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.

Level and Change in a Modified 7-Component Life's Essential 8 and Incident Cardiovascular Multimorbidity in Chinese Middle-Aged and Older Adults: The China Health and Retirement Longitudinal Study

Yanjiao Li
Inner Mongolia Medical University
Jiao Xue
Fengtai Community Health Service Center, Fengtai District, Beijing, China
Guojuan Liu
478282468@qq.com

Inner Mongolia People's Hospital

Research Article

Keywords: Life's Essential 8, cardiovascular multimorbidity, trajectories, latent class growth model, CHARLS, prospective cohort

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10850364/v1

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

Additional Declarations: No competing interests reported.

Level and Change in a Modified 7-Component Life's Essential 8 and Incident Cardiovascular Multimorbidity in Chinese Middle-Aged and Older Adults: The China Health and Retirement Longitudinal Study

Yanjiao Li[a], Jiao Xue[b], Guojuan Liu[c,*]

a Inner Mongolia Medical University, Hohhot, China

b Fengtai Community Health Service Center, Fengtai District, Beijing, China

c Inner Mongolia People's Hospital, Hohhot, China

* Correspondence: Guojuan Liu, Inner Mongolia People's Hospital, Hohhot, China. Email: 478282468@qq.com. ORCID: 0009-0002-7237-5533

Running title: LE8 level, change, and cardiovascular multimorbidity

---

Abstract

Background: This study aimed to investigate the association between the level and change of a modified, diet-excluded Life's Essential 8 (LE8) score and the risk of incident cardiovascular multimorbidity (CVM) in Chinese middle-aged and older adults.

Methods: A total of 4,646 cardiovascular disease-free participants aged ≥45 years from the China Health and Retirement Longitudinal Study (CHARLS), all of whom completed physical examinations in 2011, 2013, and 2015, were included. A modified 7-component LE8 score (diet excluded; not directly comparable with 8-component studies) was constructed at each wave, and latent class growth models were used to identify LE8 trajectories over 2011–2015. The primary outcome was incident CVM (hypertension with heart disease and/or stroke) during 2015–2020, ascertained at the 2020 wave.

Results: Over follow-up, 330 of the 4,547 at-risk participants developed CVM (7.3%). Three trajectories were identified: persistent high (44.3%), persistent moderate (45.1%), and persistent low (10.7%). Compared with the persistent high group, the persistent moderate and persistent low groups had higher risks of CVM (multiple-adjusted OR, 1.95 [95% CI, 1.45–2.62] and 4.04 [2.77–5.89]; P for trend <0.001), with consistent results across all 25 sensitivity analyses. The association was reproduced in an independent 30% validation sample (persistent low: HR 4.20,

95% CI 2.39–7.39; C-index 0.662; calibration slope 0.92) and in NHANES for cardiovascular mortality (HR 0.67, 95% CI 0.59–0.76 per 10 points). After additional adjustment for the baseline LE8 score, the association of the persistent low group remained significant (OR, 2.23 [1.23–4.00]), but the trajectories added little discriminative value beyond a single baseline score (C-index increment 0.005), whereas within-person decline in LE8 was associated with a higher risk (HR, 0.82 [0.72–0.92] per 10 points).

Conclusion: The level of the modified 7-component LE8 score showed a significant, graded, inverse association with incident CVM in this examination-compliant, cardiovascular disease-free population. Over the 4-year window, the trajectory classes mainly reflected baseline level, and within-person change, rather than class membership, carried the additional prognostic information; a single baseline measurement largely suffices for risk stratification, and repeated measurement adds value mainly for detecting within-person change.

Keywords: Life's Essential 8; cardiovascular multimorbidity; trajectories; latent class growth model; CHARLS; prospective cohort

---

1. Introduction

Cardiovascular multimorbidity (CVM) refers to the coexistence of two or more cardiovascular diseases in the same individual, typically hypertension combined with heart disease and/or stroke. CVM is common in aging societies and is associated with higher mortality, disability, and healthcare costs than any single cardiovascular disease [1,2]. In China, middle-aged and older adults are the main group affected by cardiovascular diseases. It is therefore crucial to identify modifiable risk factors for CVM in this population.

Life's Essential 8 (LE8) is a cardiovascular health score proposed by the American Heart Association in 2022 [3]. It includes four health behaviors (diet, physical activity [PA], nicotine exposure, and sleep) and four health factors (BMI, blood lipids, blood glucose, and blood pressure). Each component is scored from 0 to 100, and the total score is the mean of all components. Many studies have shown that LE8 is related to cardiovascular disease, death, and multiple chronic conditions in a dose-response manner [3–5], as summarized in a recent meta-analysis [6]. However, most of these studies used only one LE8 score measured at a single time point. Whether long-term dynamic LE8 is related to incident CVM in Chinese middle-aged and older adults remains unclear.

Several studies have examined LE8 in Chinese middle-aged and older adults, but most of them used a single score [7,8]. The Kailuan cohort found that 10-year LE8 trajectories predicted incident hypertension [9], and that changes in LE8 were related to cardiovascular disease risk [10]. A UK Biobank study showed that sleep patterns were associated with the dynamic progression of cardiovascular multimorbidity [11]. A nationwide Chinese cohort study (the 4C Study) linked higher single-measurement LE8 scores to lower cardiovascular-kidney-metabolic syndrome stage-related risk [12]. However, relevant studies on LE8 trajectories and CVM are lacking. First, LE8 trajectory patterns have not been described in a national, population-representative sample of Chinese middle-aged and older adults. The Kailuan cohort is an industrial occupational cohort, the 4C Study used a single baseline LE8 measurement, and the UK Biobank analyses focused on single components such as sleep. Second, no trajectory study has used CVM, rather than a single disease, as the outcome. Third, no study has directly tested whether LE8 trajectories add prognostic information beyond a single baseline score. This question has direct implications for how often cardiovascular health should be re-measured in practice.

The aim of this study was to use five waves of CHARLS follow-up data (2011–2020), with trajectories based on the three waves containing physical examinations (2011/2013/2015) [15], to (1) identify LE8 trajectory patterns in Chinese middle-aged and older adults; (2) examine the association between trajectory groups and incident CVM; and (3) formally test whether trajectories provide information beyond the baseline LE8 score. Because CHARLS lacks a food-frequency questionnaire, a modified 7-component LE8 (diet excluded, with component scoring adapted to the available data; Methods and eTable 7) was used. To our knowledge, the third aim represents the first direct assessment of the incremental value of repeated cardiovascular health measurement in this setting.

2. Materials and methods

2.1 Study population

CHARLS is a nationally representative longitudinal survey of Chinese adults aged ≥45 years and their spouses, conducted by Peking University across 28 provinces, 150 counties/districts, and 450 communities, with 17,708 baseline respondents in 2011 and biennial follow-ups through 2020 (response rate >80%) [15]. Physical examinations (blood pressure, BMI) were available in 2011/2013/2015, and venous blood biomarkers (glucose, lipids, HbA1c) were measured in 2011 and 2015 [16]. CHARLS was approved by the Biomedical Ethics Review Committee of Peking

University (IRB00001052-11015), and all participants provided written informed consent. This secondary analysis follows the STROBE guideline [17]; the completed STROBE checklist and a ROBINS-E risk-of-bias self-assessment [18] are provided in the Supplementary Material (eAppendix A–C). The analysis plan was developed post hoc, and the study was not pre-registered.

The inclusion criteria were as follows: (1) age ≥45 years at baseline (2011); (2) no baseline cardiovascular disease (self-reported physician-diagnosed hypertension, heart disease, or stroke; measured blood pressure was not an exclusion criterion); (3) complete blood pressure and BMI measurements in all three waves of 2011/2013/2015; and (4) at least one blood test in 2011 or 2015. Participants who entered the cohort after baseline (no 2011 record) were excluded. A total of 4,646 participants were included in the study cohort (Figure 1).

2.2 LE8 assessment

A modified 7-component LE8 score (0–100 per component; total = mean of available components, with ≥5 of 7 required) was constructed following the AHA 2022 definition [3]. Because CHARLS lacks a food-frequency questionnaire, the diet component was excluded, consistent with prior CHARLS studies [7]. PA was scored according to AHA/WHO intensity recommendations [19], the component thresholds followed the official AHA definition [3], and the scoring framework matched the Kailuan trajectory study [9]. Because the total is the mean of the available components, the score retains the 0–100 scale regardless of which components are missing, and no further re-normalization was applied. The score is therefore compositionally heterogeneous across participants (5 vs 7 components), which the trajectory model treats as fully comparable; availability-stratified classification checks support this assumption (eTable 10). Three components required adaptation to the CHARLS data, and each was tested in sensitivity analyses: (1) PA: MET-min/week were converted to moderate-equivalent minutes at 4 METs (the standard energy cost of moderate-intensity activity in the AHA/WHO framework [3,19]), and full points required twice the official moderate-equivalent minutes (≥300 vs ≥150 min/week); the doubling is a calibration adjustment, because the CHARLS questionnaire aggregates walking, moderate, and vigorous MET-minutes into a single total, and questionnaire-based totals overestimate device-measured moderate-to-vigorous activity by a factor of about two [30]; (2) nicotine exposure: a never/former/current-by-quantity scheme was used, because CHARLS does not record quit duration or second-hand smoke exposure; and (3) sleep health: more lenient thresholds than the official definition were used, because CHARLS records only night-time sleep (excluding naps). The official AHA thresholds for PA and nicotine exposure were tested in S13a,

those for sleep in S13d, and all three jointly in S15. PA was assessed in a randomly selected half of households (design-based missing at random), and missing components were excluded from the mean. The full component-by-component scoring versus the official AHA definitions is provided in eTable 7. Because venous blood was unavailable in 2013, the glucose and lipid components were linearly interpolated between the 2011 and 2015 measurements (carried forward or backward when only one measurement existed); no-interpolation and LOCF alternatives were tested in S1 and S2.

2.3 Trajectory modeling

LE8 scores at 2011/2013/2015 (time 0/2/4 years) were modeled using latent class growth models (LCGM; hlme function of the R package lcmm) with quadratic time effects [20]. The fitted model is a latent-class linear mixed model with class-specific fixed intercepts and time coefficients plus a shared random participant intercept, which allows within-class heterogeneity (a growth-mixture-type specification rather than a strict Nagin-style LCGM). A linear-time specification was also fitted for comparison (eTable 17). The quadratic term was clearly required (BIC 92,181.7 vs 97,453.0 for the three-class solutions), and the class structure was nearly identical under either specification.

