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.

Joint trajectories of depressive symptoms and frailty and 2-year stroke recurrence and all-cause mortality among older Chinese stroke survivors: a national cohort study

Ping Yu
Rehabilitation Medicine Center, Fujian Provincial Governmental Hospital, No. 67 Guping Road, Gulou District, Fuzhou 350003, Fujian, China

Lisheng Le
Rehabilitation Medicine Center, Fujian Provincial Governmental Hospital, No. 67 Guping Road, Gulou District, Fuzhou 350003, Fujian, China

Jianhao Chen
Shengli Clinical Medical College of Fujian Medical University, Fuzhou, China

Changyue Zheng
Rehabilitation Medicine Center, Fujian Provincial Governmental Hospital, No. 67 Guping Road, Gulou District, Fuzhou 350003, Fujian, China

Yiting Tang
m18106071318@163.com

The First Clinical Medical College of Fujian Medical University, No. 1 Xuefu North Road, Fuzhou University Town, Minhou County, Fuzhou 350122, Fujian, China

Research Article

Keywords: depressive symptoms, frailty, joint trajectory, group-based trajectory modelling, post-stroke depression, stroke recurrence, mortality, older adults, population-based cohort, CHARLS

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11077699/v1

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

Additional Declarations: No competing interests reported.

Joint trajectories of depressive symptoms and frailty and 2-year stroke recurrence and all-cause mortality among older Chinese stroke survivors: a national cohort study

Authors: Ping Yu[1], Lisheng Le[1],*, Jianhao Chen[2],[3], Changyue Zheng[1],†, Yiting Tang[4],†

Affiliations:

[1] Rehabilitation Medicine Center, Fujian Provincial Governmental Hospital, No. 67 Guping Road, Gulou District, Fuzhou 350003, Fujian, China

[2] Shengli Clinical Medical College of Fujian Medical University, Fuzhou, China

[3] Rehabilitation Medicine Center, Fuzhou University Affiliated Provincial Hospital, Fuzhou, China

[4] The First Clinical Medical College of Fujian Medical University, No. 1 Xuefu North Road, Fuzhou University Town, Minhou County, Fuzhou 350122, Fujian, China

*** Co-first authors:** Ping Yu and Lisheng Le contributed equally to this work and share first authorship.

† Co-corresponding authors: Yiting Tang and Changyue Zheng contributed equally to this work and share corresponding authorship.

† Correspondence: Yiting Tang, m18106071318@163.com; Changyue Zheng, 46138447@qq.com

Study design: Retrospective cohort study using a nationally representative population sample (secondary analysis of the China Health and Retirement Longitudinal Study, CHARLS).

Word count (main text, Background to Conclusions): 5,226

Number of figures: 5

Number of tables: 4

Number of additional files: 2

Abstract

Background: Depressive symptoms and frailty are among the most common and disabling conditions in older adults after stroke, and both are priorities for population-level chronic disease management in ageing societies. Most studies describe them separately and at a single time point, so whether their joint trajectories mark the risk of recurrence and death is unknown.

Methods: We used three waves of the China Health and Retirement Longitudinal Study (CHARLS, 2013–2020); the cohort comprised participants with a physician-diagnosed stroke recorded in 2018. Joint trajectories of depressive symptoms (CESD-10) and frailty (frailty index) across 2013, 2015 and 2018 were identified by group-based trajectory modelling with a parallel-process specification. Risk ratios (RR) were estimated by multivariable log-binomial regression, with robust Poisson and logistic sensitivity analyses and multinomial logit regression for the competing risk of death; missing covariates were handled by multiple imputation by chained equations (MICE, m = 10) with Rubin's rules.

Results: Among 394 stroke survivors (mean age 65.6 years, 51.0% men), two joint trajectory classes were identified: a low-rising class (n = 222, 56.3%) and a high-increasing class (n = 172, 43.7%) with elevated, rising depressive symptoms and frailty. Class 2 had a higher risk of stroke recurrence than class 1 (RR 3.09, 95% CI 1.83–5.23; P < 0.001), robust across complete-case (RR 2.66), robust Poisson (RR 2.70), high-certainty membership (RR 2.59) and logistic (OR

3.29) analyses; 2-year absolute risk was 23.3% versus 8.6% (risk difference 14.7 percentage points). Adjustment for the 2013 single-timepoint CESD-10 and frailty index attenuated but retained the contrast (RR 2.19, 95% CI 1.13–4.25; P = 0.020) while the single-timepoint terms became non-significant; discrimination was similar for the joint trajectory and simpler exposures (area under the curve 0.696 vs 0.677), so added predictive value is not established. Mortality findings were exploratory (adjusted RR 2.71, 95% CI 1.07–6.83; unadjusted RR 2.21, 0.89–5.50; 19 deaths).

Conclusions: A jointly high and increasing depressive-symptom/frailty trajectory identifies older stroke survivors at roughly threefold risk of recurrence within two years; the association was robust but could not be separated from the 2018 level, and it remains hypothesis-generating pending external validation in an independent population.

