–orientation relationship was suggestive but not conclusive. These findings highlight the central role of educational gradients in cognitive aging and underscore the complexity of disentangling stroke-related cognitive outcomes in population-based studies.

Declarations

Ethics approval and consent to participate

The China Health and Retirement Longitudinal Study (CHARLS) obtained ethical approval from the Institutional Review Board of Peking University (IRB00001052-11015), and all participants provided written informed consent prior to participation.

The Chinese Longitudinal Healthy Longevity Survey (CLHLS) was approved by the Biomedical Ethics Committee of Peking University (IRB00001052-13074), and written informed consent was obtained from all participants or their legal representatives.

The present study is a secondary analysis of publicly available, de-identified data from both CHARLS and CLHLS and therefore did not require additional ethical approval from the authors' institutions. Both studies were conducted in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable. This study used de-identified secondary data and does not contain any individual person's identifiable information.

Availability of data and materials

The datasets analyzed during the current study are publicly available.

CHARLS data can be obtained upon registration through the official CHARLS data portal.

CLHLS data are available from the CLHLS project upon application and approval via the official CLHLS website.

All analytic code used to generate the results is available from the corresponding author upon reasonable request.

Funding

This study was supported by the Letter of the Department of Personnel and Education of the National Administration of Traditional Chinese Medicine (2022) No. 256; the 9th Anhui Provincial Special Support Program for Talents (2023) No. 35; the Anhui Provincial Natural Science Foundation (Grant No. 2308085MH297); the Key Research Project of Natural Sciences in Anhui Universities (Grant No. 2023AH040099); the Special Project for Clinical Medical Research and Transformation of Anhui Province (Grant No. 202427b10020060); the Scientific Research Project of Anhui Provincial Department of Education (Grant No. 2025AHGXZK40506).

Authors' contributions

Ce Shi and Fei Li conceived the research question and designed the study. Ce Shi led data curation, constructed analytic variables, developed and validated the statistical code, conducted the primary analyses and sensitivity analyses, generated the figures and tables, and drafted the initial manuscript. Qiqi Yang and Tianxin Jiang assisted with variable construction, data harmonization, and verification of analytic datasets; they also contributed to interpretation of findings and preparation of supplementary materials. Lihua Wu provided methodological guidance on model specification, interaction assessment, and robustness analyses, and critically revised the manuscript for important intellectual content. Fei Wang and Xiang Shang contributed to the clinical and public health interpretation of results and assisted with manuscript revision. Baoguo Wang supported data checking and interpretation, and provided input on presentation of results. Jianhong Gao, Weiran Li, and Ziyu Ye contributed to methodological input, literature contextualization, and critical revision of the manuscript. Fei Li served as the corresponding author, provided overall supervision, project administration, and oversight of manuscript development. All authors reviewed and approved the final manuscript and agree to be accountable for all aspects of the work.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We thank the research teams and participants of both the CHARLS and CLHLS studies for providing open-access data that made this work possible.

References

1. Tsoy E, Kiekhofer RE, Guterman EL, Tee BL, Windon CC, Dorsman KA, Lanata SC, Rabinovici GD, Miller BL, Kind AJH, et al. Assessment of Racial/Ethnic Disparities in Timeliness and Comprehensiveness of Dementia Diagnosis in California. JAMA Neurol. 2021;78(6):657–65.

