

Hemoglobin glycation index and incident stroke in middle-aged and older Chinese adults: a prospective cohort study based on the China Health and Retirement Longitudinal Study

Miao Peng

People's Hospital of Deyang City

Ni He

People's Hospital of Deyang City

Quanhua Meng

People's Hospital of Deyang City

Sixu Liu

People's Hospital of Deyang City

Jing Tian

jingscorpion@163.com

People's Hospital of Deyang City

Research Article

Keywords: Hemoglobin glycation index, Glycated hemoglobin, Stroke, Prospective cohort study, China Health and Retirement Longitudinal Study, Competing risk

Posted Date: October 1st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11014404/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

Glycated hemoglobin (HbA1c) reflects long-term average glycemia, yet stable inter-individual differences in hemoglobin glycation exist. The hemoglobin glycation index (HGI) quantifies the gap between measured HbA1c and the HbA1c predicted from blood glucose. Whether HGI predicts incident stroke in the general population, particularly independently of HbA1c, remains unclear.

Methods

Within the China Health and Retirement Longitudinal Study, 10,643 stroke-free participants aged 45 years or older with baseline (2011) fasting plasma glucose and HbA1c were followed to 2020. The outcome was incident self-reported physician-diagnosed stroke. HGI equaled measured HbA1c minus the value predicted from a baseline-population regression of HbA1c on fasting plasma glucose; the primary exposure was high HGI (at or above the 90th percentile). We fitted sequentially adjusted Cox proportional hazards models, including a prespecified HbA1c-adjusted key test; restricted cubic splines assessed the dose–response relationship; a Fine–Gray model addressed the competing risk of death; and discrimination was compared against the stress hyperglycemia ratio, triglyceride–glucose index, HbA1c, and fasting plasma glucose. Eight prespecified sensitivity analyses were performed.

Results

Over a median follow-up of about 9 years, 841 incident strokes occurred (7.9%). In the complete-case main analysis ($n = 8,684$; 699 events), high HGI was associated with increased stroke risk (adjusted hazard ratio 1.392, 95% confidence interval 1.122–1.728, $P = 0.003$) and persisted after HbA1c adjustment (hazard ratio 1.333, 1.018–1.746, $P = 0.037$), meeting the prespecified criterion. Nonlinearity was suggested overall ($P = 0.060$) and significant among participants without diabetes ($P = 0.017$). Competing-risk results were consistent (subdistribution hazard ratio 1.386, 1.119–1.716). The association was stronger among participants without diabetes (hazard ratio 1.478, 1.118–1.954). All sensitivity analyses were directionally consistent, and no elevated risk at low HGI was observed after excluding anemia. Discrimination did not differ significantly among the five indices (C-index 0.658–0.663).

Conclusions

High HGI was independently associated with incident stroke in middle-aged and older Chinese adults, beyond HbA1c and more pronounced in those without diabetes. As a trait-like marker of glycation

propensity computable from routine measurements, HGI may help identify individuals at elevated stroke risk whom HbA1c does not flag, although its stand-alone predictive value is limited.

Introduction

Stroke is a leading cause of death and disability among Chinese adults. According to the Global Burden of Disease Study, more than 100 million people worldwide were living with stroke in 2019, and China bears a particularly heavy burden, ranking among the highest in incidence, prevalence, and disease burden [1, 2]. Diabetes and dysglycemia are established modifiable risk factors for stroke: meta-analyses indicate that individuals with diabetes have a two- to four-fold higher stroke risk than those without, and even prediabetic states are associated with elevated vascular risk [3, 4]. Accurate identification of glycemia-related high-risk individuals in the general population is therefore of considerable importance for the primary prevention of stroke.

Glycated hemoglobin (HbA1c) reflects average blood glucose over the preceding 8–12 weeks and is widely used for diabetes diagnosis and glycemic management; elevated HbA1c is associated with increased cardiovascular risk even among individuals without diabetes [5]. However, HbA1c is determined not only by glycemia but also by biological factors such as the individual rate of hemoglobin glycation and red blood cell lifespan: at identical glucose levels, HbA1c can differ consistently and stably across individuals [6–8]. To quantify this inter-individual variation, Hempe et al. introduced the hemoglobin glycation index (HGI) in 2002, defined as the difference between the measured HbA1c and the HbA1c predicted from blood glucose, thereby capturing an individual's "glycation propensity" phenotype [6].

Previous studies have linked high HGI to adverse outcomes. In type 1 diabetes, high HGI predicted retinopathy and nephropathy [9]. In type 2 diabetes, secondary analyses of the ACCORD, ADVANCE, and DEVOTE trials showed that high HGI was associated with cardiovascular events, mortality, or differential benefit from intensive glucose-lowering treatment [10–12]; however, a 15-year prospective cohort found no predictive advantage of HGI over HbA1c for cardiovascular, microvascular, or mortality outcomes [13], whereas in patients with coronary heart disease and type 2 diabetes, high HGI remained predictive of adverse cardiovascular events [14]. Beyond overt diabetes, high HGI has been associated with subclinical myocardial injury in the general population [15], carotid atherosclerosis and insulin resistance among non-diabetic individuals [16], coronary artery calcification [17], and kidney injury in patients undergoing coronary angiography [18]. For stroke specifically, prior HGI studies have largely addressed prognosis after acute ischemic stroke rather than incident events [19]. The Northeast China Rural Cardiovascular Health Study (NCRCHS) was the first to combine diabetes status with HGI to evaluate incident stroke in a regional rural population, reporting an approximately 70% higher stroke risk among participants with isolated high HGI (without diabetes) [20]; the REACTION cohort likewise reported an association between HGI and 5-year major adverse cardiovascular events [21]. A recent CHARLS-based study suggested a U-shaped association between HGI and mild cognitive impairment in middle-aged and older adults [22].

Nevertheless, three gaps remain. First, whether HGI provides risk information independent of HbA1c is debated: in populations with diabetes (DEVOTE and ADVANCE), HGI was no longer significant once HbA1c entered the model [11, 12], whereas in the general population—where the HbA1c distribution is narrow—the relative contribution of glycation propensity may be larger, and this question awaits direct testing in a nationally representative cohort. Second, prospective evidence in the general population derives solely from a regional cohort with a limited number of events (NCRCHS: 79 events), and death as an important competing risk was not addressed. Third, a U/J-shaped signal of elevated risk at the low-HGI end has been observed, but its origin—such as spuriously low HbA1c caused by anemia producing an apparent association—has not been specifically examined.

We therefore used the nationally representative CHARLS 2011–2020 cohort to systematically evaluate the association between baseline HGI and incident stroke among stroke-free adults aged 45 years or older. The HbA1c-adjusted model was prespecified as the key test, complemented by dose–response analysis, a Fine–Gray competing-risk model, subgroup and interaction analyses, and eight prespecified sensitivity analyses. We aimed to answer three questions: whether high HGI predicts incident stroke independently of HbA1c; whether the association holds among individuals without diabetes; and how HGI is positioned relative to existing glycemia-related indices (SHR, TyG, HbA1c, and FPG).

Methods

Study design and population

CHARLS is a nationwide prospective cohort conducted by Peking University, covering residents aged 45 years or older in 28 Chinese provinces. The national baseline survey ($n = 17,705$) was completed in 2011–2012, followed by four waves in 2013, 2015, 2018, and 2020; venous blood samples were collected in 2011 and 2015. Details of the study design and quality control have been published in the cohort profile [23]. This study is reported in accordance with the STROBE statement [24].