Participants with ≥2 waves of LE8 (n=4,625) were included in the trajectory analysis; 3,777 (81.7%) contributed all three waves and 848 (18.3%) contributed two waves. Missing waves were handled under the missing-at-random assumption. Models with 1–5 classes were compared using a pre-specified hierarchical decision rule: (1) minimum class proportion ≥5%, applied first, because classes below this threshold are clinically uninterpretable and unstable; (2) the lowest BIC among solutions meeting this constraint; and (3) entropy >0.7 and average posterior probability ≥0.7 as classification-quality criteria [13]. A three-class solution was selected and named by baseline level. The term "persistent" refers to relative stability over the 4-year trajectory window (2011–2015) rather than to lifelong patterns.

2.4 Outcome

Definition. The primary outcome was incident CVM, defined as the first occurrence of two or more coexisting cardiovascular diseases (hypertension with heart disease and/or stroke) during follow-up, based on self-reported physician diagnoses, as in other CHARLS-based studies [7,8]. The definition follows the cardiometabolic multimorbidity framework [2] and prior operationalizations in population surveys [21], with hypertension included given its central role in the Chinese CVD spectrum. The robustness of the definition was verified by excluding the blood-

pressure component of LE8 (S10), by outcome variants that do not depend on hypertension (S11, S12), by a cross-analysis separating circularity from pathway effects (S16), and by the exclusion of participants with undetected measured hypertension (S17a/S17b). Self-reported hypertension shows limited agreement with measured blood pressure in this population [23], which may attenuate the associations; stepwise single-disease analyses (incident hypertension, heart disease, and stroke separately; §3.5) additionally inform the interpretation.

Ascertainment and censoring. The diagnosis waves with complete data on all three diseases were 2013/2015/2020; the 2018 wave lacked stroke data and could not be used for stroke ascertainment, although hypertension and heart-disease diagnoses recorded at 2018 contributed to the disease status on which the 2020 coexistence classification rests. Among the 330 incident CVM events, the constituent combinations ascertained at the 2020 wave were hypertension plus heart disease (n=210), hypertension plus stroke (n=73), and all three diseases (n=29); the combination could not be classified in 18 events because of missing disease data at that wave. These 18 events were counted as events in the primary analysis, and their exclusion was tested in S14. Events were ascertained in 2020, strictly after the trajectory window (2015–2020), and the risk set comprised participants free of CVM through 2015 (n=4,547). Of these, 4,064 (89.4%) had complete outcome information at the 2020 wave (follow-up completion rate). The remaining 483 participants (314 who died without a confirmed CVM and 169 lost to follow-up) were censored at their last known diagnosis wave. Deaths without confirmed CVM were treated as non-events in the primary logistic analysis and were additionally examined with the Fine–Gray competing-risk model (S3). Death (radyear/radmonth) was treated as a competing event [24].

Time structure and primary model. Event times were approximated by interval midpoints (event = 3.5 years; censoring = 5.0 years; t0 = 2015). Because CVM was ascertained only at the 2020 wave, the data contain only two event-ascertainment time points, and event times carry little information beyond event status. Logistic regression of 5-year cumulative risk was therefore adopted as the primary framework, because fixed-window cumulative risk is the most directly interpretable measure for screening decisions; Cox and Fine–Gray models served as framework-consistent secondary analyses. This choice is supported by the near-identity of the logistic and Cox estimates (OR 1.95/4.04 vs HR 1.90/3.75). The interval-midpoint approximation assigns all events a single time and therefore cannot distort the ordering of event times, but it precludes inference about hazard timing. Person-time rates are accordingly not reported, and the results are interpreted as comparisons of cumulative risk within the fixed 5-year window (median follow-up is not applicable; maximum follow-up 5.0 years).

2.5 Covariates and statistical analysis

The baseline (2011) covariates were age (continuous, years), sex, education (three categories), marital status (married vs other), residence (rural vs urban), alcohol drinking, log-transformed household per capita consumption (13.8% missing), diabetes and dyslipidemia (self-reported), and the number of self-reported chronic conditions (0–6 in this sample; a count of 14 conditions: hypertension, dyslipidemia, diabetes, cancer, lung disease, liver disease, heart disease, stroke, kidney disease, digestive disease, psychiatric illness, memory-related disease, arthritis, and asthma; hypertension, heart disease, and stroke are zero by design in this cohort, and the count overlaps with the diabetes and dyslipidemia covariates, a double adjustment that is conservative and more likely to attenuate than to inflate the estimates). LE8 components constitute the exposure and were not entered as covariates, to avoid over-adjustment. Primary analyses used complete cases (n=3,848; 278 events), and covariate missingness was addressed by multiple imputation (m=20) [25].

Baseline characteristics were presented as the mean (SD) for continuous variables and as n (%) for categorical variables, with standardized mean differences (SMD); Kruskal–Wallis and chi-square tests were used for crude comparisons. Cumulative-incidence (Aalen–Johansen) curves with death as a competing event were used to display outcome rates, with Kaplan–Meier curves as descriptive material. The association between trajectory groups and CVM was assessed using logistic regression models: the crude model (Model 1) adjusted for no variables; Model 2 adjusted for age, sex, residence, education, marital status, drinking, and log consumption; and Model 3 further adjusted for diabetes, dyslipidemia, and chronic conditions. Odds ratios (ORs) with 95% CIs were reported. The primary trend test used the class-specific median baseline LE8 score as a continuous variable (per 10 points), with the trajectory ordinal (low=2/moderate=1/high=0) as a supplementary check. Cox proportional hazards models were fitted as framework-consistent analyses.

The incremental value of repeated measurement (Aim 3) was assessed through two co-primary exposures, trajectory-class membership and within-person change (ΔLE8), against the baseline score. Trajectory groups and baseline LE8 were entered jointly on the subsample with baseline scores, and ΔLE8 was analyzed overall and within baseline-LE8 tertiles, where trajectory-group associations and C-index increments conditional on baseline level were additionally compared (eTable 11). A component-contribution decomposition was performed by entering baseline component scores one at a time (eTable 16).

Sensitivity analyses were performed: (1) to test the interpolation method (no interpolation, S1; LOCF, S2) and the complete three-wave sample (S7); (2) to test component thresholds (Chinese BMI criteria, S9; sleep 6–8 h, S9b; official AHA thresholds for PA and nicotine, S13a, and for sleep, S13d; a joint analysis re-scoring PA, nicotine, and sleep simultaneously with official thresholds, S15); (3) to test score composition (six-component LE8 without blood pressure, S10; without PA, S13e) and classification quality stratified by the number of available components (eTable 10); (4) to address competing risk (Fine–Gray, S3), reverse causation (S4), unmeasured confounding (E-value [26], S5), and covariate missingness (multiple imputation, S6); (5) to test outcome definitions (inclusion of 2015, 2013, or 2018 events, S8a–S8c; heart disease plus stroke without hypertension, S11; a diabetes-anchored definition, S12; exclusion of the 18 unclassifiable events, S14); (6) to test a cross-analysis combining the blood-pressure-excluded score with the hypertension-free outcome (S16); (7) to exclude participants with undetected measured hypertension at baseline (S17a) and additionally at 2015 (S17b); and (8) to address selection by inverse-probability weighting with stabilized weights and 99th-percentile truncation (weight distribution in eTable 12). In total, 25 sensitivity analyses (S1–S17, including sub-variants) were performed; full definitions are provided in the Supplementary Material.

Subgroup analyses were performed by sex, age group (45–59 vs ≥60 years), and residence, with interaction likelihood-ratio tests. Given the number of events, the minimum interaction effect sizes detectable with 80% power were computed analytically (eTable 15), and all subgroup analyses were regarded as exploratory. A two-sided P value <0.05 was considered statistically significant. The analyses were performed using R 4.6.0 (data.table 1.18.4, lcmm 2.2.2, survival 3.8.6, cmprsk 2.2.12, mice 3.19.0).

2.6 Split-sample internal validation. The trajectory cohort was randomly split into derivation (70%) and validation (30%) samples, stratified by outcome status (event, no event, CVM before 2015) and full-sample trajectory class (fixed random seed). The LCGM was re-estimated in the derivation sample (one to five classes, identical specification and selection rules). Validation participants were assigned to derivation classes by posterior probability, based on the class-specific trajectories, the residual variance, and the class proportions. In the validation sample we report class agreement with the full-sample classification (Cohen's κ), Model 3 estimates, the C-index of the derivation Cox model applied to the validation sample (covariates-only reference), and the calibration of the derivation logistic model (observed/expected ratio, intercept, and slope; eTable 20, eFigure 2).

2.7 Independent NHANES cohort validation. The US National Health and Nutrition Examination Survey (NHANES) 2011–2014, linked to the National Death Index through 31 December 2019 (median follow-up 6.7 years), was used to examine the LE8 score in an independent cohort; these cycles match the CHARLS measurement period and provide the longest follow-up available in the public-use mortality files. Adults aged ≥45 years with the same 7-component score (identical scoring code) formed the validation cohort (n = 6,057; 905 deaths, 295 from cardiovascular causes [diseases of the heart or cerebrovascular disease]); a cardiovascular disease–free subcohort (n = 2,652) was analysed separately. Survey-weighted Cox models (MEC weights, strata, and primary sampling units) regressed all-cause and cardiovascular mortality on the LE8 score (per 10 points; AHA categories 0–49/50–79/80–100), adjusted for age, sex, and race/ethnicity. NHANES is cross-sectional, so trajectories cannot be derived; this analysis therefore tests the LE8 score rather than the trajectories.

3. Results

3.1 Study population and baseline characteristics

Of 25,873 unique CHARLS respondents, 17,708 were baseline (2011) respondents. The exclusion criteria were applied sequentially (Figure 1): 6,341 participants were aged <45 years (n=648) or had baseline cardiovascular disease (n=5,693), leaving 11,367; 6,500 had incomplete three-wave physical examinations, leaving 4,867; and 242 had no blood test in 2011 or 2015 (n=221) or LE8 available in <2 waves (n=21), leaving the trajectory cohort of 4,625 (the primary cohort before the ≥2-wave requirement comprised 4,646). After the further exclusion of 78 participants with CVM before 2015, the primary risk set comprised 4,547 participants. During follow-up, 330 of the 4,547 at-risk participants developed CVM (7.3%) over the fixed 5-year window (maximum follow-up 5.0 years); person-time rates are not reported because event timing is not identifiable under the two-point ascertainment design. A total of 314 participants died during follow-up (6.9%).