2. Kohler IV, Kämpfen F, Bandawe C, Kohler HP. Cognition and Cognitive Changes in a Low-Income Sub-Saharan African Aging Population. J Alzheimer's disease: JAD. 2023;95(1):195–212.
3. Fukuda H, Kanzaki H, Murata F, Maeda M, Ikeda M. Disease Burden and Progression in Patients with New-Onset Mild Cognitive Impairment and Alzheimer's Disease Identified from Japanese Claims Data: Evidence from the LIFE Study. J Alzheimer's disease: JAD. 2023;95(4):1559–72.
4. Guerrero-Berroa E, Luo X, Schmeidler J, Rapp MA, Dahlman K, Grossman HT, Haroutunian V, Beeri MS. The MMSE orientation for time domain is a strong predictor of subsequent cognitive decline in the elderly. Int J Geriatr Psychiatry. 2009;24(12):1429–37.
5. Tian R, Jiang Y, Zhang Y, Yan X, Zhou Y, Chen D. Cognitive training program improves cognitive ability and daily living ability in elderly patients with mild cognitive impairment. Aging Clin Exp Res. 2022;34(5):997–1005.
6. Fauth EB, Schwartz S, Tschanz JT, Østbye T, Corcoran C, Norton MC. Baseline disability in activities of daily living predicts dementia risk even after controlling for baseline global cognitive ability and depressive symptoms. Int J Geriatr Psychiatry. 2013;28(6):597–606.
7. Bangen KJ, Weigand AJ, Thomas KR, Delano-Wood L, Clark LR, Eppig J, Werhane ML, Edmonds EC, Bondi MW. Cognitive dispersion is a sensitive marker for early neurodegenerative changes and functional decline in nondemented older adults. Neuropsychology. 2019;33(5):599–608.
8. Wilcock D, Jicha G, Blacker D, Albert MS, D'Orazio LM, Elahi FM, Fornage M, Hinman JD, Knoefel J, Kramer J, et al. MarkVCID cerebral small vessel consortium: I. Enrollment, clinical, fluid protocols. Alzheimer's Dement J Alzheimer's Assoc. 2021;17(4):704–15.
9. Zhi N, Zhang L, Wang Y, Bai S, Geng J, Yu L, Cao W, Zhuang L, Zhou Y, Guan Y. Modified cerebral small vessel disease score is associated with vascular cognitive impairment after lacunar stroke. Aging. 2021;13(7):9510–21.
10. Zhao J, Jin X, Chen B, Fu C, Ji S, Shen W, Wei J, Zheng H, Zhang Y. Apathy symptoms increase the risk of dementia conversion: a case-matching cohort study on patients with post-stroke mild cognitive impairment in China. Psychogeriatrics: official J Japanese Psychogeriatr Soc. 2021;21(2):149–57.
11. Abdullah AH, Sharip S, Rahman AHA, Bakar L. Cognitive reserve in stroke patients. PsyCh J. 2021;10(3):444–52.
12. Umarova RM, Schumacher LV, Schmidt CSM, Martin M, Egger K, Urbach H, Hennig J, Klöppel S, Kaller CP. Interaction between cognitive reserve and age moderates effect of lesion load on stroke outcome. Sci Rep. 2021;11(1):4478.
13. Mungas D, Fletcher E, Gavett BE, Widaman K, Zahodne LB, Hohman TJ, Mayeda ER, Dowling NM, Johnson DK, Tomaszewski Farias S. Comparison of Education and Episodic Memory as Modifiers of Brain Atrophy Effects on Cognitive Decline: Implications for Measuring Cognitive Reserve. J Int Neuropsychological Society: JINS. 2021;27(5):401–11.
14. Albrecht P, Perisse J, Sauleau EA, Blanc F. Cognitive reserve as one of the determinants of the progression of Alzheimer disease. *Geriatrie et psychologie neuropsychiatrie du vieillissement* 2022.

15. Choi B, Choi J, Lee J, Choi E, Jang S-N. Socioeconomic disparities in hypertension and diabetes among older adults in South Korea: a trend analysis from a national survey, 2011–2023. Soc Sci Med. 2025;382:118350.
16. Jensen HAR, Davidsen M, Brønnum-Hansen H, Eliasen MH, Christensen AI. Trends in educational inequality in healthy life expectancy in Denmark between 2010 and 2021: a population-based study. BMJ Open. 2025;15(7):e100851.
17. Niu J, Li X, Chen Q, Yang W, Suo L, Chen Z. Association Between Educational Inequality and Income Inequality With Metabolic Diseases and Cause-Specific Mortality. Clin Cardiol. 2025;48(7):e70173.
18. Langa KM, Ryan LH, McCammon RJ, Jones RN, Manly JJ, Levine DA, Sonnega A, Farron M, Weir DR. The Health and Retirement Study Harmonized Cognitive Assessment Protocol Project: Study Design and Methods. Neuroepidemiology. 2020;54(1):64–74.
19. Zhang M, Lv X, Chen Y, Tu L, Fan Z, Yao Y, Yu X, Guan N, Wang H. Excessive sleep increased the risk of incidence of cognitive impairment among older Chinese adults: a cohort study based on the Chinese Longitudinal Healthy Longevity Survey (CLHLS). Int Psychogeriatr. 2022;34(8):725–34.
20. Nieboer D, van der Ploeg T, Steyerberg EW. Assessing Discriminative Performance at External Validation of Clinical Prediction Models. PLoS ONE. 2016;11(2):e0148820.
21. Jonsson H, Nyström L, Blom J. Poisson regression with adjustment for contamination and non-compliance in cohort studies conducted to estimate intervention effectiveness. J Med Screen 2025:9691413251388380.
22. Beker N, Ganz A, Hulsman M, Klausch T, Schmand BA, Scheltens P, Sikkes SAM, Holstege H. Association of Cognitive Function Trajectories in Centenarians With Postmortem Neuropathology, Physical Health, and Other Risk Factors for Cognitive Decline. JAMA Netw open. 2021;4(1):e2031654.
23. Hu D, Liu C, Xia K, Abramowitz A, Wu G. Characterizing the Resilience Effect of Neurodegeneration for the Mechanistic Pathway of Alzheimer's Disease. J Alzheimer's disease: JAD. 2021;84(3):1351–62.
24. Delgado J, Masoli J, Hase Y, Akinyemi R, Ballard C, Kalaria RN, Allan LM. Trajectories of cognitive change following stroke: stepwise decline towards dementia in the elderly. Brain Commun. 2022;4(3):fcac129.
25. Lo JW, Crawford JD, Lipnicki DM, Lipton RB, Katz MJ, Preux PM, Guerchet M, d'Orsi E, Quialheiro A, Rech CR, et al. Trajectory of Cognitive Decline Before and After Stroke in 14 Population Cohorts. JAMA Netw open. 2024;7(10):e2437133.
26. Chen LY, Norby FL, Gottesman RF, Mosley TH, Soliman EZ, Agarwal SK, Loehr LR, Folsom AR, Coresh J, Alonso A. Association of Atrial Fibrillation With Cognitive Decline and Dementia Over 20 Years: The ARIC-NCS (Atherosclerosis Risk in Communities Neurocognitive Study). J Am Heart Association 2018, 7(6).
27. Jamrozik E, Hyde Z, Alfonso H, Flicker L, Almeida O, Yeap B, Norman P, Hankey G, Jamrozik K. Validity of self-reported versus hospital-coded diagnosis of stroke: a cross-sectional and