The inclusion and exclusion flow is shown in Fig. 1. Among the 17,705 baseline respondents, we sequentially excluded participants without 2011 blood measurements or with missing FPG/HbA1c ($n = 6,173$), those aged < 45 years ($n = 227$), those with self-reported prevalent stroke at baseline ($n = 267$), those lost to all follow-up waves ($n = 387$), and those with missing key covariates ($n = 8$), yielding a full analytic sample of 10,643 participants. The complete-case main analysis further required non-missing values for all model covariates, resulting in 8,684 participants. Missing covariate data were dominated by systolic blood pressure ($n = 1,579$), and event rates were similar between excluded and included participants (7.25% vs 8.05%), suggesting limited selection bias (a multiple-imputation sensitivity analysis is described below).

Exposure measurement: hemoglobin glycation index

FPG and HbA1c were measured in fasting venous blood at the 2011 baseline by the CHARLS central laboratory team. A linear regression of HbA1c on FPG in the baseline analytic population yielded the prediction equation: predicted HbA1c = $3.609 + 0.0151 \times \text{FPG (mg/dL)}$ ($R^2 = 0.44$). HGI was defined as measured HbA1c minus predicted HbA1c [6, 25]. The primary exposure was high HGI, defined as $\text{HGI} \geq$ the 90th percentile of the full population ($\text{P90} = 0.574$); a secondary cutoff of $\text{HGI} \geq 0.5$ was also examined. HGI was additionally modeled as a continuous variable using restricted cubic splines.

To assess the influence of equation derivation, two sensitivity analyses were prespecified: S1 refit the regression equation in the baseline population without diabetes, and S2 applied the standardized equation proposed by Hempe et al. based on the NHANES non-medicated population (predicted HbA1c = $3.1 + 0.024 \times \text{FPG [mg/dL]}$) [25, 26].

Outcome ascertainment

The outcome was incident stroke during follow-up, defined as a self-reported physician diagnosis of stroke at waves 2–5 (2013/2015/2018/2020) among participants without stroke at baseline. Event time was assigned to the midpoint between the wave at which stroke was first reported and the preceding wave; participants without events were censored at their last attended wave, with follow-up ending in 2020. Information on all-cause death was obtained from the 2013 Exit Interview and the 2020 Exit Module (1,169 deaths in total); for the 663 deaths with a missing death year, the death time was imputed as the midpoint between the last wave at which the participant was observed alive and 2020, according to the prespecified rule (sensitivity analysis S4 examined this imputation). Death was treated as a competing event in the Fine–Gray analysis.

Covariates

All covariates were measured at the 2011 baseline: age (years, continuous), sex, urban/rural residence (classified according to the National Bureau of Statistics community attribute codes), educational attainment, self-reported physician-diagnosed hypertension, diabetes, dyslipidemia, and heart disease, systolic blood pressure (mmHg), C-reactive protein, triglycerides, and high-density lipoprotein cholesterol. In the adjusted models, diabetes was entered as self-reported physician-diagnosed diabetes. For the subgroup analysis, a broader definition was used: self-reported diagnosis, use of glucose-lowering medication, or meeting glycemic criteria at baseline ($\text{HbA1c} \geq 6.5\%$ or fasting plasma glucose ≥ 126 mg/dL). The longitudinal change in HGI between 2011 and 2015 (ΔHGI) was additionally constructed for sensitivity analysis S6, in which the outcome window was strictly restricted to incident strokes occurring after 2015 and baseline HGI was adjusted for.

Statistical analysis

The analysis plan, including all decision rules, was prespecified in a statistical analysis plan. The main analysis proceeded as follows.

(1) Main analysis: Cox proportional hazards models estimated hazard ratios (HRs) with 95% confidence intervals (CIs) for the association between high HGI ($\geq$ P90, and the secondary cutoff ≥ 0.5) and incident stroke. Covariates were adjusted sequentially: model 1 adjusted for age and sex; model 2 additionally for urban/rural residence and education; model 3 additionally for hypertension, diabetes, dyslipidemia, heart disease, systolic blood pressure, C-reactive protein, triglycerides, and high-density lipoprotein cholesterol. The proportional hazards assumption was assessed using scaled Schoenfeld residuals (Grambsch–Therneau test) [27]; for covariates violating the assumption, time-by-covariate interaction terms (time-varying coefficients) were added as a robustness check.

(2) Key test: HbA1c (continuous) was added to model 3 to examine whether the association of high HGI was independent of HbA1c. The prespecified decision rule was: $HR > 1.2$ with $P < 0.05$ indicates that the independent association holds.

(3) Dose–response: restricted cubic splines with four knots characterized the dose–response curve of continuous HGI against stroke risk, with HGI = 0 as the reference; the P value for nonlinearity is reported.

(4) Discrimination comparison: high HGI, SHR (= $FPG [mmol/L] / (1.59 \times HbA1c - 2.59)$) [28], TyG index (= $\ln[\text{triglycerides} \times FPG / 2]$) [29], HbA1c, or FPG was added separately to the same base model, and Harrell's C-indices [30] were compared; 500 bootstrap resamples estimated the 95%CI of the C-index difference (ΔC).

(5) Competing risk: a Fine–Gray subdistribution hazards model estimated the subdistribution HR of high HGI for stroke, with death as the competing event [31].

(6) Subgroups and interactions: analyses were stratified by diabetes status, sex, age ($< 65 / \geq 65$ years), and urban/rural residence, with multiplicative interaction terms testing effect modification.

(7) Prespecified sensitivity analyses: S1, non-diabetic-population equation; S2, NHANES standardized equation; S3, exclusion of baseline anemia (defined by sex-specific hemoglobin thresholds), with a symmetric test at the low-HGI end ($< P10$); S4, midpoint imputation of death timing; S5a, a model without systolic blood pressure ($n = 10,170$); S5b, multiple imputation by chained equations (mice, $m = 20$, 5 iterations, seed 20260903) [32]; S6, ΔHGI and incident stroke after 2015.

Two-sided $\alpha = 0.05$ was used throughout. All analyses were performed in R (version 4.x), primarily using the survival, cmprsk, mice, and tableone packages.

Ethics

CHARLS was approved by the Biomedical Ethics Review Committee of Peking University (IRB00001052-11015 for the questionnaire survey; IRB00001052-11014 for blood sample collection), and all

respondents provided written informed consent. The present study is a secondary analysis of publicly available de-identified data.

Results

Study population and baseline characteristics

Among the 10,643 participants in the full analytic sample, 841 incident strokes (7.9%) and 1,169 all-cause deaths were recorded during follow-up. The complete-case main analysis sample comprised 8,684 participants with 699 events (8.05%). Baseline characteristics by HGI group are shown in Table 1. Compared with the low-HGI group, participants with high HGI ($\geq$ P90) had higher HbA1c but similar FPG (inherent to the definition of HGI), and displayed a noteworthy feature—a lower prevalence of anemia (9.2% vs 13.5%)—which resonates with the mechanistic interpretation discussed below.