Keywords: depressive symptoms; frailty; joint trajectory; group-based trajectory modelling; post-stroke depression; stroke recurrence; mortality; older adults; population-based cohort; CHARLS

Background

Stroke is the leading cause of death and acquired disability in China, and the number of community-dwelling survivors continues to grow as the population ages (1). These survivors remain at high risk of recurrent vascular events and premature death, so their long-term health represents a substantial and expanding public health burden. Population-level strategies for reducing that burden have traditionally emphasised pharmacological and lifestyle targets, but a growing body of evidence indicates that two widely prevalent and at least partly modifiable health states — depressive symptoms and frailty — carry considerable prognostic weight.

Two such phenotypes have attracted particular attention. Post-stroke depression (PSD) is highly prevalent in Chinese stroke survivors, with estimates ranging from 49.4% to 57.8% (2, 3, 4, 5) — markedly higher than figures reported in many Western cohorts. Frailty, an age-related state of diminished physiological reserve, is also common after stroke: a recent analysis of four cohorts spanning 18 countries found that the frailty index rises before stroke, jumps sharply at the event, and continues to accelerate afterwards (6). Notably, comparable acceleration has been documented for cognitive and functional decline even *before* the stroke itself (7), suggesting that stroke is superimposed on a pre-existing trajectory rather than initiating it de novo.

Depression and frailty are not independent. Functional disability is among the strongest predictors of PSD (8, 9), while depressive symptoms in turn predict subsequent disability and impaired activities of daily living (10). This bidirectional coupling has been described repeatedly but almost always as a static, cross-sectional or single-lag relationship. Yet both constructs evolve, and their co-evolution — not their value at any one visit — is the biologically plausible driver of downstream vascular and mortality risk. Recent work has begun to model these constructs longitudinally: Wang et al. characterised frailty-index trajectories before and after stroke (6), and Cai et al. described pre- and post-stroke cognitive trajectories (7). However, these studies examined one phenotype at a time, and none combined depressive symptoms and frailty into a *joint* trajectory in stroke survivors.

A second limitation concerns outcomes. Among studies of psychosocial and functional phenotypes around stroke, the overwhelming majority use self-rated or clinician-rated soft endpoints such as depressive symptom scores, activities of daily living, or cognitive test performance (5, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17). Only a handful address hard endpoints, and those that do generally treat metabolic or lipid markers — for example the estimated glucose

disposal rate (18), remnant cholesterol inflammatory index (19) or HDL-C (20) — as the exposure rather than a peri-stroke trajectory. The one recurrence-prediction model identified in this literature was internally validated only (1). Whether joint depressive-symptom/frailty trajectories spanning the period before and at stroke predict recurrence and death therefore remains an open question.

Third, external validity is a recurring weakness of the existing evidence. Prediction models built on Chinese cohorts routinely acknowledge that their findings have not been validated beyond China (8, 9, 10, 21), and those that claim internal validation typically split the same sampling frame, the same questionnaire and the same respondents — a procedure that cannot supply evidence of independence.

We therefore used the China Health and Retirement Longitudinal Study (CHARLS), a nationally representative cohort of Chinese adults aged 45 years and older, to address three questions. First, can depressive symptoms and frailty be jointly characterised as distinct latent trajectory classes in older stroke survivors at the population level? Second, does membership of a high, increasing joint trajectory mark an elevated risk of stroke recurrence and all-cause mortality? Third, are any observed associations robust to analytic strategy, missing-data handling and alternative outcome definitions? Framing the exposure as the joint course of two population-level health states rather than as a single clinical measurement, we ask whether this joint pattern identifies a subgroup of older survivors who might be prioritised within population-level chronic disease management. We are not aware of a previous study that has modelled joint depressive-symptom/frailty trajectories in relation to hard post-stroke outcomes in a Chinese national cohort; given the breadth of this literature we state this as "we are unaware of" rather than as a first-in-field claim.

Methods

Study design and data source

We performed a retrospective longitudinal cohort study using data from CHARLS, a nationally representative survey of Chinese adults aged 45 years and older that has been fielded in successive waves (2011, 2013, 2015, 2018 and 2020). The CHARLS design and sampling frame have been described in detail elsewhere (22). The survey was approved by the Biomedical Ethics Review Committee of Peking University (IRB00001052-11015), and all participants provided written informed consent. The present analysis used de-identified publicly available data and was conducted in accordance with the Declaration of Helsinki. No separate institutional ethics approval was required for this secondary analysis of de-identified public data.

Study population

The analytic cohort was constructed from the merged CHARLS 2013–2020 records. Participants were eligible if they had complete CESD-10 and frailty-index data in all three of the 2013, 2015 and 2018 waves, had a physician-diagnosed stroke recorded at the 2018 wave (which served as the baseline for outcome ascertainment), and were aged 30–100 years at baseline. Participants without ascertainable follow-up outcome status at the 2020 wave (recurrence assessment, freedom from recurrence, or death) were excluded. The selection process is shown in **Figure 1** and detailed in Additional file 1: Table S1. From the initial 25,380 merged records, 8,176 had complete CESD-10 and FI data in all three waves, 797 had ever had a stroke, 415 had a stroke diagnosis recorded at the 2018 wave (all 415 met the baseline age criterion, so no participant was removed at that step), and 394 had ascertainable follow-up outcomes, constituting the final analytic sample.