longitudinal study. Cerebrovasc Dis. 2014;37(4):256–62.
28. Yan PJ, Hou LS, Li ME, Lu ZX, Zhan FY, Ran MD, Li JJ, Zhang L, Yang R, Zhou MK, et al. Associations between Lesion Locations and Stroke Recurrence in Survivors of First-ever Ischemic Stroke: A Prospective Cohort Study. Curr Med Sci. 2020;40(4):708–18.
29. Nakling AE, Aarsland D, Næss H, Wollschlaeger D, Fladby T, Hofstad H, Wehling E. Cognitive Deficits in Chronic Stroke Patients: Neuropsychological Assessment, Depression, and Self-Reports. Dement geriatric Cogn disorders extra. 2017;7(2):283–96.
30. Joundi RA, Fang J, Austin PC, Smith EE, Yu AYX, Hachinski V, Sposato LA, Ganesh A, Sharma M, Kapral MK. Magnitude and Time-Course of Dementia Risk in Stroke Survivors: A Population-Wide Matched Cohort Study. Neurology. 2025;104(1):e210131.
31. Weaver NA, Kuijf HJ, Aben HP, Abrigo J, Bae HJ, Barbay M, Best JG, Bordet R, Chappell FM, Chen C, et al. Strategic infarct locations for post-stroke cognitive impairment: a pooled analysis of individual patient data from 12 acute ischaemic stroke cohorts. Lancet Neurol. 2021;20(6):448–59.
32. Ohlmeier L, Nannoni S, Pallucca C, Brown RB, Loubiere L, Markus HS. Prevalence of, and risk factors for, cognitive impairment in lacunar stroke. Int J stroke: official J Int Stroke Soc. 2023;18(1):62–9.
33. Arce Rentería M, Manly JJ, Vonk JMJ, Mejia Arango S, Michaels Obregon A, Samper-Ternent R, Wong R, Barral S, Tosto G. Midlife Vascular Factors and Prevalence of Mild Cognitive Impairment in Late-Life in Mexico. J Int Neuropsychological Society: JINS. 2022;28(4):351–61.
34. Leon AC, Heo M. Sample Sizes Required to Detect Interactions between Two Binary Fixed-Effects in a Mixed-Effects Linear Regression Model. Comput Stat Data Anal. 2009;53(3):603–8.
35. Heo M, Leon AC. Sample sizes required to detect two-way and three-way interactions involving slope differences in mixed-effects linear models. J Biopharm Stat. 2010;20(4):787–802.
36. Yang C, Berkalieva A, Mazumdar M, Kwon D. Power calculation for detecting interaction effect in cross-sectional stepped-wedge cluster randomized trials: an important tool for disparity research. BMC Med Res Methodol. 2024;24(1):57.
37. Oh H, Park J, Seo W. A 2-year prospective follow-up study of temporal changes associated with post-stroke cognitive impairment. Int J Nurs Pract. 2018;24(2):e12618.
38. Kessels RP, Eikelboom WS, Schaapsmeerders P, Maaijwee NA, Arntz RM, van Dijk EJ, de Leeuw FE. Effect of Formal Education on Vascular Cognitive Impairment after Stroke: A Meta-analysis and Study in Young-Stroke Patients. J Int Neuropsychological Society: JINS. 2017;23(3):223–38.
39. Roldán-Tapia MD, Cánovas R, León I, García-Garcia J. Cognitive Vulnerability in Aging May Be Modulated by Education and Reserve in Healthy People. Front Aging Neurosci. 2017;9:340.
40. Mantri S, Nwadiogbu C, Fitts W, Dahodwala N. Quality of education impacts late-life cognition. Int J Geriatr Psychiatry. 2019;34(6):855–62.
41. Ho S, Kozhevnikov M. Cognitive style and creativity: The role of education in shaping cognitive style profiles and creativity of adolescents. Br J Educ Psychol. 2023;93(4):978–96.

42. Pfeiffer RM, Petracci E. Variance computations for functional of absolute risk estimates. Stat Probab Lett. 2011;81(7):807–12.