Table 1

Baseline characteristics of the study population by baseline hemoglobin glycation index (HGI) group

Characteristic	Overall (n = 8,684)	HGI < P90 (n = 7,832)	HGI ≥ P90 (n = 852)	P value	SMD
n	8,684	7,832	852		
Age, years, mean (SD)	59.35 (9.27)	59.30 (9.33)	59.83 (8.76)	0.110	0.059
Sex, n (%)				0.064	0.068
Male	4,017 (46.3)	3,649 (46.6)	368 (43.2)		
Female	4,667 (53.7)	4,183 (53.4)	484 (56.8)		
Residence, n (%)				0.005	0.102
Rural	5,684 (65.5)	5,164 (65.9)	520 (61.0)		
Urban	3,000 (34.5)	2,668 (34.1)	332 (39.0)		
Educational attainment, n (%) †				0.137	0.090
1 (no formal education)	2,518 (29.0)	2,251 (28.7)	267 (31.3)		
2 (did not finish primary school)	1,591 (18.3)	1,450 (18.5)	141 (16.5)		
3 (sishu/home schooling)	44 (0.5)	39 (0.5)	5 (0.6)		
4 (primary school)	1,961 (22.6)	1,770 (22.6)	191 (22.4)		
5 (middle school)	1,724 (19.9)	1,554 (19.8)	170 (20.0)		
6 (high school)	580 (6.7)	529 (6.8)	51 (6.0)		
7 (vocational school)	149 (1.7)	134 (1.7)	15 (1.8)		
8 (junior college)	89 (1.0)	80 (1.0)	9 (1.1)		
9 (bachelor's degree)	27 (0.3)	25 (0.3)	2 (0.2)		
10 (postgraduate degree)	1 (0.0)	0 (0.0)	1 (0.1)		
Data are mean (SD) or n (%). P values were derived from t tests (continuous variables) or chi-squared tests (categorical variables) comparing the two HGI groups. SMD, standardized mean difference; HbA1c, glycated hemoglobin; HGI, hemoglobin glycation index (P90 = 0.574); HDL, high-density lipoprotein.					
† Educational attainment coded according to the CHARLS questionnaire (1 = lowest, 10 = highest); category labels to be verified against the analysis coding.					
‡ "Self-reported diabetes" denotes a self-reported physician diagnosis (the covariate in the adjusted models); "diabetes including undiagnosed" additionally includes use of glucose-lowering medication or meeting glycemic criteria at baseline (HbA1c ≥ 6.5% or fasting plasma glucose ≥ 126 mg/dL), and was used for the subgroup analysis. Anemia was defined by sex-specific hemoglobin thresholds.					

Characteristic	Overall (n = 8,684)	HGI < P90 (n = 7,832)	HGI ≥ P90 (n = 852)	P value	SMD
Hypertension, n (%)				< 0.001	0.174
No	6,617 (76.2)	6,027 (77.0)	590 (69.2)		
Yes	2,067 (23.8)	1,805 (23.0)	262 (30.8)		
Self-reported diabetes, n (%) ‡				< 0.001	0.485
No	8,204 (94.5)	7,515 (96.0)	689 (80.9)		
Yes	480 (5.5)	317 (4.0)	163 (19.1)		
Dyslipidemia, n (%)				< 0.001	0.172
No	7,912 (91.1)	7,177 (91.6)	735 (86.3)		
Yes	772 (8.9)	655 (8.4)	117 (13.7)		
Heart disease, n (%)				< 0.001	0.129
No	7,682 (88.5)	6,962 (88.9)	720 (84.5)		
Yes	1,002 (11.5)	870 (11.1)	132 (15.5)		
Diabetes including undiagnosed, n (%) ‡				< 0.001	0.695
No	7,283 (83.9)	6,795 (86.8)	488 (57.3)		
Yes	1,401 (16.1)	1,037 (13.2)	364 (42.7)		
Systolic blood pressure, mmHg, mean (SD)	130.47 (21.55)	130.23 (21.52)	132.65 (21.75)	0.002	0.112
Body mass index, kg/m ² , mean (SD)	23.49 (3.87)	23.38 (3.81)	24.46 (4.19)	< 0.001	0.270
Fasting plasma glucose, mg/dL, mean (SD)	109.44 (35.08)	107.28 (30.17)	129.28 (61.17)	< 0.001	0.456
HbA1c, %, mean (SD)	5.26 (0.81)	5.10 (0.47)	6.72 (1.49)	< 0.001	1.460
Data are mean (SD) or n (%). P values were derived from t tests (continuous variables) or chi-squared tests (categorical variables) comparing the two HGI groups. SMD, standardized mean difference; HbA1c, glycated hemoglobin; HGI, hemoglobin glycation index (P90 = 0.574); HDL, high-density lipoprotein.					
† Educational attainment coded according to the CHARLS questionnaire (1 = lowest, 10 = highest); category labels to be verified against the analysis coding.					
‡ “Self-reported diabetes” denotes a self-reported physician diagnosis (the covariate in the adjusted models); “diabetes including undiagnosed” additionally includes use of glucose-lowering medication or meeting glycemic criteria at baseline (HbA1c ≥ 6.5% or fasting plasma glucose ≥ 126 mg/dL), and was used for the subgroup analysis. Anemia was defined by sex-specific hemoglobin thresholds.					

Characteristic	Overall (n = 8,684)	HGI < P90 (n = 7,832)	HGI ≥ P90 (n = 852)	P value	SMD
HGI, mean (SD)	0.00 (0.60)	-0.12 (0.41)	1.16 (0.77)	< 0.001	2.068
Hemoglobin, g/dL, mean (SD)	14.37 (2.21)	14.36 (2.23)	14.47 (2.01)	0.155	0.054
C-reactive protein, mg/L, mean (SD)	2.66 (7.23)	2.60 (7.17)	3.15 (7.76)	0.038	0.073
Triglycerides, mg/dL, mean (SD)	125.90 (75.44)	124.55 (74.63)	138.30 (81.55)	< 0.001	0.176
HDL cholesterol, mg/dL, mean (SD)	51.57 (15.11)	51.67 (15.05)	50.66 (15.65)	0.065	0.065
Anemia, n (%)				0.001	0.137
No	7,404 (86.9)	6,641 (86.5)	763 (90.8)		
Yes	1,112 (13.1)	1,035 (13.5)	77 (9.2)		
Incident stroke during follow-up, n (%)	699 (8.0)	596 (7.6)	103 (12.1)		
Data are mean (SD) or n (%). P values were derived from t tests (continuous variables) or chi-squared tests (categorical variables) comparing the two HGI groups. SMD, standardized mean difference; HbA1c, glycated hemoglobin; HGI, hemoglobin glycation index (P90 = 0.574); HDL, high-density lipoprotein.					
† Educational attainment coded according to the CHARLS questionnaire (1 = lowest, 10 = highest); category labels to be verified against the analysis coding.					
‡ “Self-reported diabetes” denotes a self-reported physician diagnosis (the covariate in the adjusted models); “diabetes including undiagnosed” additionally includes use of glucose-lowering medication or meeting glycemic criteria at baseline (HbA1c ≥ 6.5% or fasting plasma glucose ≥ 126 mg/dL), and was used for the subgroup analysis. Anemia was defined by sex-specific hemoglobin thresholds.					

Main analysis and the HbA1c-adjusted key test

High HGI was associated with an increased risk of incident stroke, remaining stable across adjustment levels (Table 2): in the fully adjusted model, HR = 1.392 (95%CI 1.122–1.728, P = 0.003); results were consistent for the secondary cutoff HGI ≥ 0.5 (HR = 1.316, 95%CI 1.074–1.613, P = 0.008). The proportional hazards test indicated that the exposure satisfied the assumption (P = 0.253); hypertension and systolic blood pressure showed minor violations, and adding time-varying coefficient terms yielded HR = 1.393, virtually identical to the main result.

Table 2
Association between high hemoglobin glycation index (HGI) and incident stroke in Cox proportional hazards models

Model	HR (95% CI)	P value	No. of participants	No. of events
Model 1 (unadjusted)	1.625 (1.319–2.004)	< 0.001	8,684	699
Model 2 (demographic adjustment)	1.615 (1.310–1.992)	< 0.001	8,684	699
Model 3 (fully adjusted)	1.392 (1.122–1.728)	0.003	8,684	699
Robustness: cutoff HGI ≥ 0.5 (fully adjusted)	1.316 (1.074–1.613)	0.008	8,684	699
Key test: P90 cutoff + HbA1c	1.333 (1.018–1.746)	0.037	8,684	699
Key test: 0.5 cutoff + HbA1c	1.241 (0.966–1.594)	0.091	8,684	699
HR, hazard ratio; CI, confidence interval; HbA1c, glycated hemoglobin. High HGI was defined as HGI at or above the 90th percentile (P90 = 0.574) unless otherwise indicated. Model 1 was unadjusted; model 2 adjusted for age and sex; model 3 additionally adjusted for urban/rural residence, educational attainment, hypertension, diabetes, dyslipidemia, heart disease, systolic blood pressure, C-reactive protein, triglycerides, and high-density lipoprotein cholesterol. Key tests additionally included HbA1c (continuous) in model 3.				