PA was missing in 58.2%, 64.0%, and 52.9% of participants in 2011, 2013, and 2015, respectively (primary cohort, n=4,646). Approximately half of this missingness reflects the design-based half-household sampling of the PA module, and the remainder arose from within-module non-response. In the trajectory cohort, the availability of the seven components per wave ranged from 36–47% (PA) and 74–99% (nicotine exposure, sleep, blood pressure) to 82–100% (glucose, lipids) and 100% (BMI); the full distribution of available component counts per wave,

and classification quality and estimates stratified by the number of available components, are shown in eTable 10.

Compared with baseline respondents who were not included (n=13,062), the included participants (n=4,646) were similar in age and sex (age P=0.072; sex P=0.593) but had lower baseline cardiovascular disease (hypertension 35.3%, heart disease 16.6%, and stroke 3.8% in the excluded group) and lower cardiometabolic burden, BMI, and blood pressure (eTable 6). Selection was therefore toward a healthier sample, which may attenuate the association. Inverse-probability-weighted analysis gave similar estimates (Model 3: HR 2.00, 95% CI 1.41–2.84 for persistent moderate; HR 3.61, 95% CI 2.28–5.71 for persistent low; S13b). Because the inverse-probability weights address selection into the analytic sample rather than population representativeness, the class proportions describe this examination-compliant cohort rather than national prevalence.

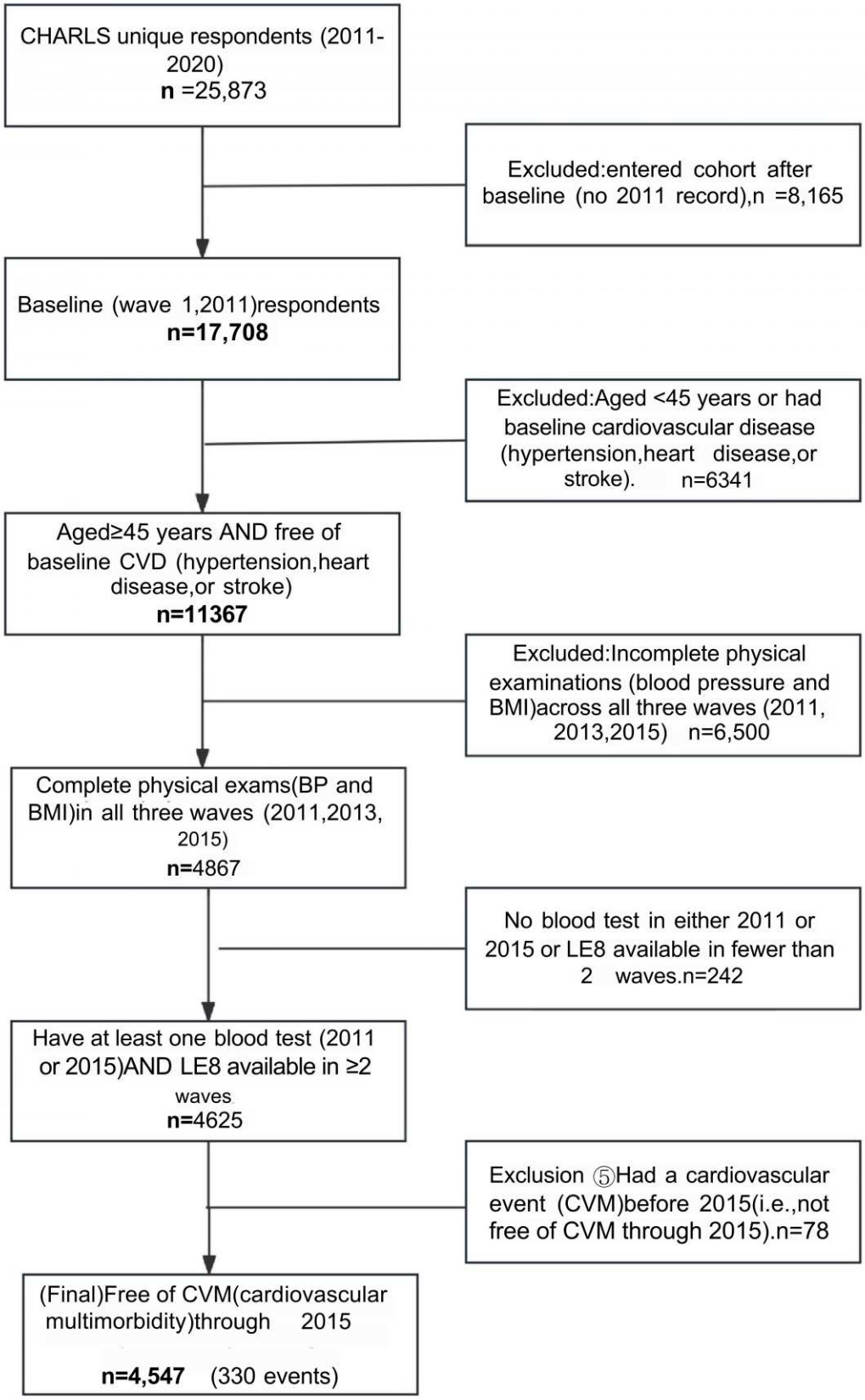

CHARLS unique respondents (2011-2020)
n =25,873
Excluded:entered cohort after baseline (no 2011 record),n =8,165
Baseline (wave 1,2011)respondents
n=17,708
Excluded:Aged <45 years or had baseline cardiovascular disease (hypertension,heart disease,or stroke). n=6341
Aged≥45 years AND free of baseline CVD (hypertension,heart disease,or stroke)
n=11367
Excluded:Incomplete physical examinations (blood pressure and BMI)across all three waves (2011, 2013,2015) n=6,500
Complete physical exams(BP and BMI)in all three waves (2011,2013, 2015)
n=4867
No blood test in either 2011 or 2015 or LE8 available in fewer than 2 waves.n=242
Have at least one blood test (2011 or 2015)AND LE8 available in ≥2 waves
n=4625
Exclusion ⑤Had a cardiovascular event (CVM)before 2015(i.e.,not free of CVM through 2015).n=78
(Final)Free of CVM(cardiovascular multimorbidity)through 2015
n=4,547 (330 events)

Figure 1. Study flow diagram (STROBE). Of 25,873 unique respondents across 2011–2020, 8,165 who entered after baseline (no 2011 record) were excluded, leaving 17,708 baseline (2011) respondents. Sequential exclusions: aged <45 years or baseline cardiovascular disease (hypertension, heart disease, or stroke; n=6,341, comprising 648 age and 5,693 baseline CVD) → 11,367; incomplete three-wave physical examinations (blood pressure and BMI; n=6,500) → 4,867; no blood test in 2011 or 2015 or LE8 available in <2 waves (n=242, comprising 221 without blood tests and 21 with LE8 <2 waves) → 4,625 (trajectory cohort); CVM before 2015 (n=78) → primary risk set n=4,547 (330 incident CVM events, 7.3%).

Table 1 shows the baseline characteristics of the study cohort stratified by trajectory group. Significant differences were found in age, sex, marital status, residence, drinking, diabetes, dyslipidemia, and LE8 scores (P<0.001 for each). Starting with the demographic variables, there was a steady increase in age (57.0 ± 8.3 to 60.3 ± 8.9 years) and in the percentage of males (39.1% to 65.3%) from the persistent high to the persistent low group. Rural residence declined across the groups (72.7% to 63.6%), whereas drinking progressively increased (29.8% to 44.1%). Education (P=0.073) and log household consumption (P=0.988) showed no significant differences. The prevalence of diabetes (1.5% to 9.6%) and dyslipidemia (3.4% to 8.5%) increased steadily across the groups, while the proportion with ≥2 chronic conditions was similar (P=0.431). Baseline LE8 scores showed a strong gradient (89.2 ± 7.1 to 62.1 ± 8.9; SMD 3.373), which persisted at the 2015 wave (86.2 ± 7.3 to 59.9 ± 8.9; SMD 3.232), indicating stable separation of cardiovascular health levels across the trajectory window.

Variable	**Persistent high (n=2,048)**	**Persistent moderate (n=2,084)**	**Persistent low (n=493)**	**P**	**SMD**
Age (years)	57.0 ± 8.3	59.3 ± 8.9	60.3 ± 8.9	<0.001	−0.382
Male	794 (39.1%)	1080 (53.0%)	314 (65.3%)	<0.001	−0.542
Education (primary/secondary/tertiary, %)	90.1 / 9.0 / 0.9	92.1 / 7.3 / 0.6	93.3 / 6.2 / 0.4	0.073	−0.117/0.106
Married	1862	1790	413	<0.001	0.099

	(91.8%)	(87.9%)	(85.9%)		
Rural	1475 (72.7%)	1425 (70.0%)	306 (63.6%)	<0.001	0.196
Drinking	605 (29.8%)	773 (37.9%)	212 (44.1%)	<0.001	−0.299
Diabetes	30 (1.5%)	86 (4.2%)	46 (9.6%)	<0.001	−0.361
Dyslipidemia	67 (3.4%)	107 (5.3%)	40 (8.5%)	<0.001	−0.218
Chronic conditions ≥2	426 (21.0%)	438 (21.5%)	114 (23.7%)	0.431	−0.065
Log consumption	8.3 ± 0.8	8.3 ± 0.9	8.3 ± 0.8	0.988	−0.009
LE8 total (w1)	89.2 ± 7.1	77.4 ± 7.9	62.1 ± 8.9	<0.001	3.373
LE8 total (w3)	86.2 ± 7.3	72.6 ± 7.8	59.9 ± 8.9	<0.001	3.232

Table 1. Baseline characteristics of participants by LE8 trajectory group. Continuous variables as mean ± SD (Kruskal–Wallis); categorical as n (%) (chi-square); SMD = standardized mean difference (persistent high vs persistent low); |SMD|>0.1 indicates a meaningful difference. Education is shown as one three-category variable (1 = primary school or below; 2 = secondary; 3 = tertiary; row percentages within each group; SMD for the primary vs secondary categories). Although the chi-square P for education was 0.073, the |SMD| for individual categories slightly exceeded 0.1 (−0.117/0.106); the between-group differences are small in absolute terms, and education was adjusted in Model 2. Sample sizes refer to the trajectory cohort (n=4,625); the primary analysis risk set comprised 4,547 participants (330 events).