Figures

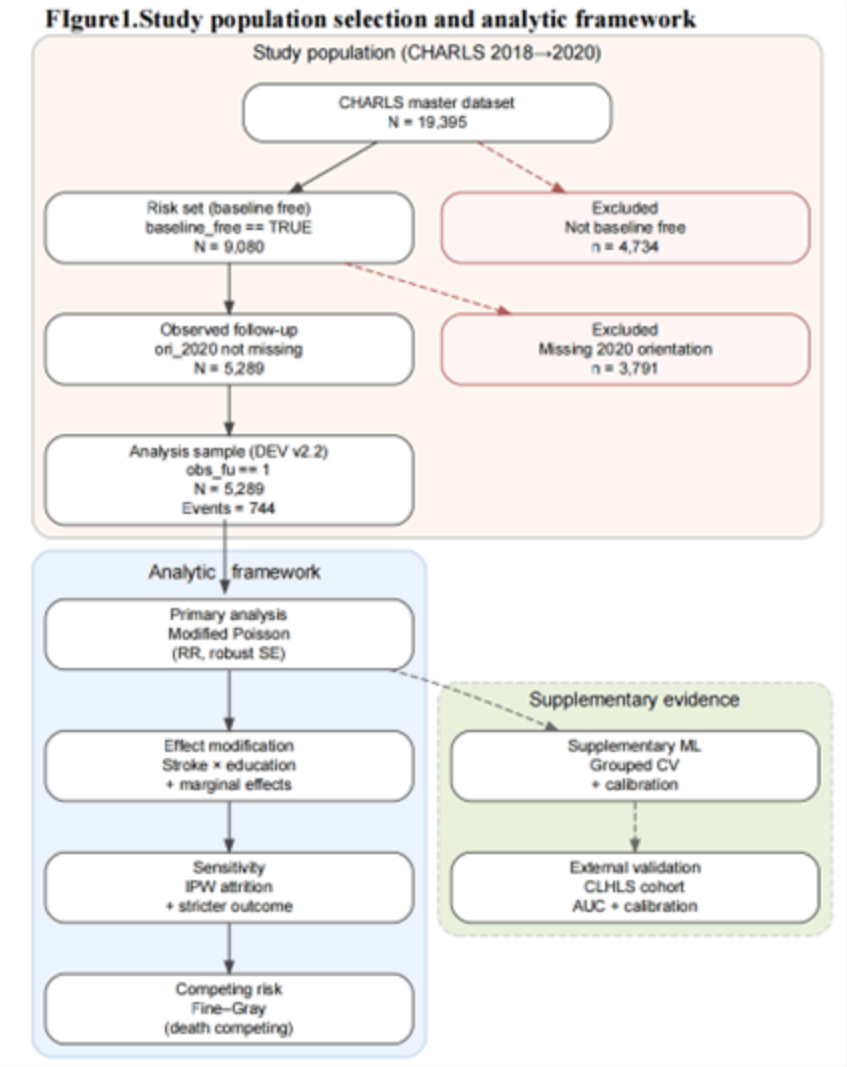


Figure 1

Flowchart of cohort derivation and analytic sample selection (CHARLS 2018–2020).

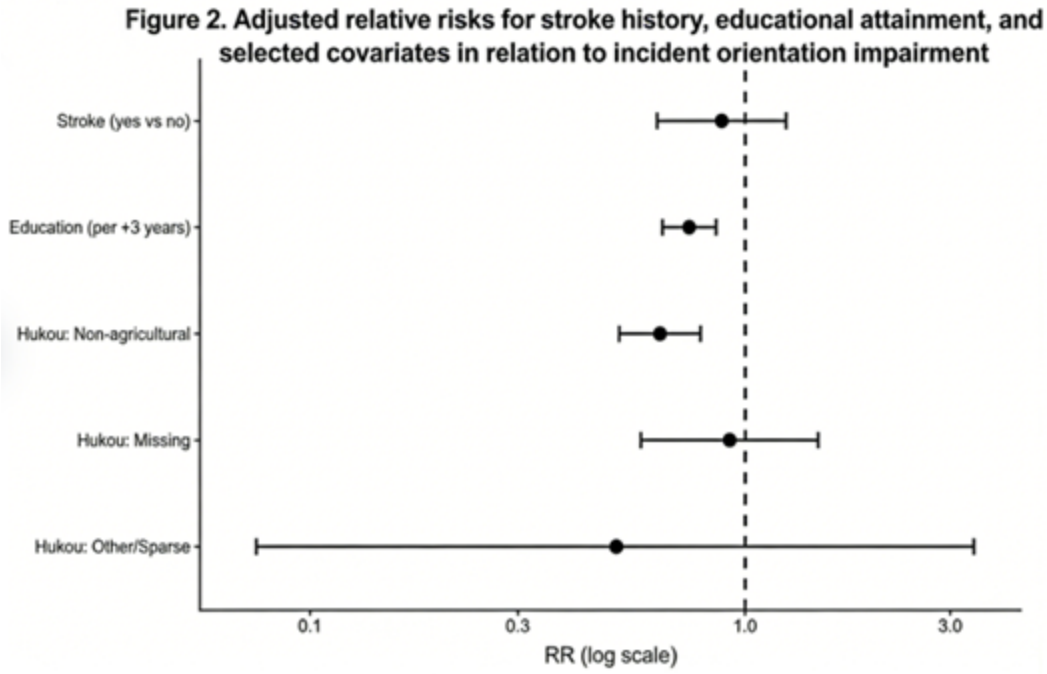


Figure 2

Adjusted relative risks for stroke history, educational attainment, and selected covariates in relation to incident orientation impairment.