In the key test including HbA1c, the association of high HGI remained significant (HR = 1.333, 95%CI 1.018–1.746, $P = 0.037$), meeting the prespecified decision criterion (HR > 1.2 and $P < 0.05$), whereas HbA1c itself was not significantly associated with incident stroke in the same model (HR = 1.029). This indicates that, in this general-population sample, the glycation propensity captured by HGI provides risk information beyond HbA1c.

Dose–response relationship

Restricted cubic splines showed an overall positive dose–response between HGI and stroke risk, with a curve shape suggestive of nonlinearity (P for nonlinearity = 0.060, Fig. 2): an approximate plateau in the mid-range, rising risk at the high-HGI end, and a mild upward turn at the low-HGI end suggestive of a J shape. Among participants without diabetes, the nonlinearity test was statistically significant ($P = 0.017$).

Discrimination comparison with other glycemia-related indices

When each index was added separately to the same base model, the C-indices ranged from 0.6576 to 0.6631 (Fig. 3). Relative to the high-HGI model, the 95%CI of ΔC for the SHR, TyG, HbA1c, and FPG

models all crossed zero, indicating no statistically significant difference in discrimination. At the continuous level, each standard-deviation increase in TyG was associated with an HR of 1.308, suggesting that these indices capture overlapping rather than complementary dimensions of risk.

Competing-risk analysis

In the Fine–Gray model with death as the competing event, the subdistribution HR for high HGI was 1.386 (95%CI 1.119–1.716, $P = 0.003$), highly consistent with the main analysis, indicating that the primary result is not materially affected by competing-risk bias from death.

Subgroup analyses

Subgroup results are shown in Table 3 and Fig. 4. The association was stronger and statistically significant among participants without diabetes (HR = 1.478, 95%CI 1.118–1.954, $P = 0.006$) and directionally consistent but non-significant among those with diabetes (HR = 1.335, 95%CI 0.940–1.896); point estimates exceeded 1 across strata defined by sex, age, and urban/rural residence. None of the interaction tests was significant ($P = 0.552–0.908$), indicating a directionally consistent association across subgroups.

Table 3
Subgroup analyses of the association between high hemoglobin glycation index (HGI, $\geq$ P90) and incident stroke

Subgroup	Stratum	n	Events	HR (95% CI)	P value	P for interaction
Diabetes status	Non-diabetic	7,283	552	1.478 (1.118–1.954)	0.006	0.676
	Diabetic	1,401	147	1.335 (0.940–1.896)	0.107	
Sex	Male	4,017	323	1.284 (0.922–1.789)	0.139	0.741
	Female	4,667	376	1.414 (1.062–1.882)	0.018	
Age group	< 65 years	6,326	474	1.335 (1.027–1.736)	0.031	0.908
	≥ 65 years	2,358	225	1.463 (1.002–2.137)	0.049	
Residence	Rural	5,684	467	1.466 (1.123–1.914)	0.005	0.552
	Urban	3,000	232	1.294 (0.896–1.869)	0.170	
HRs were estimated within each stratum using the fully adjusted Cox model (model 3 in Table 2), with the stratifying variable excluded from its own model. P values for interaction were obtained from multiplicative product terms. HR, hazard ratio; CI, confidence interval.						

Sensitivity analyses

All prespecified sensitivity analyses were directionally consistent with the main analysis (Table 4): non-diabetic-population equation (S1), HR = 1.381; NHANES standardized equation (S2), HR = 1.428; after excluding baseline anemia (S3), the high-HGI end retained a significant association (HR = 1.282, 95%CI 1.017–1.615, $P = 0.035$), whereas the low-HGI end ($< P10$) showed no elevated risk (HR = 0.865, 95%CI 0.652–1.148, non-significant); midpoint imputation of death timing (S4), HR = 1.410; model without systolic blood pressure (S5a, $n = 10,170$), HR = 1.376; multiple imputation (S5b, $m = 20$), HR = 1.338. Δ HGI between 2011 and 2015 was not significantly associated with incident stroke after 2015 (continuous HR = 0.960, $P = 0.401$; binary HR = 0.933, $P = 0.531$).

Table 4

Summary of robustness and sensitivity analyses for the association between high hemoglobin glycation index (HGI) and incident stroke

Analysis	HR/SHR (95% CI)	P value
Main analysis (anchor)	1.392 (1.122–1.728)	0.003
Competing risk (Fine–Gray; subdistribution HR)	1.386 (1.119–1.716)	0.003
S1 HGI equation refit in participants without diabetes	1.381 (1.110–1.719)	0.004
S2 Published NHANES standardized HGI equation	1.428 (1.153–1.768)	0.001
S3 Excluding participants with anemia (high-HGI end)	1.282 (1.017–1.615)	0.035
S3a Low-HGI end (after excluding anemia)	0.865 (0.652–1.148)	0.314
S4 Midpoint imputation of event/death timing	1.410 (1.136–1.748)	0.002
S5a Model without systolic blood pressure (n = 10,170)	1.376 (1.125–1.684)	0.002
S5b Multiple imputation (m = 20)	1.338 (1.097–1.631)	0.004
S6 Δ HGI 2011–2015 per 1-0053BD increase (outcome after 2015)	0.960 (0.873–1.056)	0.401
S6 Δ HGI highest quartile vs the rest (outcome after 2015)	0.933 (0.751–1.159)	0.531
HR, hazard ratio; SHR, subdistribution hazard ratio; CI, confidence interval; HGI, hemoglobin glycation index; Δ HGI, change in HGI between the 2011 and 2015 blood examinations. Except where indicated, estimates are for high HGI ($\geq$ P90) from the fully adjusted Cox model. S6 analyses were restricted to incident strokes occurring after 2015 and adjusted for baseline HGI.		

Discussion

In this prospective study based on a nationally representative cohort with approximately 9 years of follow-up, we found that high baseline HGI ($\geq$ P90) was independently associated with an approximately 39% higher risk of incident stroke among middle-aged and older Chinese adults. The association remained significant after adjustment for HbA1c, meeting the prespecified decision criterion for the key test, and was consistent across the competing-risk model and eight sensitivity analyses. To our knowledge, this is the first study to directly answer—with a prespecified key test in a nationally representative general population—whether HGI is independent of HbA1c, and the first to provide mechanistic verification of the J-shaped signal at the low-HGI end.

Comparison with previous studies

The study closest to ours is the NCRCHS rural Northeast China cohort [20], which reported a 70% higher stroke risk among participants with isolated high HGI without diabetes (HR = 1.701), directionally consistent with and similar in magnitude to our subgroup finding (HR = 1.478). However, the NCRCHS involved a regional population with a median follow-up of 4.23 years and only 79 events, and neither

included HbA1c and HGI in the same model nor addressed the competing risk of death. With a national sample, 841 events, and a more complete analytic chain, our study substantially strengthens that finding.