3.2 LE8 trajectories

The three-class solution met all selection criteria (BIC 92,181.7; entropy 0.897; average posterior probability 0.866; minimum class proportion 10.7%). For transparency, the four-class (BIC 92,035.9) and five-class (BIC 92,007.2) solutions had lower BIC values than the three-class

solution, but both contained classes below 5% of the sample (1.2% and 0.7%, respectively), which are clinically uninterpretable and unstable. The pre-specified decision rule therefore selected the three-class solution. The four-class solution nonetheless reproduced the same monotonic outcome gradient (OR 1.60, 3.55, and 8.22 for the three lower classes vs the highest; eTable 1 note), indicating that the class-count choice does not affect the conclusions. The full model comparison is provided in eTable 1. The class-specific baseline component-score profiles are summarized in eTable 2.

Observed mean LE8 scores declined from 2011 to 2015 in all groups (Figure 2): from 89.2 to 86.2 in the persistent high group, from 77.4 to 72.6 in the persistent moderate group, and from 62.1 to 59.9 in the persistent low group. The median maximum posterior probability was 0.912, and the class-specific mean posterior probabilities were 0.881, 0.834, and 0.850, indicating acceptable classification quality in every class (eTable 17). A linear-time specification gave a nearly identical class structure but a decisively worse fit (BIC 97,453.0 vs 92,181.7 for three classes), confirming that the quadratic time term is required (eTable 17). Class-specific parameters are reported in eTable 8. The Hessian of the fitted model was numerically singular at the optimum, so the standard errors of the class-specific parameters could not be estimated; this limitation affects the parametric uncertainty bands of Figure 2, which should be interpreted as indicative (see Limitations). Although the three trajectories declined roughly in parallel, the group × time interaction was significant in a linear mixed model ($\chi^2 = 35.0$, df = 2, $P < 0.001$), indicating modest but detectable shape differences; the per-2-year linear declines were −1.52 ± 0.12 points (persistent high), −2.40 ± 0.22 (persistent moderate), and −1.10 ± 0.31 (persistent low).

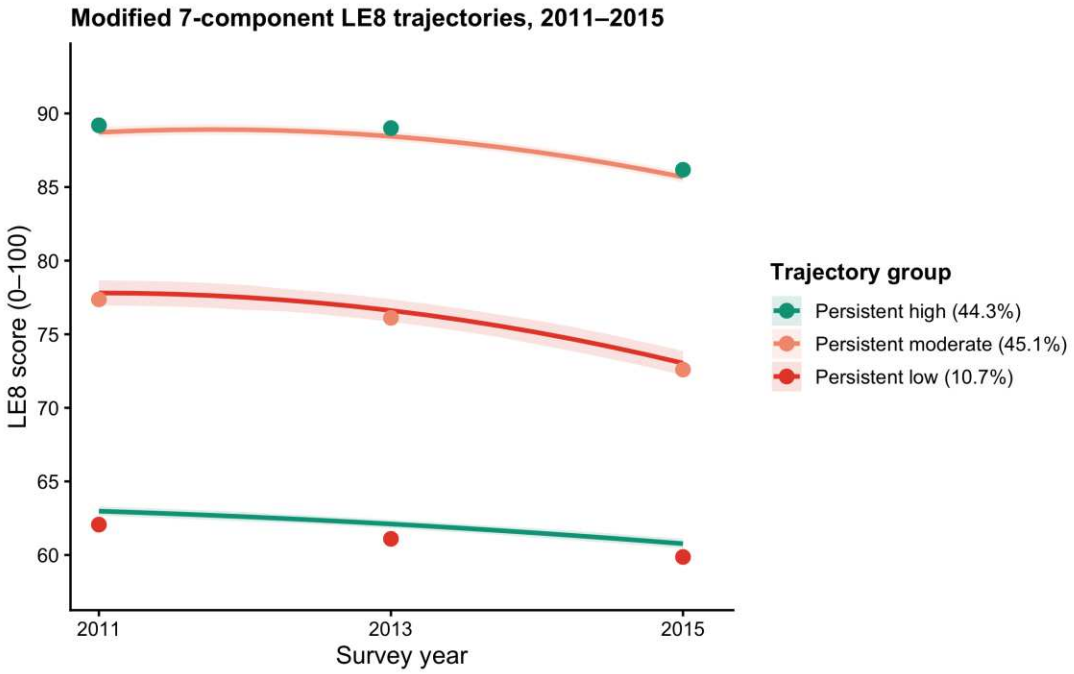
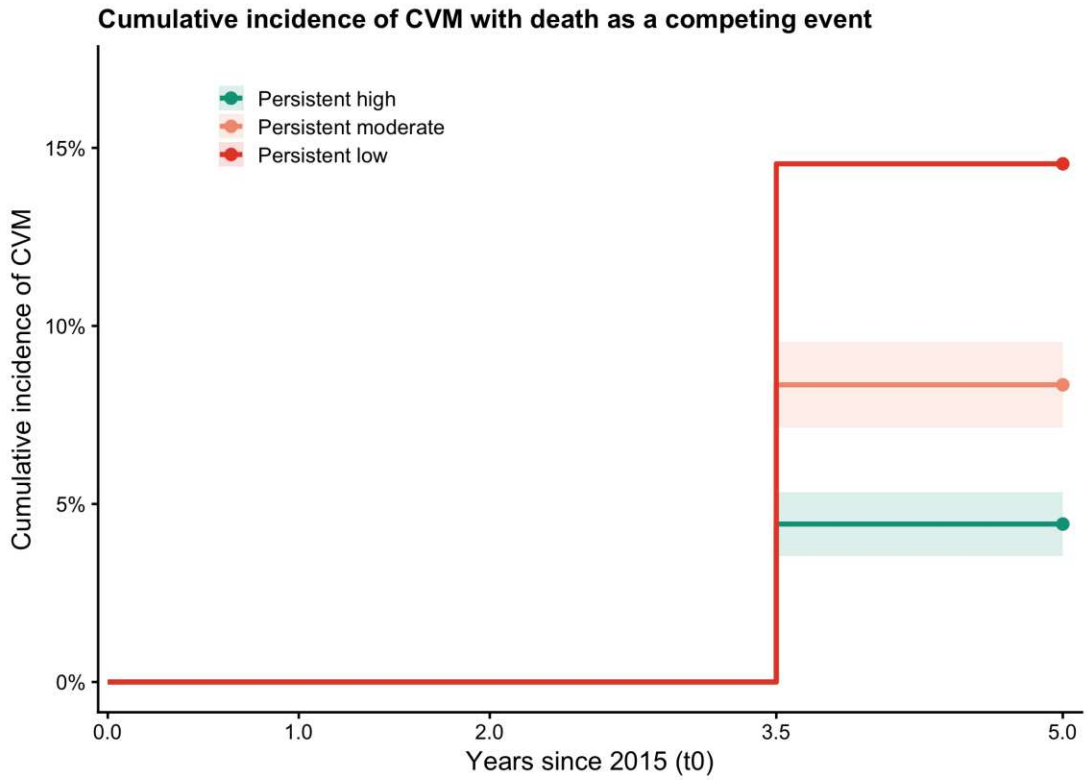


Figure 2. Modified 7-component LE8 trajectories, 2011–2015 (observed means with LCGM-predicted curves and 95% confidence bands). Persistent high (green, 44.3%), persistent moderate (orange, 45.1%), and persistent low (red, 10.7%); all groups declined slightly, with clear between-group level separation (class-specific parameters in eTable 8). "Persistent" denotes relative level stability over this 4-year window, not lifelong patterns; the bands are indicative given the singular model Hessian (see Limitations). The y-axis is truncated to display trajectory detail.

3.3 Association between LE8 trajectory groups and incident CVM

Cumulative-incidence curves with death as a competing event showed a graded separation across the groups (Figure 3). The 5-year cumulative incidences of CVM were 4.4% (95% CI, 3.5–5.3) in the persistent high group, 8.3% (7.1–9.5) in the persistent moderate group, and 14.6% (11.4–17.7) in the persistent low group. The Fine–Gray subdistribution HR for persistent low was 3.52 (95% CI 2.51–4.93; eTable 4). Kaplan–Meier curves with a log-rank test (P<0.001) are provided as descriptive material in eFigure 1.



Group	0 y	3.5 y	5 y
Persistent high	2029	2029	1939
Persistent moderate	2037	2037	1867
Persistent low	481	481	411

Figure 3. Cumulative incidence of CVM with death as a competing event, by trajectory group (Aalen–Johansen; t0 = 2015; n=4,547; 330 CVM events and 314 competing deaths). Shaded areas are 95% pointwise confidence intervals; the number-at-risk table uses the interval-midpoint convention (all events at 3.5 years). "Persistent" denotes relative level stability over the 4-year window (2011–2015), not lifelong patterns.

In the multivariable logistic regression analysis (Table 2), the association between trajectory groups and CVM was significant in the crude model, Model 2, and Model 3. Compared with the persistent high group, the ORs of the persistent moderate group were 1.96 (95% CI 1.51–2.56), 1.95 (1.46–2.62), and 1.95 (1.45–2.62), and those of the persistent low group were 3.67 (2.63–5.10), 4.09 (2.82–5.91), and 4.04 (2.77–5.89), respectively (P for trend <0.001 in all models; Figure 4). A significant trend across the class-specific median baseline LE8 score was also observed (OR 0.60 per 10 points, 95% CI 0.53–0.69, P < 0.001, after multiple adjustment). The

Cox models gave nearly identical estimates (HR 1.90, 95% CI 1.43–2.52 and 3.75, 95% CI 2.63–5.34; Table 3).