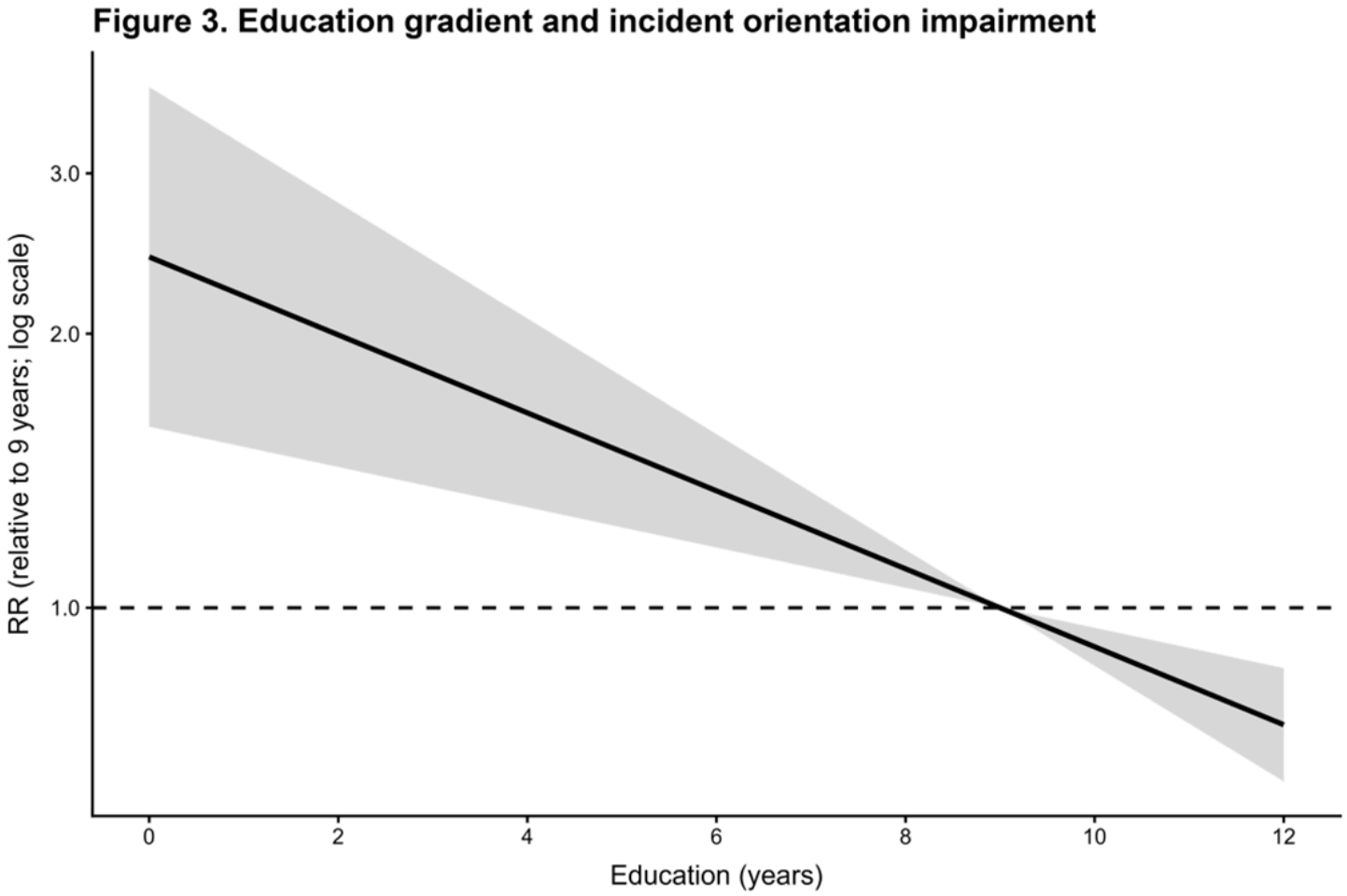


Figure 3

Education gradient and incident orientation impairment.