Notably, evidence from populations with diabetes does not support an independent value of HGI: secondary analyses of DEVOTE and ADVANCE showed that HGI no longer predicted cardiovascular events once HbA1c entered the model [11, 12], the DCCT data in type 1 diabetes were likewise negative [7], and a 15-year prospective cohort in type 2 diabetes found no advantage of HGI over HbA1c for any outcome [13]. In contrast, we observed an association that remained significant after HbA1c adjustment in the general population, while HbA1c itself was not significant in the same model. This discrepancy may be population-dependent: in diabetes, the HbA1c distribution is wide and glycemia itself dominates risk, diluting the marginal contribution of glycation propensity; in the general population, where HbA1c is narrowly distributed and glucose is mostly normal, glycation differences at identical glucose levels become a more discernible dimension of risk. This interpretation is corroborated by studies linking high HGI to subclinical myocardial injury, carotid atherosclerosis, coronary artery calcification, insulin resistance, and kidney injury [15–18, 33], and by the REACTION cohort, in which HGI predicted 5-year major adverse cardiovascular events [21].

Potential mechanisms

The biological basis of elevated HGI is thought to involve prolonged red blood cell lifespan, differences in membrane glucose transport, and inter-individual heterogeneity in glycation rates [6, 34, 35]. A high-HGI phenotype is also accompanied by higher levels of inflammation, with TNF α mediating downstream effects [36], and a propensity for generating advanced glycation end products (AGEs), which promote atherosclerosis through endothelial dysfunction and vascular inflammation—a plausible mechanistic pathway from “glycation propensity” to vascular events. The dissociation between normal FPG and elevated HbA1c in high-HGI individuals may additionally reflect postprandial glucose exposure not captured by fasting glucose.

Our study provides new mechanistic evidence regarding the left arm of the J-shaped curve. Conditions that accelerate red blood cell turnover, such as anemia, produce spuriously low HbA1c [37], thereby artificially lowering HGI and placing some high-risk individuals in the “low-HGI” group, generating an apparent upturn in risk at the low-HGI end. Three independent lines of internal evidence support this explanation: first, the high-HGI group paradoxically had a lower prevalence of anemia at baseline (9.2% vs 13.5%); second, after excluding participants with anemia, the low-HGI group showed no elevated risk (HR = 0.865, 95%CI 0.652–1.148, non-significant; the point estimate below 1 with a confidence interval crossing 1 supports a null rather than a protective effect), whereas the association at the high-HGI end was preserved (HR = 1.282, P = 0.035); third, the J-shaped nonlinearity test was significant among participants without diabetes, in whom interference from anemia is more cleanly discernible. This suggests that anemia status is an essential factor to consider when applying and studying HGI.

Stability and positioning of HGI

The null association between Δ HGI and subsequent stroke risk (S6) should be interpreted cautiously. DEVOTE data showed that HGI migrates across tertiles at the individual level and stabilizes after approximately 2 years [12], suggesting that HGI behaves more like a relatively stable trait marker; moreover, the 2011 and 2015 CHARLS blood samples were assayed in different batches, and cross-batch measurement differences may have attenuated any true Δ HGI signal through regression dilution. We therefore interpret S6 as indicating that “a single baseline HGI measurement is sufficient to capture its trait-like risk information,” rather than as evidence against the value of longitudinal change.

In head-to-head comparison with SHR, TyG, HbA1c, and FPG, discrimination did not differ significantly across indices (all 95% CIs of Δ C crossed zero). We report this negative result transparently: the value of HGI lies not in replacing or outperforming existing indices in predictive performance, but in revealing an orthogonal dimension of risk—identifying individuals with high glycation propensity at identical HbA1c or FPG levels. This positioning of “mechanistic information over predictive gain” is consistent with the HGI standardization initiative of Hempe et al. [25, 26].

Clinical and public health implications

Given that routine health examinations already include FPG and HbA1c, HGI can be computed from existing data at zero additional cost. Our findings suggest that among middle-aged and older adults whose HbA1c is only mildly elevated and whose FPG is normal, those with HGI in the top decile of the population have an approximately 40% higher actual stroke risk, with particular identification value among individuals without diabetes. Such “high-risk individuals whom HbA1c does not flag” are candidates for intensified management of vascular risk factors (blood pressure, lipids, lifestyle). Nevertheless, until external validation and positive evidence of incremental predictive value are available, HGI should not be used as a stand-alone stroke risk prediction tool.

Strengths and limitations

Strengths of this study include the nationally representative prospective design, approximately 9 years of follow-up with 841 events; a prespecified statistical analysis plan with explicit decision criteria for the key test, reducing the risk of data-driven analysis; systematic handling of the competing risk of death; completion and transparent reporting of all eight prespecified sensitivity analyses; and testable mechanistic evidence for the left arm of the J-shaped curve.

Several limitations should be acknowledged. First, the stroke outcome was self-reported physician diagnosis without imaging confirmation, and ischemic and hemorrhagic strokes could not be distinguished; misclassification is possible, although this definition is widely used in large cohorts such as CHARLS and non-differential misclassification generally biases associations toward the null. Second, HGI was based on a single baseline measurement, and intra-individual fluctuation cannot be fully excluded. Third, the main analysis used complete cases (with missingness dominated by systolic blood pressure); although multiple imputation yielded consistent results (HR = 1.338), residual selection bias cannot be entirely excluded. Fourth, variables such as atrial fibrillation and antithrombotic or glucose-lowering medications were unavailable, and residual confounding may exist. Fifth, death year was

missing for 663 of the 1,169 deaths and was imputed at interval midpoints (verified robust in S4). Sixth, ΔC was negative, indicating limited incremental predictive value of HGI. Seventh, findings from a single cohort require external validation.

Conclusions

Among middle-aged and older Chinese adults in the general population, high HGI was independently associated with an increased risk of incident stroke; the association was not explained by HbA1c and was more pronounced among individuals without diabetes. Spuriously low HbA1c caused by anemia can account for the previously reported J-shaped signal at the low-HGI end. As a trait-like marker of individual glycation propensity that can be computed at zero cost from routine measurements, HGI may help identify individuals at genuinely elevated stroke risk whom conventional indicators do not flag as high risk; its stand-alone predictive value is limited, and clinical translation awaits external validation and mechanistic research.

Abbreviations

AGEs: advanced glycation end products; CHARLS: China Health and Retirement Longitudinal Study; CI: confidence interval; FPG: fasting plasma glucose; HbA1c: glycated hemoglobin; HGI: hemoglobin glycation index; HR: hazard ratio; NCRCHS: Northeast China Rural Cardiovascular Health Study; RCS: restricted cubic spline; SHR: stress hyperglycemia ratio; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TyG: triglyceride–glucose index; ΔC : C-index difference; ΔHGI : change in hemoglobin glycation index.

Declarations

Ethics approval and consent to participate

CHARLS was approved by the Biomedical Ethics Review Committee of Peking University (IRB00001052-11015; blood sample collection IRB00001052-11014), and all participants provided written informed consent. This study is a secondary analysis of de-identified publicly available data.

Consent for publication

Not applicable.

Availability of data and materials

The CHARLS data used in this study are available upon application from the National School of Development, Peking University (<https://charls.pku.edu.cn/>). The analysis code for this study is available from the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by Sichuan Medical Association Youth Innovation Project(Q21046)

Authors' contributions

MP and JT conceived and designed the study. MP performed the statistical analyses and drafted the manuscript. NH contributed to data extraction, cleaning, and analysis. QM contributed to the methodology and interpretation of the results. SL contributed to data curation and visualization. JT supervised the study, acquired funding, and critically revised the manuscript for important intellectual content. All authors read and approved the final manuscript.

Acknowledgements

We thank the CHARLS project team and all participants.

Use of generative AI

AI-based language editing tools were used to improve the readability and language of this manuscript. The authors reviewed and revised all content and take full responsibility for the final version of the paper.