Model	Variable	n (events)	OR (95% CI)	P
Model 1	Persistent moderate	3,848 (278)	1.96 (1.51–2.56)	<0.001
Model 1	Persistent low	3,848 (278)	3.67 (2.63–5.10)	<0.001
Model 1	Ordinal trend	3,848 (278)	—	<0.001
Model 2	Persistent moderate	3,848 (278)	1.95 (1.46–2.62)	<0.001
Model 2	Persistent low	3,848 (278)	4.09 (2.82–5.91)	<0.001
Model 2	Ordinal trend	3,848 (278)	—	<0.001
Model 3	Persistent moderate	3,848 (278)	1.95 (1.45–2.62)	<0.001
Model 3	Persistent low	3,848 (278)	4.04 (2.77–5.89)	<0.001
Model 3	Ordinal trend (per category)	3,848 (278)	2.04 (1.67–2.49)	<0.001
Model 3 + baseline LE8	Persistent moderate	3,434 (251)	1.41 (0.98–2.02)	0.065
Model 3 + baseline LE8	Persistent low	3,434 (251)	2.23 (1.23–4.00)	0.008
Model 3 + baseline LE8	Baseline LE8 (per 10 points)	3,434 (251)	0.781 (0.661–0.922)	0.003

Table 2. Multivariable logistic regression analysis to assess the association between LE8 trajectory groups and incident CVM (5-year cumulative risk; reference: persistent high). Model 1 = unadjusted; Model 2 = adjusted for age, sex, residence, education, marital status,

drinking, and log consumption; Model 3 = adjusted for Model 2 plus diabetes, dyslipidemia, and chronic conditions (complete-case sample n=3,848; 278 events). The last three rows show the joint model additionally including the baseline LE8 score (per 10 points), fitted on the subsample of complete cases who also had a baseline (2011) LE8 total (n=3,434; 251 events; 501 risk-set participants lacked a 2011 total because their scored waves were 2013 and 2015; see eTable 18 for the sample flow); in this joint model all participants share a common baseline level, so the trajectory coefficient is interpreted at a given baseline score. The primary trend test used the class-specific median baseline LE8 score as a continuous variable (OR 0.60 per 10 points, 95% CI 0.53–0.69); the trajectory ordinal (low=2/moderate=1/high=0) gave a per-category OR of 2.04 (1.67–2.49). Cox models gave nearly identical estimates (Table 3, primary-analysis row). Group sizes and events refer to the full risk set (n=4,547): persistent high n=2,029 (90 events; 5-year incidence 4.4%), persistent moderate n=2,037 (170 events; 8.3%), persistent low n=481 (70 events; 14.6%).

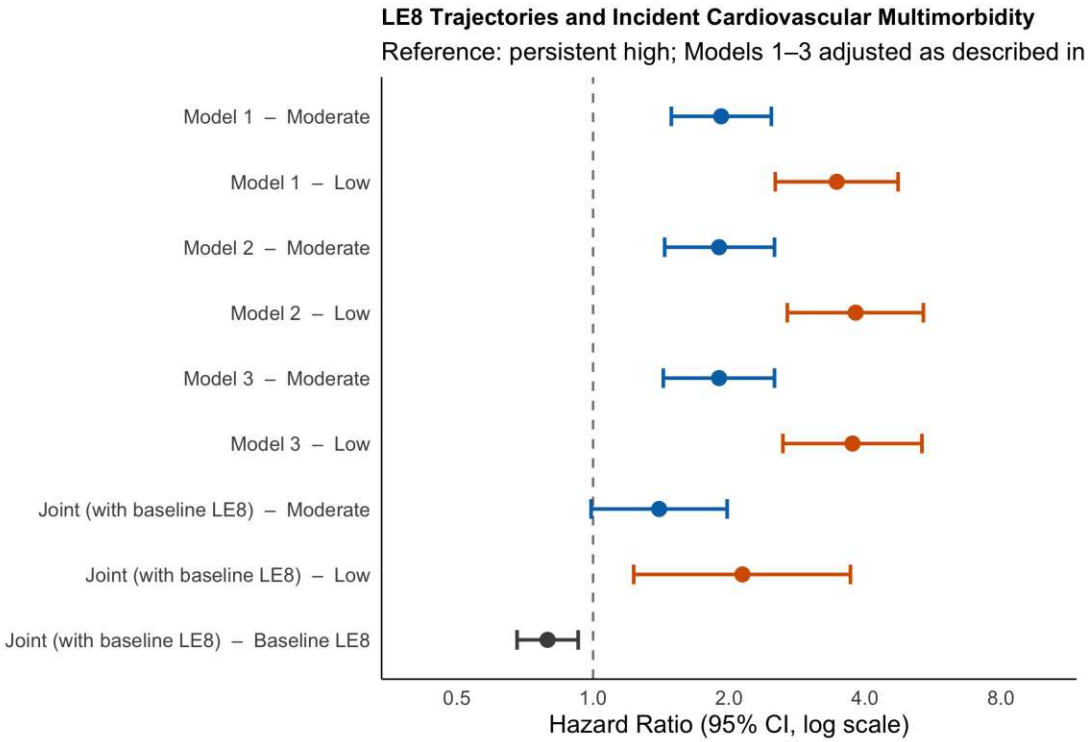


Figure 4. Forest plot of odds ratios (Models 1–3 and the joint model including baseline LE8; reference: persistent high; complete-case sample n=3,848). ORs for the persistent low group ranged from 3.67 to 4.09 across Models 1–3 (persistent moderate ≈1.95); in the joint logistic model, the persistent low group retained OR 2.23 (95% CI 1.23–4.00), the persistent moderate

group OR 1.41 (95% CI 0.98–2.02), and baseline LE8 corresponded to OR 0.781 per 10 points (95% CI 0.661–0.922); the Cox framework gave nearly identical estimates.

With respect to the incremental value of repeated measurement, 501 of the 4,547 risk-set participants had no baseline (2011) LE8 total, because their scored waves were 2013 and 2015. All incremental-value analyses were therefore restricted to participants with a baseline total (n=4,046; 295 events; complete-case analyses n=3,434, 251 events); the ΔLE8 analysis additionally required a 2015 score (n=3,761; 283 events). The full sample flow is given in eTable 18. In the joint model fitted on this common-baseline subsample, the persistent low group remained significant (OR 2.23, 95% CI 1.23–4.00), whereas the persistent moderate group did not reach statistical significance (OR 1.41, 95% CI 0.98–2.02); baseline LE8 corresponded to an OR of 0.781 per 10 points (0.661–0.922). The joint Cox model gave similar estimates (HR 2.14, 1.23–3.71 and 1.40, 0.99–1.98). Because baseline LE8 and trajectory membership were strongly correlated (between-group SMD 3.37; $R^2 = 0.556$), these joint estimates should be regarded as descriptive. The incremental C-index for trajectories over baseline LE8 alone was 0.005, and the NRI and IDI were near zero, whereas the likelihood-ratio tests were significant for both additions; this discrepancy reflects the difference between global fit and rank-based discrimination, and bootstrap-corrected estimates confirmed the small increments (eTable 9). Within baseline-LE8 tertiles, trajectories retained associations in the lowest (low vs high: HR 4.96, 95% CI 1.54–16.03) and middle (moderate 1.75, 1.08–2.82; low 4.92, 1.16–20.93, an estimate based on only 12 low-group members) tertiles, and the within-tertile C-index increments remained small (ΔC = 0.015, 0.012, and 0.012; eTable 11). Therefore, even among individuals with a similar baseline level, trajectory membership carried some prognostic information but little additional discrimination.

3.4 Within-person change in LE8 (ΔLE8)

The individual change in LE8 (ΔLE8 = 2015 score − 2011 score; n=3,761; 283 events) was not associated with CVM in the overall analysis (HR 1.03 per 10 points, 95% CI 0.92–1.15). However, after additional adjustment for the baseline score, a 10-point increase in LE8 was associated with a lower risk of CVM (HR 0.82, 95% CI 0.72–0.92, P = 0.001; baseline LE8 HR 0.62, 0.55–0.70). Stratified by baseline tertile, the adjusted ΔLE8 effect was similar in the lowest and middle tertiles (HR 0.80, 0.67–0.95 and 0.77, 0.63–0.96) and attenuated in the highest tertile (0.95, 0.72–1.27; eTable 13). This pattern is consistent with some regression to the mean but not with regression to the mean being the sole explanation. Taken together with the negligible class increment described above, these results indicate that within-person change, rather than class

membership, is the component of repeated measurement that carries information beyond a single baseline score: among individuals with the same baseline score, a greater decline (or less improvement) in LE8 was associated with a higher risk of CVM.

3.5 Sensitivity analysis

All 25 sensitivity analyses (S1–S17, including sub-variants; definitions in the Supplementary Material) were directionally consistent with the primary results (Table 3). The LOCF variant (S2) reproduced the primary estimates exactly, because the interpolation method for 2013 did not alter the trajectory classification. Excluding the blood-pressure component attenuated the association to HR 1.96 (95% CI 1.27–3.04). Across the remaining variants, the HRs of the persistent low group ranged from 3.5 to 6.1 (all $P<0.001$). Rescoring PA and nicotine exposure with the official AHA thresholds (S13a) re-assigned 328 participants (7.1%) to different trajectory classes but left the estimates essentially unchanged (HR 1.89, 95% CI 1.42–2.53 and 3.75, 95% CI 2.63–5.35). In the joint rescoring analysis (S15), in which PA, nicotine, and sleep were simultaneously re-scored with official thresholds, the class enumeration again selected three classes, 459 participants (9.9%) were re-assigned, and the estimates were essentially unchanged (HR 1.93, 95% CI 1.44–2.59 and 3.79, 95% CI 2.64–5.42). Of note, the class composition changed materially under the official thresholds (moderate class 45.1% to 11.4%; low class 10.7% to 40.7%; score correlation $r = 0.95–0.98$; trajectory agreement 90.1%, $\kappa = 0.832$; eTable 14), so S15 demonstrates the robustness of the association across scorings rather than the interchangeability of the classes. Cluster-robust standard errors (S13c), official sleep thresholds (S13d: HR 1.97, 95% CI 1.46–2.65 and 3.96, 95% CI 2.77–5.67), exclusion of the 18 unclassifiable events (S14: OR 1.93, 95% CI 1.44–2.62 and 4.08, 95% CI 2.78–5.98), and the PA-excluded score (S13e: HR 1.96, 95% CI 1.47–2.60 and 4.22, 95% CI 2.96–6.02) likewise did not change the conclusions.

With stabilized inverse-probability weights (mean 0.99, SD 0.69; un-stabilized weights: mean 3.60, maximum 46.3, 2.6% >10; 99th-percentile truncation at 3.59; eTable 12), the estimates were similar (HR 1.91, 95% CI 1.36–2.69 and 3.53, 95% CI 2.28–5.47 after truncation). Classification quality and estimates stratified by the number of available LE8 components were consistent across strata (mean posterior probabilities 0.86–0.87; HR 4.73 and 2.93 in the two larger strata; eTable 10), indicating that the component-missingness pattern does not drive the results.

When CVM was defined as the co-occurrence of heart disease and stroke without hypertension, the HR of the persistent low group was 4.26 (95% CI 1.83–9.92), and when a diabetes-anchored

definition was used, the HR was 6.13 (95% CI 3.84–9.80). Both definitions yielded stronger rather than weaker associations, arguing against mechanical inflation through the hypertension component; these two analyses had limited events (50 and 164) and should be regarded as exploratory. A cross-analysis combining the blood-pressure-excluded score with the hypertension-free outcome (S16; 38 events) gave HRs of 2.21 (95% CI 1.07–4.56) for persistent moderate and 1.86 (95% CI 0.50–6.87) for persistent low (P for trend = 0.067); removing the blood-pressure component attenuated the estimate regardless of whether the outcome contained hypertension.