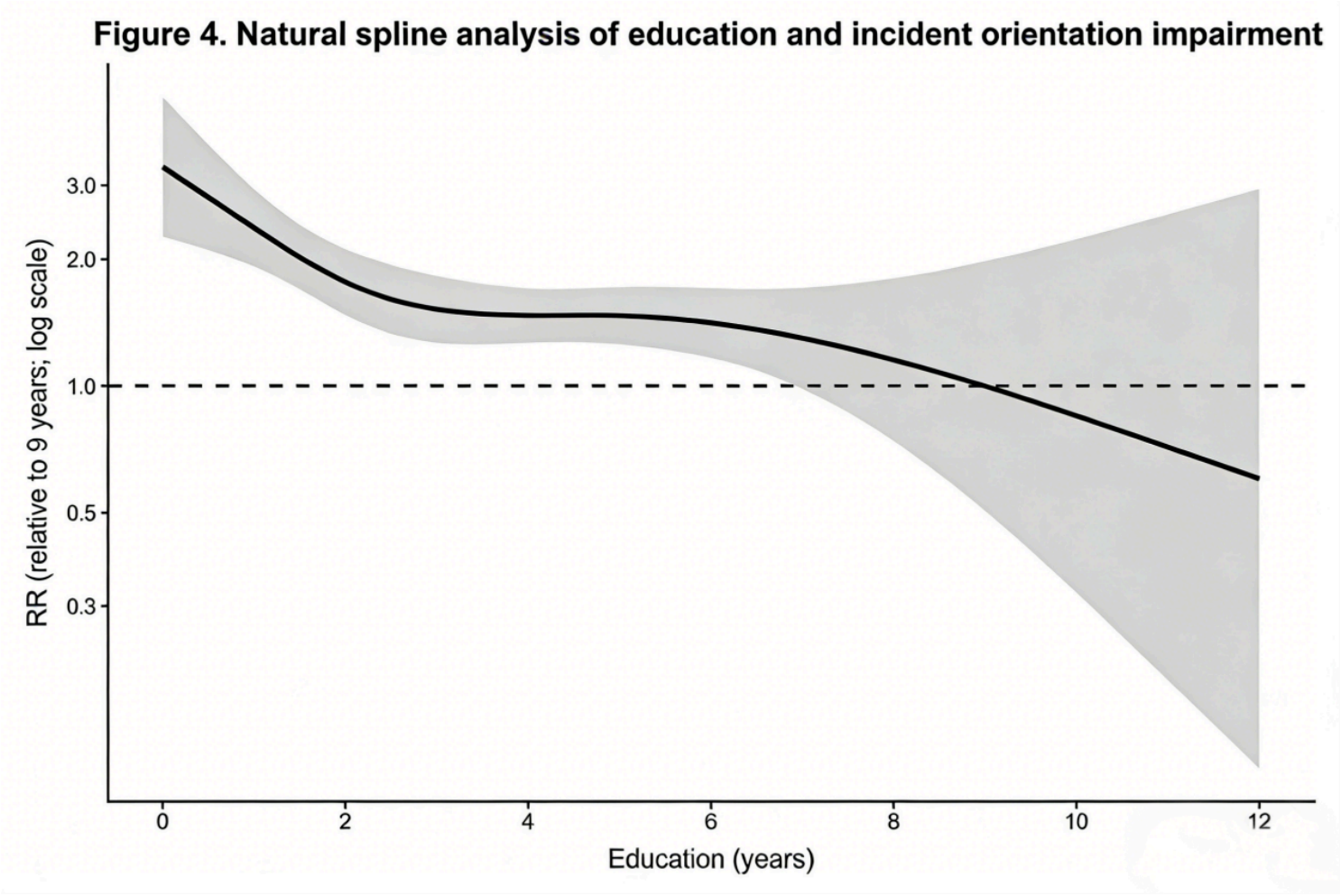


Figure 4

Natural spline analysis of education and incident orientation impairment.