References

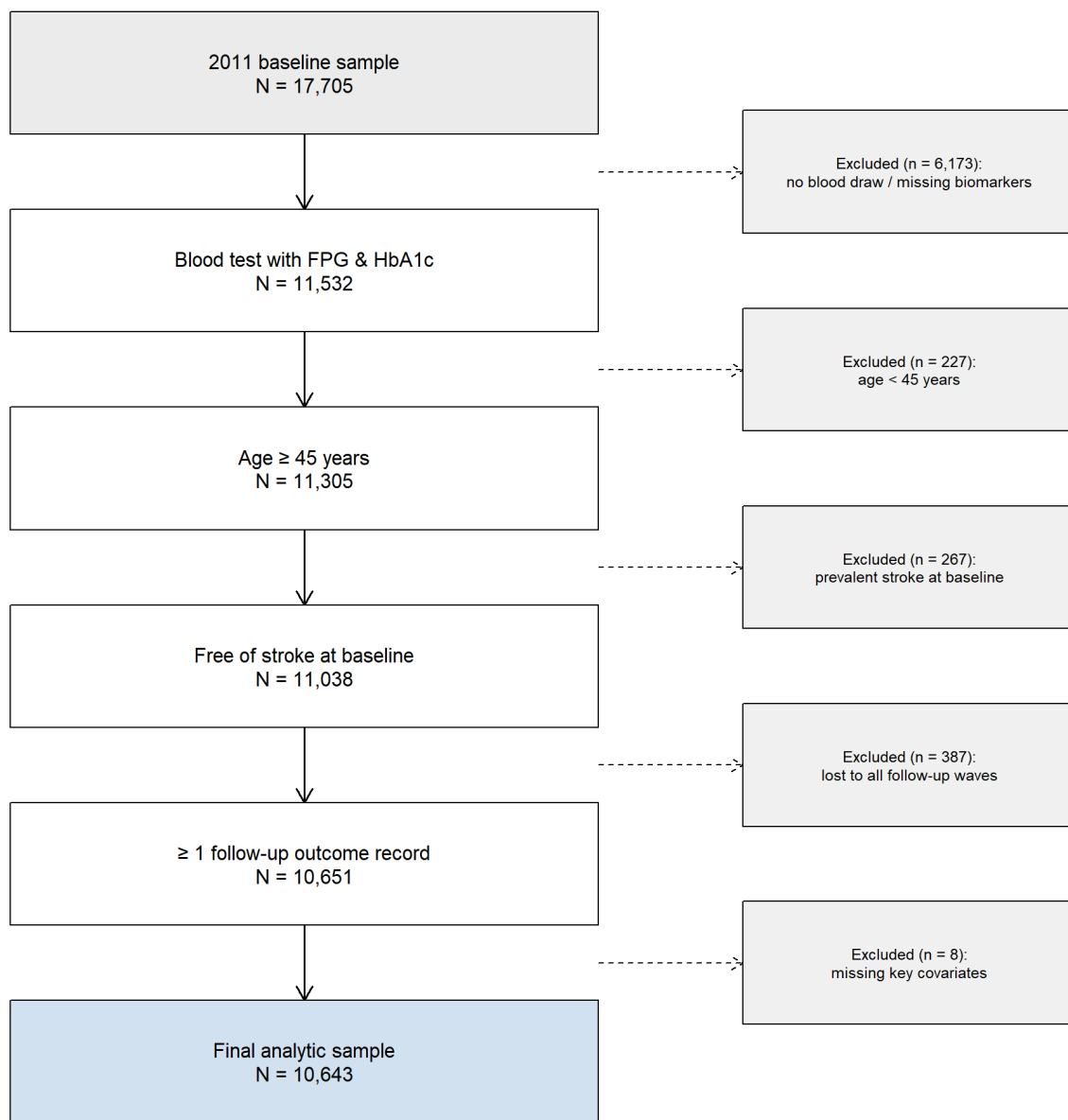
1. GBD 2019 Stroke Collaborators, Feigin VL, Stark BA, Johnson CO, et al. Global, regional, and national burden of stroke and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Neurol.* 2021;20(10):795–820.
2. Tu WJ, Wang LD, Special Writing Group of China Stroke Surveillance Report. China stroke surveillance report 2021. *Mil Med Res.* 2023;10(1):33.
3. Peters SAE, Huxley RR, Woodward M. Diabetes as a risk factor for stroke in women compared with men: a systematic review and meta-analysis of 64 cohorts, including 775,385 individuals and 12,539 strokes. *Lancet.* 2014;383(9933):1973–80.
4. Schlesinger S, Neuenschwander M, Barbaresco J, et al. Prediabetes and risk of mortality, diabetes-related complications and comorbidities: umbrella review of meta-analyses of prospective studies. *Diabetologia.* 2022;65(2):275–85.
5. Butalia S, Chu LM, Dover DC, et al. Association between hemoglobin A1c and development of cardiovascular disease in Canadian men and women without diabetes at baseline: a population-based study of 608,474 adults. *J Am Heart Assoc.* 2024;13(9):e031095.

6. Hempe JM, Gomez R, McCarter RJ Jr, Chalew SA. High and low hemoglobin glycation phenotypes in type 1 diabetes: a challenge for interpretation of glycemic control. *J Diabetes Complications*. 2002;16(5):313–20.
7. Lachin JM, Genuth S, Nathan DM, Rutledge BN. The hemoglobin glycation index is not an independent predictor of the risk of microvascular complications in the Diabetes Control and Complications Trial. *Diabetes*. 2007;56(7):1913–21.
8. Hempe JM, Hsia DS. Variation in the hemoglobin glycation index. *J Diabetes Complications*. 2022;36(7):108223.
9. McCarter RJ, Hempe JM, Gomez R, Chalew SA. Biological variation in HbA1c predicts risk of retinopathy and nephropathy in type 1 diabetes. *Diabetes Care*. 2004;27(6):1259–64.
10. Hempe JM, Liu S, Myers L, McCarter RJ, Buse JB, Fonseca V. The hemoglobin glycation index identifies subpopulations with harms or benefits from intensive treatment in the ACCORD trial. *Diabetes Care*. 2015;38(6):1067–74.
11. van Steen SC, Woodward M, Chalmers J, et al. Haemoglobin glycation index and risk for diabetes-related complications in the Action in Diabetes and Vascular Disease: Preterax and Diamicron Modified Release Controlled Evaluation (ADVANCE) trial. *Diabetologia*. 2018;61(4):780–9.
12. Klein KR, Franek E, Marso S, et al. Hemoglobin glycation index, calculated from a single fasting glucose value, as a prediction tool for severe hypoglycemia and major adverse cardiovascular events in DEVOTE. *BMJ Open Diabetes Res Care*. 2021;9(2):e002339.
13. Cardoso CRL, Leite NC, Salles GF. Importance of the hemoglobin glycation index for risk of cardiovascular and microvascular complications and mortality in individuals with type 2 diabetes. *Endocrinol Metab (Seoul)*. 2024;39(5):732–47.
14. Xu S, Qin Z, Yuan R, et al. The hemoglobin glycation index predicts the risk of adverse cardiovascular events in coronary heart disease patients with type 2 diabetes mellitus. *Front Cardiovasc Med*. 2022;9:992252.
15. Wang Z, Liu Y, Xie J, Liu NF. Association between hemoglobin glycation index and subclinical myocardial injury in the general population free from cardiovascular disease. *Nutr Metab Cardiovasc Dis*. 2022;32(2):469–78.
16. Marini MA, Fiorentino TV, Succurro E, et al. Association between hemoglobin glycation index with insulin resistance and carotid atherosclerosis in non-diabetic individuals. *PLoS ONE*. 2017;12(4):e0175547.
17. Rhee EJ, Cho JH, Kwon H, et al. Association between coronary artery calcification and the hemoglobin glycation index: the Kangbuk Samsung Health Study. *J Clin Endocrinol Metab*. 2017;102(12):4634–41.
18. Chen Z, Li D, Lin M, et al. Association of hemoglobin glycation index with contrast-induced acute kidney injury in patients undergoing coronary angiography: a retrospective study. *Front Physiol*. 2022;13:870694.