At baseline, 879 of the 4,547 at-risk participants (19.4%) had measured blood pressure in the hypertensive range (SBP ≥140 mmHg, DBP ≥90 mmHg, or antihypertensive treatment) despite denying a hypertension diagnosis; the proportions were 7.6%, 23.8%, and 50.7% across the persistent high, moderate, and low groups, and measured hypertension was itself associated with CVM (adjusted OR 2.76, 95% CI 2.11–3.59; eTable 19). Excluding these participants (S17a; n=3,110; 171 events) attenuated but did not remove the association (HR 1.54, 95% CI 1.11–2.15 and 3.02, 95% CI 1.85–4.91). After the additional exclusion of participants with measured hypertension at 2015 (S17b; n=2,597; 98 events), the association remained significant (HR 1.17, 95% CI 0.76–1.80 and 2.57, 95% CI 1.27–5.23; P for trend = 0.043). The E-value was 6.96 (lower CI bound 4.70; based on the Cox HR of 3.75), indicating that unmeasured confounding would need to exceed an RR of 6.96 to explain the association.

To clarify which disease transitions drive the CVM association, the three constituent diseases were examined separately (Model 3, Cox framework). The association was strongest for incident hypertension (1,110 events; HR 2.21, 95% CI 1.92–2.54 for persistent moderate and 4.12, 95% CI 3.43–4.94 for persistent low), but incident heart disease (532 events; HR 1.33, 1.10–1.60 and 1.51, 1.14–2.01) and incident stroke (174 events; HR 2.15, 1.49–3.12 and 3.93, 2.49–6.19) were also elevated. In the component-contribution decomposition (eTable 16), entering the baseline blood-pressure component alone attenuated the HR of the persistent low group by 48% (4.34 to 2.14) and the BMI component by 13%, with negligible contributions from PA, nicotine, and sleep; entering all seven components attenuated the association by 57%. Furthermore, stratification by baseline blood-pressure component score showed no significant interaction (P=0.716); within the lowest stratum (<50 points; 171 events), the HR of the persistent low group was 3.11 (95% CI 1.88–5.14), and the estimates in the middle (50–74; 42 events) and high (≥75; 64 events) strata were directionally consistent but did not reach statistical significance (1.84, 0.65–5.19 and 2.22, 0.85–5.82; eTable 5).

Sensitivity	n	Events	Moderate HR (95% CI)	Low HR (95% CI)	P trend
Cox model (complete cases)†	3848	278	1.90 (1.43–2.52)	3.75 (2.63–5.34)	<0.001
S1 No interpolation	3278	243	2.20 (1.64–2.95)	3.48 (2.26–5.35)	<0.001
S2 LOCF	3848	278	1.90 (1.43–2.52)	3.75 (2.63–5.34)	<0.001
S7 Complete 3-wave	3161	234	1.81 (1.34–2.46)	3.94 (2.67–5.81)	<0.001
S9 Chinese BMI	3848	278	1.96 (1.47–2.61)	3.92 (2.74–5.60)	<0.001
S9b Sleep 6–8h	3848	278	1.99 (1.50–2.62)	3.75 (2.60–5.41)	<0.001
S10 Without BP (6 items)	3501	257	1.52 (1.16–1.99)	1.96 (1.27–3.04)	<0.001
S3 Competing risk	3848	278	1.86 (1.41–2.46)	3.52 (2.51–4.93)	—
S4 Reverse causation	3215	148	2.09 (1.43–3.06)	3.96 (2.40–6.55)	<0.001
S5 E-value	—	—	—	6.96 (4.70)	—
S6 Multiple imputation	4547	330	1.94 (1.50–2.52)	3.70 (2.67–5.13)	<0.001
S8a Including 2015 events	3916	331	2.02 (1.56–2.63)	3.70 (2.66–5.15)	<0.001

S8b Including 2013 events	3916	347	2.03 (1.57–2.62)	3.57 (2.58–4.93)	<0.001
S8c Including 2018 two-disease	3847	313	2.14 (1.63–2.81)	3.60 (2.56–5.08)	<0.001
S11 Heart disease+stroke (no HTN)	3917	50	2.04 (1.01–4.10)	4.26 (1.83–9.92)	<0.001
S12 Diabetes-anchored (Di Angelantonio)	3778	164	2.66 (1.79–3.95)	6.13 (3.84–9.80)	<0.001
S13a Official AHA scoring (PA + nicotine)	3853	278	1.89 (1.42–2.53)	3.75 (2.63–5.35)	<0.001
S13b Inverse-probability weighting	3848	278	2.00 (1.41–2.84)	3.61 (2.28–5.71)	<0.001
S13c Cluster-robust SE (community)	3848	278	1.90 (1.44–2.50)	3.75 (2.67–5.27)	<0.001
S13d Official AHA sleep scoring	3848	278	1.97 (1.46–2.65)	3.96 (2.77–5.67)	<0.001
S14 Excluding 18 unclassifiable events (logistic)	3848	312	1.93 (1.44–2.62)	4.08 (2.78–5.98)	<0.001
S13e Excluding PA	3848	278	1.96 (1.47–2.60)	4.22 (2.96–6.02)	<0.001

(6 components)					
S15 Joint official AHA scoring (PA + nicotine + sleep)	3848	278	1.93 (1.44–2.59)	3.79 (2.64–5.42)	<0.001
S16 No-BP score × hypertension-free outcome	3501	38	2.21 (1.07–4.56)	1.86 (0.50–6.87)	0.067
S17a Excluding baseline measured hypertension (Cox)	3110	171	1.54 (1.11–2.15)	3.02 (1.85–4.91)	<0.001
S17b Excluding baseline + 2015 measured hypertension (Cox)	2597	98	1.17 (0.76–1.80)	2.57 (1.27–5.23)	0.043

Table 3. Sensitivity analyses (persistent low vs high; Model 3). †The primary-analysis row reports the Model 3 complete-case sample (full sample n=4,547/330 events; see the S6 multiple-imputation row). All rows are Cox-based except S14 (logistic; marked in the row label) and S5 (E-value); trajectories were re-estimated (LCGM) in all score-definition variants (S1, S2, S7, S9, S9b, S10, S13a, S13d, S13e, S15, S16). Full definitions of S1–S17 are provided in the Supplementary Material (Sensitivity Analysis Definitions).

3.6 Subgroup analysis

In the assessment of potential effect modifiers for the association between trajectory groups and CVM, sex and age group did not significantly modify the association (P for interaction = 0.297

and 0.685, respectively; Figure 5). The HR of the persistent low group was 3.70 (95% CI 2.25–6.10) in women and 3.32 (2.00–5.54) in men, and 3.97 (2.35–6.73) in participants aged 45–59 years and 3.44 (2.13–5.56) in those aged ≥60 years. However, the association was nominally modified by residence (P for interaction = 0.023; FDR-adjusted P = 0.069). The HR of the persistent low group was 6.17 (95% CI 2.54–14.97) in urban participants and 3.51 (2.38–5.17) in rural participants, and the HR of the persistent moderate group was likewise elevated in urban areas (4.53, 2.16–9.50). With 330 events, only interaction effect sizes of approximately ≥2.5-fold (ratio of subgroup HRs; ≥3.1-fold for residence) were detectable with 80% power (eTable 15). Given this limitation, the multiple testing across three subgroup analyses, and the small number of urban events (62), all subgroup analyses, including the residence interaction, should be regarded as exploratory (detailed results in eTable 3).

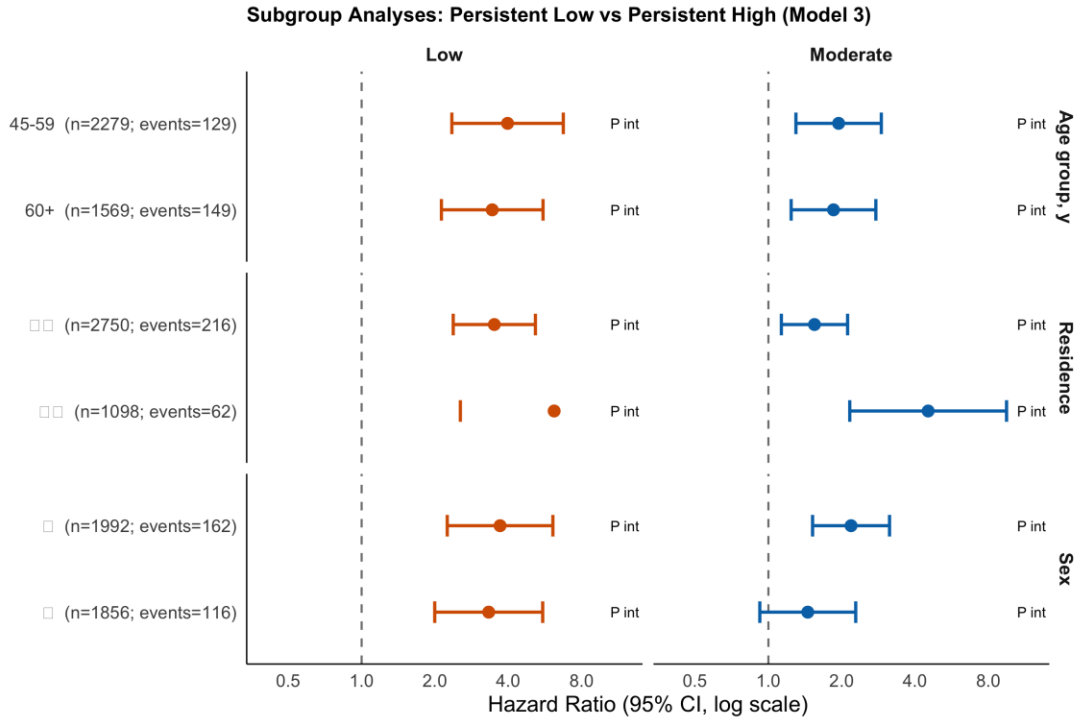


Figure 5. Subgroup forest plot (sex/age group/residence; persistent low vs high; Model 3). The residence interaction was nominally significant (P = 0.023; FDR-adjusted P = 0.069) and should be regarded as exploratory given the small number of urban events (62); no heterogeneity was observed by sex or age group.