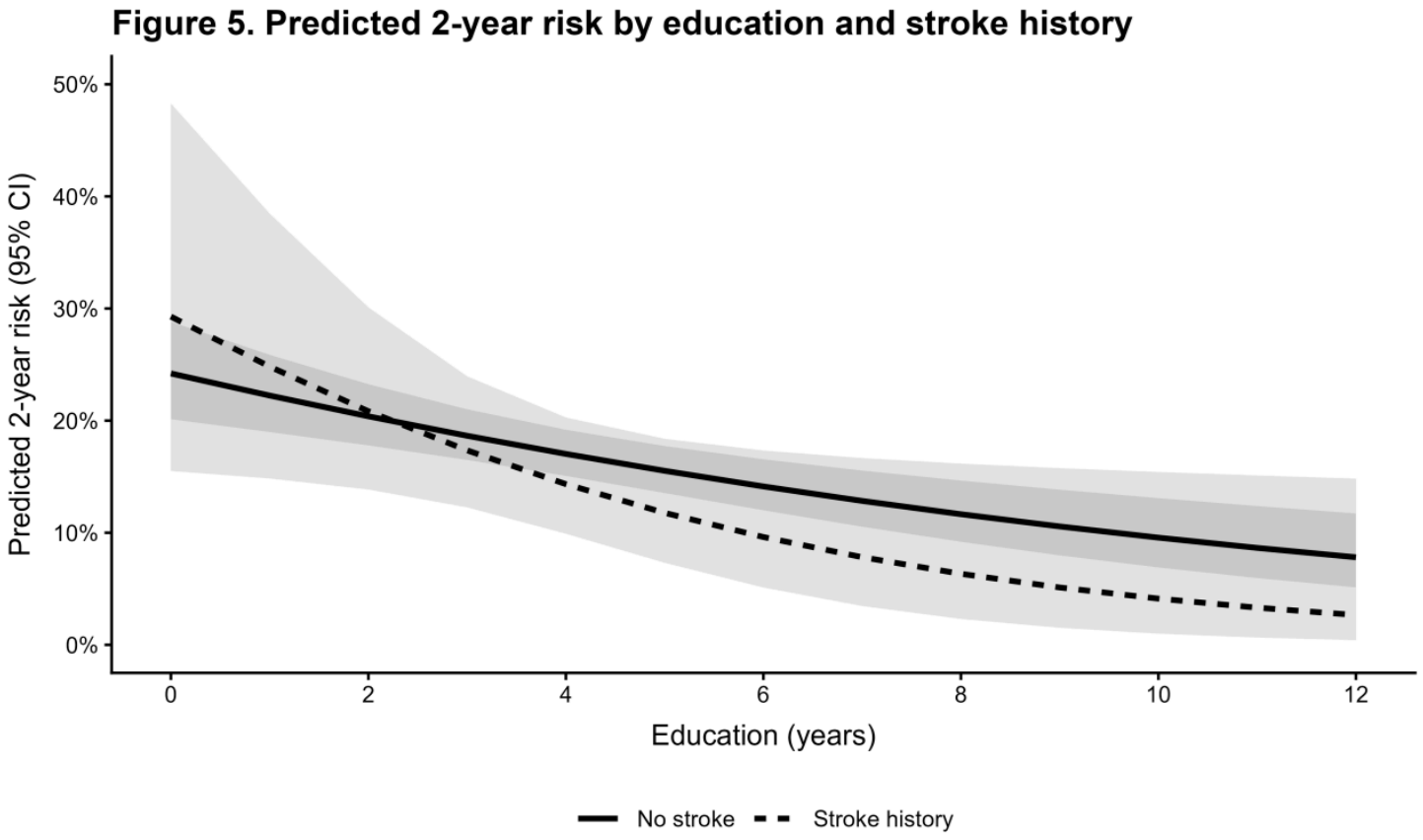


Figure 5

Predicted risk of incident orientation impairment by stroke status and educational attainment.