19. Yang Y, Liu M, Huang S, et al. Association between hemoglobin glycation index and poor outcome after endovascular thrombectomy in acute ischemic stroke. *Front Aging Neurosci.* 2025;17:1533584.
20. Wang P, Li Q, Guo X, et al. The value of hemoglobin glycation index–diabetes mellitus system in evaluating and predicting incident stroke in the Chinese population. *J Clin Med.* 2022;11(19):5814.
21. Wang Y, Liu H, Hu X, et al. Association between hemoglobin glycation index and 5-year major adverse cardiovascular events: the REACTION cohort study. *Chin Med J (Engl).* 2023;136(20):2468–75.
22. Guo D, Xu X. Relationship between hemoglobin glycation index and mild cognitive impairment risk in middle-aged and elderly people: a cohort study from CHARLS 2011–2018. *Front Neurol.* 2026;17:1852076.
23. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). *Int J Epidemiol.* 2014;43(1):61–8.
24. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet.* 2007;370(9596):1453–7.
25. Hempe JM, Yang S, Liu S, Hsia DS. Standardizing the haemoglobin glycation index. *Endocrinol Diabetes Metab.* 2021;4(4):e00299.
26. Soros AA, Chalew SA, McCarter RJ, Shepard R, Hempe JM. Hemoglobin glycation index: a robust measure of hemoglobin A1c bias in pediatric type 1 diabetes patients. *Pediatr Diabetes.* 2010;11(7):455–61.
27. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika.* 1994;81(3):515–26.
28. Roberts GW, Quinn SJ, Valentine N, et al. Relative hyperglycemia, a marker of critical illness: introducing the stress hyperglycemia ratio. *J Clin Endocrinol Metab.* 2015;100(12):4490–7.
29. Simental-Mendía LE, Rodríguez-Morán M, Guerrero-Romero F. The product of fasting glucose and triglycerides as surrogate for identifying insulin resistance in apparently healthy subjects. *Metab Syndr Relat Disord.* 2008;6(4):299–304.
30. Harrell FE Jr. *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis.* 2nd ed. Cham: Springer; 2015.
31. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc.* 1999;94(446):496–509.
32. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Softw.* 2011;45(3):1–67.
33. Hsia DS, Rasouli N, Pittas AG, et al. Implications of the hemoglobin glycation index on the diagnosis of prediabetes and diabetes. *J Clin Endocrinol Metab.* 2020;105(3):e130–8.

34. Cohen RM, Franco RS, Khera PK, et al. Red cell life span heterogeneity in hematologically normal people is sufficient to alter HbA1c. *Blood*. 2008;112(10):4284–91.
35. Wang J, Zhang L, Bai Y, et al. The influence of shorter red blood cell lifespan on the rate of HbA1c target achieved in type 2 diabetes patients with a HbA1c detection value lower than 7%. *J Diabetes*. 2023;15(1):7–14.
36. Lyu L, Yu J, Liu Y, et al. High hemoglobin glycation index is associated with telomere attrition independent of HbA1c, mediated by TNF α . *J Clin Endocrinol Metab*. 2022;107(2):462–73.
37. English E, Idris I, Smith G, Dhatariya K, Kilpatrick ES, John WG. The effect of anaemia and abnormalities of erythrocyte indices on HbA1c analysis: a systematic review. *Diabetologia*. 2015;58(7):1409–21.
38. Zhang X, Yang Z, Zhang M, et al. Independent effects of 2hPG, FPG and HbA1c on cardiovascular risk: analysis of a nationally representative sample from China. *Diabetes Res Clin Pract*. 2021;173:108672.

Figures



Incident stroke: 841 (7.9%); complete-case for main models: 8,684

Figure 1

Flow diagram of participant inclusion and exclusion (STROBE).

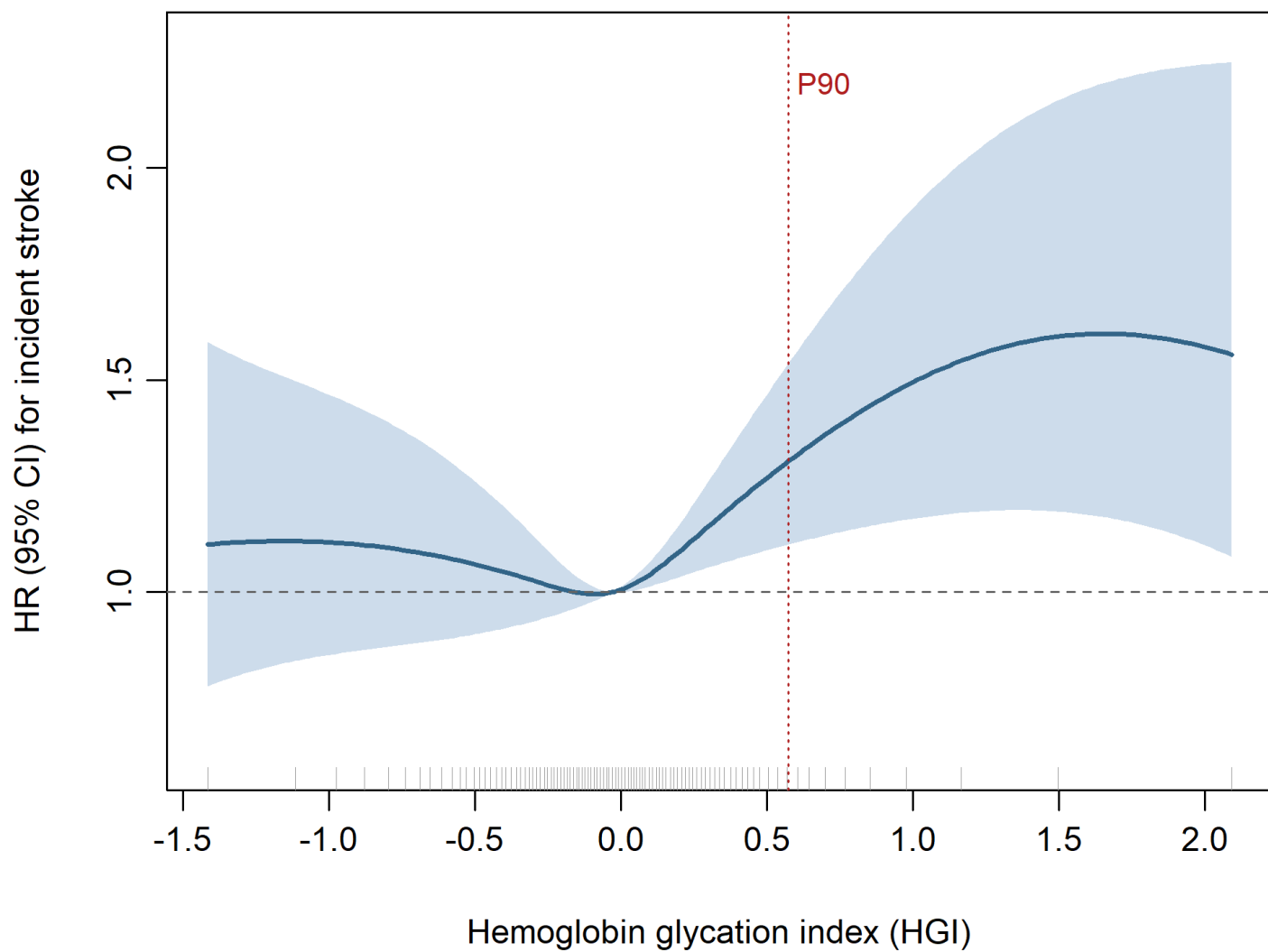


Figure 2

Restricted cubic spline curve of the association between continuous HGI and incident stroke risk (fully adjusted model; reference HGI = 0; rug plot of HGI distribution at the bottom).