3.7 Split-sample internal validation and independent NHANES cohort validation

The derivation sample comprised 3,183 participants (231 events, 7.3%) and the validation sample 1,364 (99 events, 7.3%), with class proportions preserved by design. In the derivation sample, the

class enumeration again selected three classes (BIC 68,100.1 versus 68,612.2 for two classes; entropy 0.890; mean posterior probability 0.856; minimum class 11.3%), and the four- and five-class solutions were again rejected for classes below 5%; the class mean trajectories (baseline 63.4/78.5/89.0) were close to the full-sample values (62.1/77.4/89.2). Assignment by the derivation model agreed with the full-sample classification in 93.1% of the validation sample (κ = 0.886). The Model 3 HRs were 1.84 (95% CI 1.30–2.60) for the persistent moderate and 3.51 (2.30–5.36) for the persistent low group in the derivation sample, and 1.35 (0.78–2.34) and 4.20 (2.39–7.39) in the validation sample (per-category HR 2.10, P for trend < 0.001; logistic OR for the persistent low group 4.61, 2.52–8.46). The derivation Cox model applied to the validation sample had a C-index of 0.662 (SE 0.032), versus 0.590 (0.031) for covariates alone, and the logistic model calibrated adequately (observed/expected 1.00; slope 0.92; intercept −0.20; eTable 20, eFigure 2).

In NHANES, the LE8 score showed graded associations with all-cause mortality (per 10 points: weighted HR 0.73, 95% CI 0.67–0.79; low [<50] versus high [≥80]: 3.87, 2.58–5.80) and with cardiovascular mortality (0.67, 0.59–0.76 and 4.31, 2.49–7.47; 295 cardiovascular deaths). In the cardiovascular disease–free subcohort the estimate for all-cause mortality was 0.77 (0.65–0.92); cardiovascular deaths (n = 46) were too few for stable estimates (eTable 21).

4. Discussion

This study demonstrated that the level of a modified 7-component LE8 score was independently and inversely associated with the risk of incident CVM in Chinese middle-aged and older adults. Three trajectories were identified over 2011–2015: persistent high (44.3%), persistent moderate (45.1%), and persistent low (10.7%). Compared with the persistent high group, the persistent moderate and persistent low groups had higher risks of CVM, with multiple-adjusted ORs of 1.95 (95% CI 1.45–2.62) and 4.04 (2.77–5.89), and the risk increased as the trajectory level decreased (OR 0.60 per 10 points). The findings remained consistent across all 25 sensitivity analyses. Notably, the trajectories added little discriminative value beyond a single baseline score (C-index increment 0.005), whereas within-person change carried prognostic information after baseline adjustment (HR 0.82 per 10 points).

A previous study by Zhong et al. demonstrated that 10-year LE8 trajectories predicted incident hypertension in the Kailuan cohort [9], and Li et al. reported that changes in LE8 were related to cardiovascular disease risk [10]. A UK Biobank study showed that sleep patterns were associated with the progression of cardiovascular multimorbidity [11], and transitions in sleep duration and

quality were related to incident cardiovascular disease in CHARLS [8]. A nationwide Chinese cohort study (the 4C Study) linked higher single-measurement LE8 scores to lower cardiovascular-kidney-metabolic syndrome stage-related risk [12], in the same direction as our per-10-point estimate. The present study extends this literature in a national, population-representative sample, uses CVM rather than a single disease as the outcome, and, for the first time, directly quantifies the increment of LE8 trajectories over a single baseline measurement.

With respect to the incremental value of repeated measurement, two findings should be highlighted. On the one hand, for risk stratification, a single baseline LE8 score already captures most of the prognostic information; the trajectory classes added minimal discriminative value beyond the baseline score (C-index increment 0.005; NRI/IDI near zero), and the moderate class no longer reached statistical significance after baseline adjustment. On the other hand, repeated measurement retains a narrower but real value: among individuals with the same baseline score, within-person change carried prognostic information, and trajectory membership remained associated with risk within baseline-LE8 tertiles. This apparent discrepancy can be explained by the level-versus-change structure of the data. The trajectory classes were determined almost entirely by the baseline level (between-group SMD 3.37; $R^2 = 0.556$), and all three classes declined slowly and in parallel (per-2-year changes of −1.1 to −2.4 points), so class membership adds little information once the baseline score is known. In this sense, the "persistent" label denotes relative level stability over the 4-year window rather than distinct dynamic behavior. Our trajectory window (4 years, three waves) is shorter than that of the Kailuan study (10 years) [9]; longer windows allow more pronounced divergence between individuals and may yield greater incremental value for class-based trajectories than observed here.

These findings have clinical implications. The persistent low group formed about 11% of this examination-compliant cohort (not a design-weighted national prevalence) and had a 5-year CVM cumulative incidence of 14.6%, which was 3.3 times that of the persistent high group (4.4%), with an absolute risk difference of 10.2 percentage points. This group was older, more often male, and had more cardiometabolic problems, and these characteristics are readily identifiable in clinical practice. For risk stratification and the targeting of preventive interventions, a single baseline LE8 measurement identifies these individuals nearly as well as repeated measurement. Repeated measurement is more useful for a narrower purpose, namely detecting within-person decline among people whose health is changing, particularly at similar baseline levels. These findings apply to middle-aged and older adults free of baseline

cardiovascular disease who are compliant with periodic health examinations; extrapolation to the general population or to populations with established cardiovascular disease requires caution.

The association between LE8 and CVM may be associated with the pathophysiology of cardiovascular risk accumulation. A persistently low LE8 score means long-term exposure to smoking, physical inactivity, obesity, and elevated blood pressure, glucose, and lipids. These factors are believed to damage the vasculature through shared pathways, including endothelial dysfunction, atherosclerosis, and chronic inflammation [3,27], although these mechanisms are hypothesized rather than directly tested here. Blood pressure played an important role in our study. When the blood-pressure component was removed from LE8 and the trajectories were re-estimated, the HR decreased from 3.75 to 1.96, and in the component-contribution decomposition the baseline blood-pressure component alone attenuated the association by 48% and BMI by 13%. When CVM was defined without hypertension, or with a diabetes-anchored definition, the associations became stronger rather than weaker. In addition, adiposity-related phenotypes are independently associated with cardiometabolic multimorbidity [28,29]. These findings support the view that multiple risk factors act together, with the blood-pressure dimension as the dominant pathway.

Hypertension-anchored outcome: circularity versus pathway mediation (consolidated assessment). Because hypertension anchors the outcome, the exposure contains a blood-pressure component, and baseline exclusion relied on self-report, all relevant evidence was consolidated in Table 4, which maps the three candidate explanations to their supporting analyses and verdicts. Two facts are decisive. First, removing the blood-pressure component from the exposure attenuated the estimate regardless of whether the outcome contained hypertension (S10: HR 1.96 with the standard outcome; S16: HR 1.86 with the hypertension-free outcome), whereas removing hypertension from the outcome did not reduce, and if anything strengthened, the association (S11: 4.26; S12: 6.13). This pattern is consistent with pathway mediation rather than definitional circularity. Second, the detection-bias channel is real but bounded: 19.4% of the risk set had undetected measured hypertension at baseline (overlapping the low-BP-score groups), yet the exclusion of these participants (S17a) and, additionally, of those measured hypertensive at 2015 (S17b) still left a significant graded association (HR 3.02 and 2.57). We therefore conclude that the blood-pressure attenuation mainly reflects pathway mediation, with a bounded contribution from detection bias, and that residual circularity is unlikely to explain the observed associations, although the imprecision of the small-sample cross-analysis (S16) means that a modest circular contribution cannot be entirely excluded.

Explanation	Key supporting analyses	Verdict
Pathway mediation (BP is the dominant pathway)	S10 (BP-excluded score: HR 3.75→1.96); S16 (no-BP score × no-HTN outcome: 1.86); eTable 16 (baseline BP component alone attenuates low-group HR by 48%); single-disease analyses (incident hypertension strongest)	**Supported (primary explanation)**
Circularity amplification (hypertension in the outcome definition)	S11 (no-HTN outcome: HR 4.26, stronger not weaker); S12 (diabetes-anchored: 6.13); S16 (attenuation persists without HTN in outcome); baseline BP-component stratification (interaction P=0.716; eTable 5)	**Not supported**
Detection/ascertainment bias (undetected measured hypertension)	19.4% undetected at baseline; associated with CVM (OR 2.76); S17a exclusion: HR 3.02; S17b (also excluding 2015 measured HTN): 2.57, both significant	**Present but bounded; does not explain the association**

Table 4. Circularity versus pathway mediation for the hypertension-anchored outcome: candidate explanations, supporting analyses, and verdicts.

The strengths of this study include its national representativeness, prospective design, comprehensive sensitivity framework (25 analyses), and the direct assessment of the incremental value of repeated measurement, which has rarely tested in trajectory research. In the split-sample analysis the association was reproduced in the validation set (HR 4.20 for the persistent low group; C-index 0.662; calibration slope 0.92), and the same scoring code showed graded associations of the LE8 score with all-cause and cardiovascular mortality in NHANES. However, our study has several limitations that should be acknowledged. (1) CHARLS has no food-frequency questionnaire, so the diet component was excluded, and the scoring of PA, nicotine, and sleep deviates from the official AHA definition; joint rescoring with official thresholds (S15)

and the exclusion of PA (S13e) left the results essentially unchanged, but comparability with official 8-component scores remains limited by data constraints. (2) PA was collected in only half of households; the mean of available components was used, and classification quality and estimates were consistent across component-availability strata (eTable 10). (3) Blood glucose and lipids in 2013 were interpolated from the 2011 and 2015 measurements; the no-interpolation and LOCF variants gave consistent results. (4) The outcome was based on self-reported physician diagnoses; self-reported hypertension shows limited agreement with measured blood pressure in this population [23], and misclassification usually attenuates the associations. In addition, because hypertension anchors the outcome and the LE8 score contains a blood-pressure component, part of the observed association reflects the blood-pressure pathway: removing the blood-pressure component halved the estimate (S10; HR 1.96), and the cross-analysis combining the blood-pressure-excluded score with the hypertension-free outcome, although directionally consistent, was based on only 38 events (S16; P for trend 0.067), leaving the magnitude of the non-blood-pressure contribution imprecisely estimated. (5) Events were confirmed only at the 2020 wave (two event-ascertainment time points), so the analysis is effectively a comparison of 5-year cumulative risk; outcome information was complete for 89.4% of the risk set, and none of the 314 participants who died had a confirmed CVM before death, so events occurring shortly before death may have been missed, biasing the estimates toward underestimation. (6) The trajectory window was only 4 years, and participants who died or dropped out during this window were excluded, which may introduce survival selection; in addition, the Hessian of the fitted LCGM was numerically singular, so the parametric uncertainty bands of Figure 2 should be read as indicative. (7) CHARLS sampling weights were not applied, because the longitudinal subsample is not a simple probability subset of the baseline sample; cluster-robust standard errors gave similar results. (8) The included participants were healthier and more examination-compliant than the excluded ones; inverse-probability weighting with stabilized weights showed that this selection had limited influence (eTable 12), and the results apply to relatively healthy, examination-compliant middle-aged and older adults rather than to the general population. (9) The subgroup analyses had limited power and should be regarded as exploratory. (10) This observational study was not pre-registered, and the analysis plan was developed post hoc.