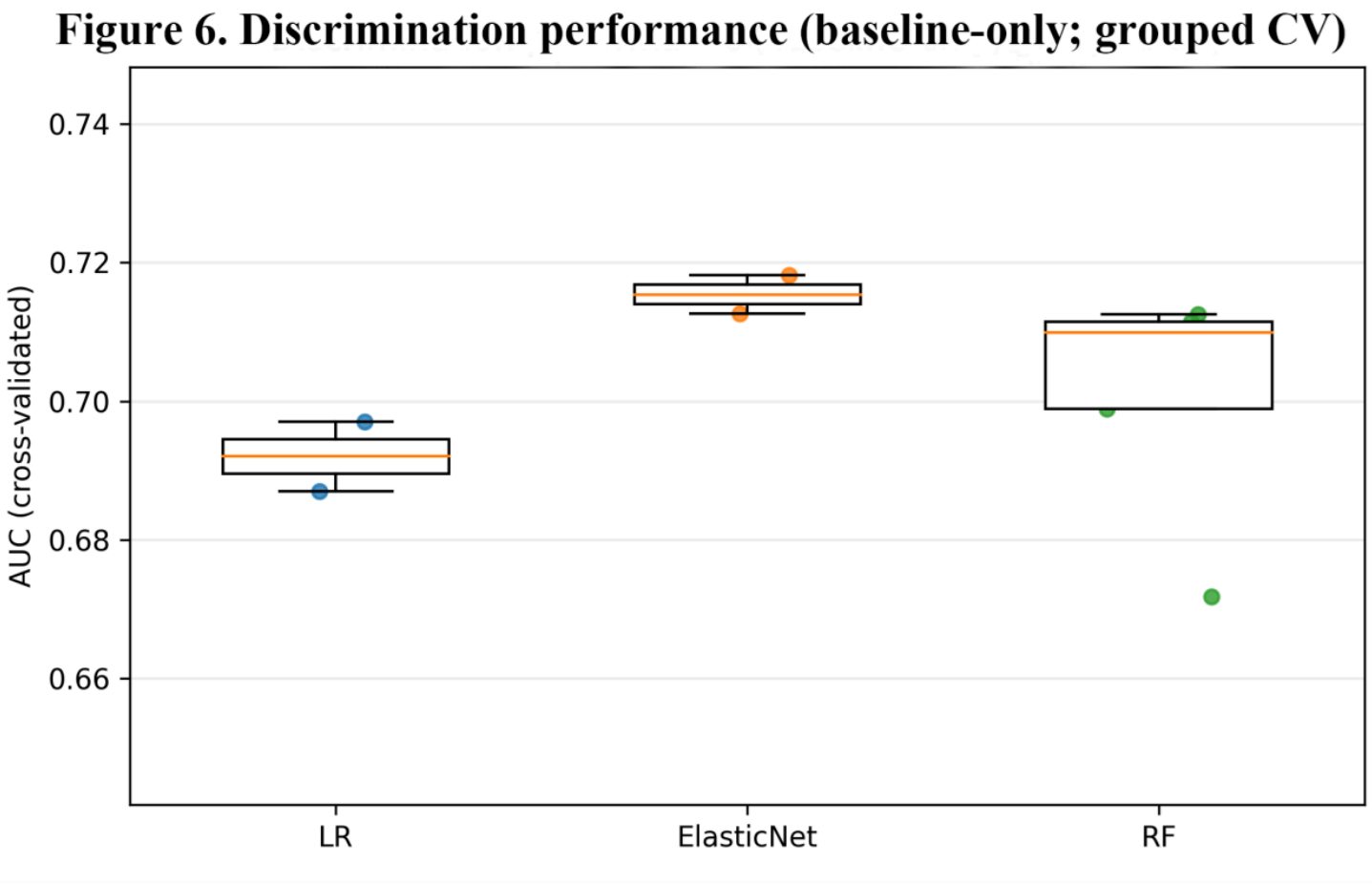


Figure 6

Cross-validated discrimination of machine-learning models.

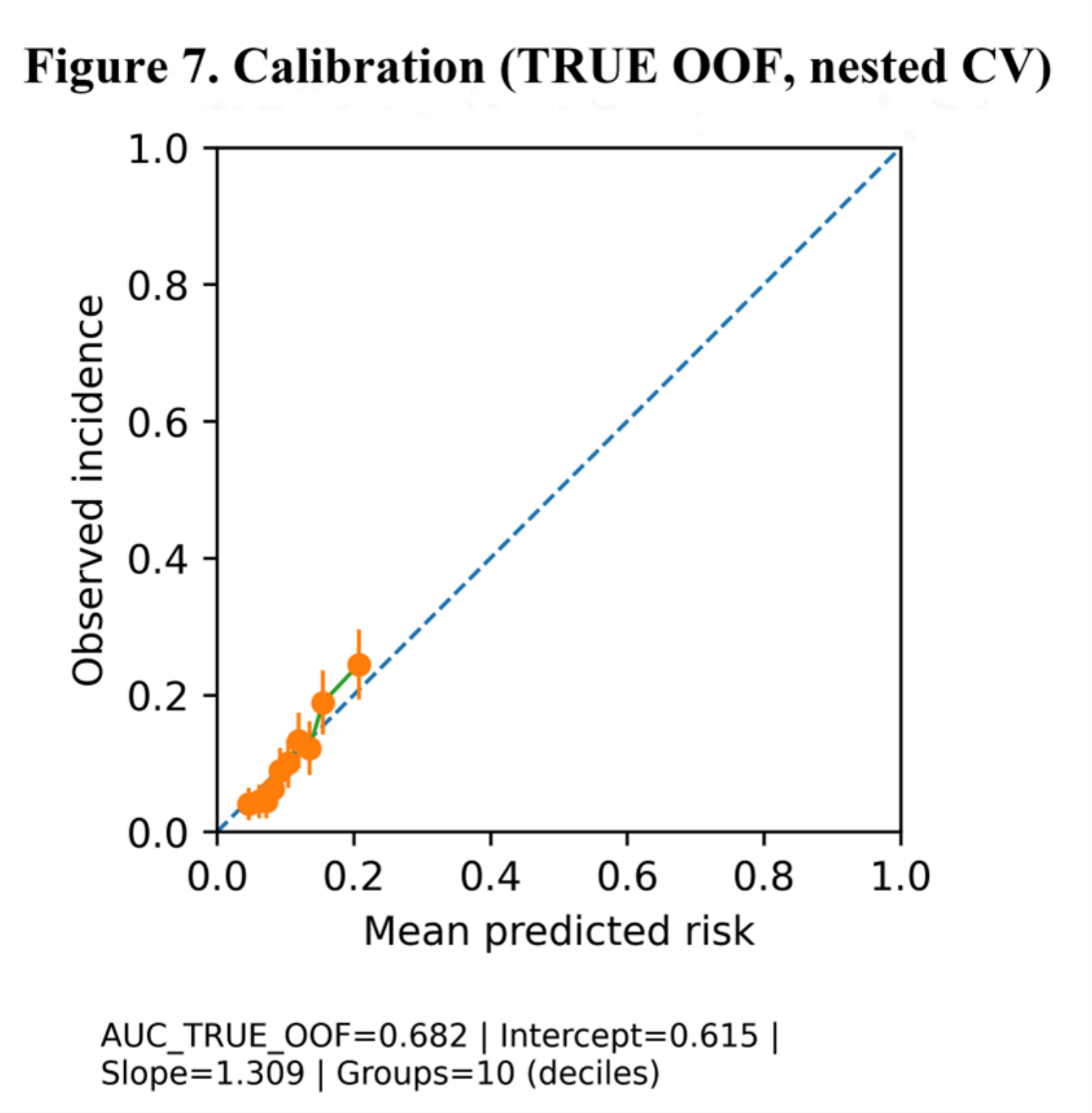


Figure 7

External validation calibration plot (CLHLS cohort).