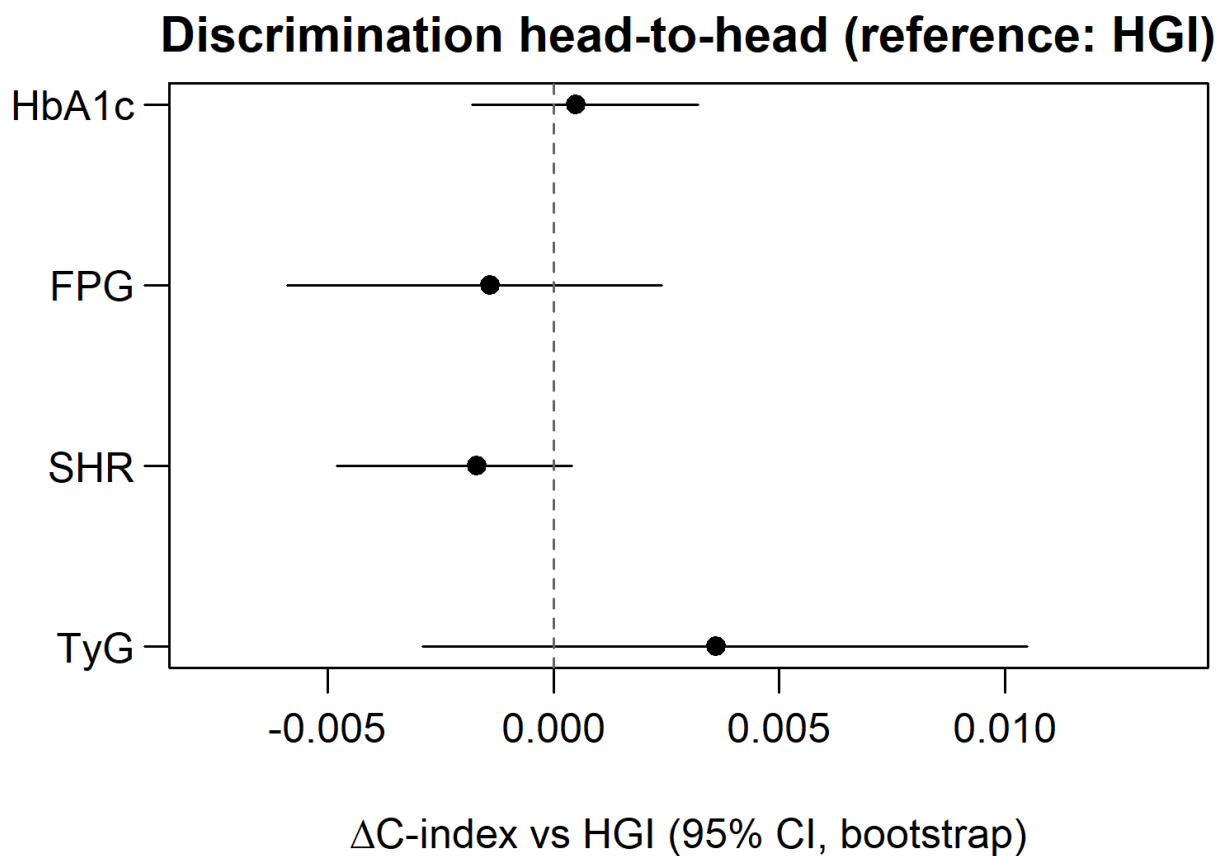


Figure 3

Forest plot of C-index differences (ΔC with 95% CIs) between each glycemia-related index model and the high-HGI model.

High HGI ($\geq$ P90) and incident stroke: subgroups

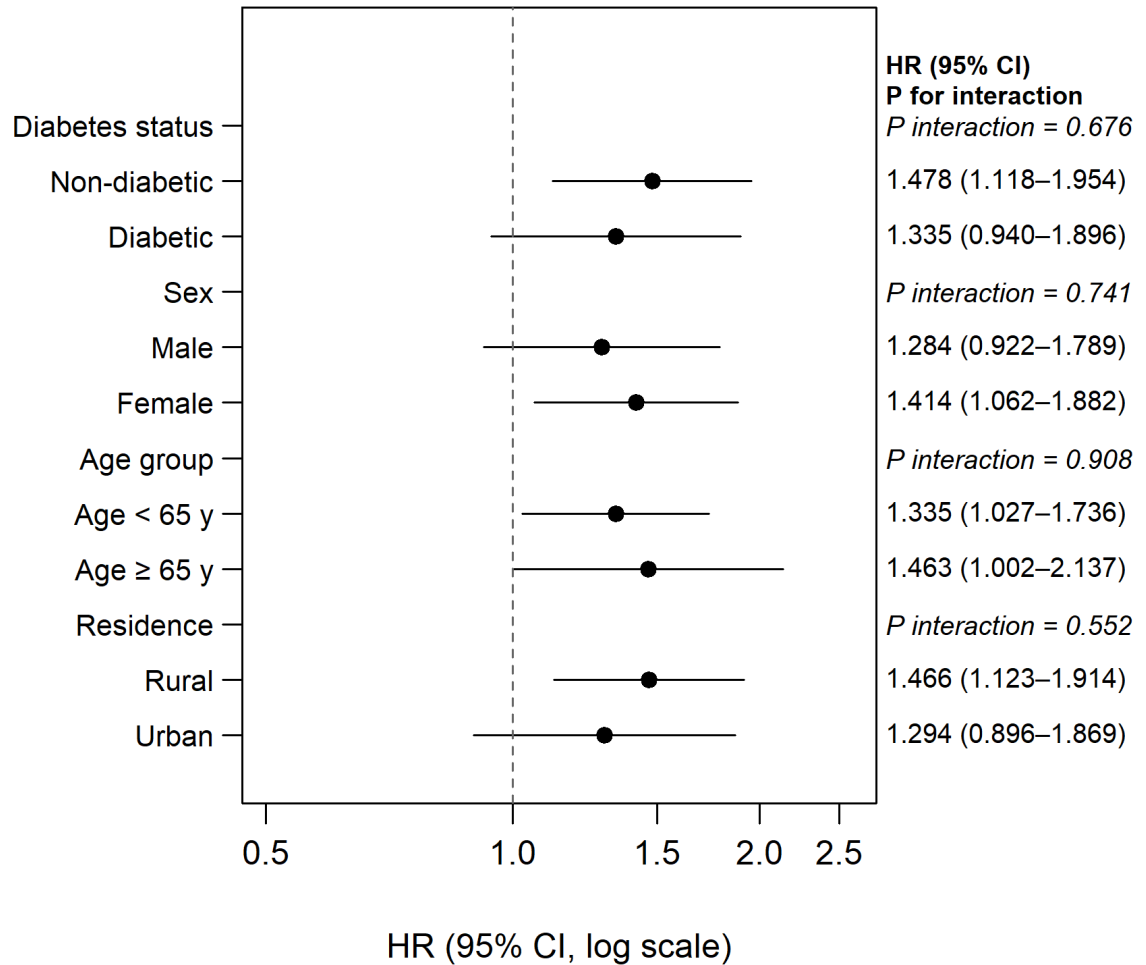


Figure 4

Forest plot of the association between high HGI ($\geq$ P90) and incident stroke across subgroups (diabetes status, sex, age, urban/rural residence), with P values for interaction.