5. Conclusion

In summary, this study indicated that lower levels of the modified 7-component LE8 score were significantly associated with a higher risk of incident CVM in Chinese middle-aged and older adults: the persistent low trajectory had a 5-year cumulative incidence of 14.6% versus 4.4% for

the persistent high trajectory (absolute difference 10.2 percentage points; adjusted OR 4.04). Notably, over the short 4-year window the trajectory classes mainly reflected baseline level, and trajectories offered minimal additional discriminative value beyond a single baseline score, whereas within-person change itself carried prognostic information (HR 0.82 per 10 points). These findings provide a new insight into the interpretation of repeated cardiovascular health measurements and emphasize that a single baseline LE8 measurement largely suffices for risk stratification, while repeated measurement adds value mainly for detecting within-person change. Further studies with longer follow-up and more measurement waves are needed to validate these findings and to explore the precise mechanisms involved, which could pave the way for more targeted prevention of cardiovascular multimorbidity.

6. Declarations

Ethics approval and consent to participate: CHARLS was approved by the Biomedical Ethics Review Committee of Peking University (IRB00001052-11015); all participants provided written informed consent. This study is a secondary analysis of publicly available data. The NHANES protocol was approved by the National Center for Health Statistics (NCHS) Research Ethics Review Board (Protocol #2011–17); all participants provided written informed consent, and the public-use linked mortality files are de-identified.

Consent for publication: Not applicable. This manuscript does not contain any individual person's data in any form (including individual details, images, or videos).

Availability of data and materials: CHARLS data are publicly available upon registration and data-use agreement at http://charls.pku.edu.cn/. NHANES public-use data and the linked mortality files are available from the NCHS website (https://wwwn.cdc.gov/nchs/nhanes/). Analysis code is available from the corresponding author upon reasonable request and will be deposited in a public repository (e.g., Zenodo or GitHub) upon acceptance.

Competing interests: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding: This study received no external funding.

Declaration of Generative AI and AI-assisted technologies in the writing process: During the preparation of this work the authors used GLM (Z.ai) in order to assist with language editing and

the preparation of statistical analysis code. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Authors' contributions: Yanjiao Li: Conceptualization, Methodology, Formal analysis, Writing – original draft. Jiao Xue: Data curation, Validation, Writing – review & editing. Guojuan Liu: Investigation, Supervision, Project administration, Writing – review & editing. All authors read and approved the final manuscript.

References

[1] GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2020;396(10258):1204-1222. DOI: 10.1016/S0140-6736(20)30925-9

[2] Di Angelantonio E, Kaptoge S, Wormser D, et al. Association of cardiometabolic multimorbidity with mortality. JAMA. 2015;314(1):52-60. DOI: 10.1001/jama.2015.7008

[3] Lloyd-Jones DM, Allen NB, Anderson CAM, et al. Life's Essential 8: updating and enhancing the American Heart Association's construct of cardiovascular health: a presidential advisory from the American Heart Association. Circulation. 2022;146(5):e18-e43. DOI: 10.1161/CIR.0000000000001078

[4] Lloyd-Jones DM, Hong Y, Labarthe D, et al. Defining and setting national goals for cardiovascular health promotion and disease reduction: the American Heart Association's strategic Impact Goal through 2020 and beyond. Circulation. 2010;121(4):586-613. DOI: 10.1161/CIRCULATIONAHA.109.192703

[5] Chai L, Zhang K, Zhang Y, Wang W, Zhang D, Fan J. Associations between Life's Essential Eight cardiovascular health metrics and cardiovascular mortality risk across frailty statuses: evidence from a UK Biobank cohort study. Front Public Health. 2025;13:1508274. DOI: 10.3389/fpubh.2025.1508274

[6] Sebastian SA, Shah Y, Paul H, Arsene C. Life's Essential 8 and the risk of cardiovascular disease: a systematic review and meta-analysis. Eur J Prev Cardiol. 2025;32(5):358-373. DOI: 10.1093/eurjpc/zwae280

[7] Wu F, Yang J, Zhang Y, Peng L. Exploring the impact of individual components of the Life's Essential 8 on the relationship between atherogenic index of plasma and adverse cardiovascular events: a population-based cohort study in China. Front Physiol. 2025;16:1538938. DOI: 10.3389/fphys.2025.1538938

[8] Zhang Y, Guo J, Hu X, Xie H. Transition of nighttime sleep duration and sleep quality with incident cardiovascular disease among middle-aged and older adults: results from a national cohort study. Arch Public Health. 2025;83(1):91. DOI: 10.1186/s13690-025-01577-5

[9] Zhong J, Jiang J, Guo L, et al. 10-year trajectory of Life's Essential 8 and incident hypertension: a community-based cohort study. Lipids Health Dis. 2024;23(1):278. DOI: 10.1186/s12944-024-02257-z

[10] Li W, Xing A, Lamballais S, Xu W, Chen S, Zhou S. Changes in Life's Essential 8 and risk of cardiovascular disease in Chinese people. Eur J Public Health. 2024;34(4):766-773. DOI: 10.1093/eurpub/ckae063

[11] Zhang J, Chen L, Zhang S, Cai M, Zou H, Vaughn MG. Associations of sleep patterns with dynamic trajectory of cardiovascular multimorbidity and mortality: a multistate analysis of a large cohort. J Am Heart Assoc. 2023;12(19):e029463. DOI: 10.1161/JAHA.123.029463

[12] Li M, Xu M, Ding Y, Lin H, Qin G, Wang T, et al. Life's Essential 8 cardiovascular health, cardiovascular-kidney-metabolic syndrome stages, and incident cardiovascular events: a nationwide 10-year prospective cohort study in China. Cardiovasc Diabetol. 2025;24(1):197. DOI: 10.1186/s12933-025-02735-3

[13] Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. Annu Rev Clin Psychol. 2010;6:109-138. DOI: 10.1146/annurev.clinpsy.121208.131413

[14] Nguena Nguefack HL, Pagé MG, Katz J, Choinière M. Trajectory modelling techniques useful to epidemiological research: a comparative narrative review of approaches. Clin Epidemiol. 2020;12:1205-1222. DOI: 10.2147/CLEP.S265287

[15] 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):61-68. DOI: 10.1093/ije/dys203

[16] Chen X, Crimmins E, Hu PP, et al. Venous Blood-Based Biomarkers in the China Health and Retirement Longitudinal Study: Rationale, Design, and Results From the 2015 Wave. Am J Epidemiol. 2019;188(11):1871-1877. DOI: 10.1093/aje/kwz170

[17] von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007;4(10):e296. DOI: 10.1371/journal.pmed.0040296

[18] Higgins JPT, Morgan RL, Rooney AA, et al. A tool to assess risk of bias in non-randomized follow-up studies of exposure effects (ROBINS-E). Environ Int. 2024;186:108602. DOI: 10.1016/j.envint.2024.108602

[19] Bull FC, Al-Ansari SS, Biddle S, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med. 2020;54(24):1451-1462. DOI: 10.1136/bjsports-2020-102955

[20] Proust-Lima C, Philipps V, Liquet B. Estimation of extended mixed models using latent classes and latent processes: the R package lcmm. J Stat Softw. 2017;78(2):1-56. DOI: 10.18637/jss.v078.i02

[21] Sakakibara BM, Obembe AO, Eng JJ. The prevalence of cardiometabolic multimorbidity and its association with physical activity, diet, and stress in Canada: evidence from a population-based cross-sectional study. BMC Public Health. 2019;19(1):1361. DOI: 10.1186/s12889-019-7682-4

[22] Bao Y, Chen H, Zhu B, Tang M, Wu Q, Jia Y, Jiang Y. Triglyceride-glucose index as a predictor of cardiovascular multimorbidity: a prospective cohort study. Front Cardiovasc Med. 2025;12:1647444. DOI: 10.3389/fcvm.2025.1647444

[23] Wu J, Chen D, Li C, Wang Y. Agreement between self-reported and objectively measured hypertension diagnosis and control: evidence from a nationally representative sample of community-dwelling middle-aged and older adults in China. Arch Public Health. 2024;82(1):245. DOI: 10.1186/s13690-024-01456-5

[24] Fine JP, Gray RJ. A proportional hazards model for the subdistribution of the competing risk. J Am Stat Assoc. 1999;94(446):496-509. DOI: 10.1080/01621459.1999.10474144

[25] van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Softw. 2011;45(3):1-67. DOI: 10.18637/jss.v045.i03

[26] VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. Ann Intern Med. 2017;167(4):268-274. DOI: 10.7326/M16-2607

[27] Libby P, Buring JE, Badimon L, et al. Atherosclerosis. Nat Rev Dis Primers. 2019;5(1):56. DOI: 10.1038/s41572-019-0106-z

[28] Yu B, Jia S, Sun T, Liu J, Jin J, Zhang S, Xiao Q, Dong H, Ou Y. Sarcopenic obesity is associated with cardiometabolic multimorbidity in Chinese middle-aged and older adults: a cross-sectional and longitudinal study. J Nutr Health Aging. 2024;28(10):100353. DOI: 10.1016/j.jnha.2024.100353

[29] Li L, Xu J, Huang L, et al. The synergy between Life's Essential 8 and muscle strength on cardiovascular disease risk. Front Med (Lausanne). 2025;12:1628066. DOI: 10.3389/fmed.2025.1628066

[30] Dyrstad SM, Hansen BH, Holme IM, Anderssen SA. Comparison of self-reported versus accelerometer-measured physical activity. Med Sci Sports Exerc. 2014;46(1):99-106. DOI: 10.1249/MSS.0b013e3182a0595f

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryLE8ENResubmit.docx