Measures

Depressive symptoms. Depressive symptoms were measured with the 10-item Center for Epidemiologic Studies Depression scale (CESD-10), a validated instrument widely used in CHARLS. Scores range from 0 to 30, with higher scores indicating more depressive symptoms.

Frailty. Frailty was operationalised as a deficit-accumulation frailty index (FI) comprising 58 health deficits spanning chronic diseases, functional limitations, disability, sensory impairment, depressive symptoms and self-rated health. Each deficit was coded 0 or 1; the FI was calculated as the proportion of deficits present and therefore ranges from 0 to 1, with higher values indicating greater frailty. The FI was computed at each of the 2013, 2015 and 2018 waves.

Outcomes. The primary outcome was stroke recurrence, ascertained at the 2020 wave. The primary definition classified a participant as having recurrent stroke if the 2020 wave recorded a worsened stroke status. This definition derives from the CHARLS item `da002_8_`, which asks the respondent, in Chinese, whether their doctor-diagnosed disease (here, stroke) is "better, worse or the same compared with the last interview" (wording in English as archived in the CHARLS codebook: *Doctor Diagnosed Disease[8] Compared to [previous wave]*; response options 1 = better, 2 = worse, 3 = same, 99 = never had the disease). We classified a participant as having recurrent stroke when this item took the value 2 ("worse") at the 2020 wave, and we refer to this throughout as *patient-reported global worsening* of stroke status. All-cause death was the secondary outcome and was ascertained from the 2020 wave death indicator. Because eligibility required a physician-diagnosed stroke already recorded at the 2018 baseline, a *new* 2020 diagnosis could not occur within the analytic sample by construction: a composite definition combining worsened status with a new 2020 diagnosis was numerically identical to the primary definition (0 additional events; Additional file 1: Table S9), and no separate composite sensitivity analysis is reported.

Covariates. Covariates were age, sex, current smoking, body mass index (BMI) and educational attainment (categorised as low [illiterate or primary], middle [secondary] and high [tertiary]), all ascertained at the 2018 baseline wave.

Statistical analysis

Joint trajectory modelling. Joint trajectories of CESD-10 and FI were identified with group-based trajectory modelling (GBTM) using a parallel-process specification, in which the two outcomes were stacked into a long-format dataset and modelled with a shared latent class structure and class-specific, independently estimated slopes over time. In plain terms, GBTM asks whether the sample can be sorted into a small number of groups, each following its own average course over the three waves; the parallel-process variant estimates these groups for depressive symptoms and frailty *simultaneously*, so that a participant is assigned to the group that best fits both measures together rather than one at a time. Two readers unfamiliar with the method can think of it as a data-driven form of cluster analysis applied to change over time, but one that returns, for each participant, a probability of belonging to each group rather than a hard label. Models were fitted by expectation–maximisation across solutions specifying one to four classes. The optimal number of classes was selected by the Bayesian information criterion (BIC), subject to the constraints that every class contain at least 5% of the sample and that the mean posterior probability of class membership exceed 0.70. The best-fitting solution was the two-class model (BIC 3,836.17, versus 3,858.06 for one class and 3,877.02 for three), with a mean posterior probability of class membership of 0.957. Based on their estimated trajectories, the two classes were labelled *low-rising* (class 1) and *high-increasing* (class 2); both classes showed rising depressive symptoms and frailty over time, but from markedly different starting levels and with different rates of increase. Following maximum-probability assignment, class 1 comprised

222 participants (56.3%) and class 2 comprised 172 (43.7%) of the 394-person analytic sample; these assigned class sizes are used throughout the Results. As is usual in group-based trajectory modelling, the model-selection statistics computed on the joint likelihood (Additional file 1: Table S2) refer to the fitted classes rather than to the final assigned groups and therefore need not reproduce the assigned percentages exactly. Class-membership posterior probabilities were retained for sensitivity analyses.

Missing data. Covariates were incomplete (current smoking missing for 149 participants, 37.8%; BMI for 83, 21.1%; education for 5, 1.3%; age and sex were complete). Missing covariates were handled by multiple imputation by chained equations (MICE) with 10 imputations, implemented with scikit-learn's `IterativeImputer`. The imputation model was specified as follows. Only the five analysis covariates — age, sex, current smoking, BMI and education group — were entered into the imputation step; the outcome variables (recurrence, death) and the derived trajectory-class variables were deliberately **excluded**, so that the imputation is outcome-blind and cannot manufacture an exposure–outcome association. The imputer used a Bayesian-ridge regression (a Gaussian, that is, continuous-variable, conditional model) with `max_iter = 15`, posterior sampling (`sample_posterior = True`), median initialisation and a fixed random seed per imputation; the dichotomous variables sex and current smoking were rounded back to 0/1 at a threshold of 0.5 after imputation. Convergence was assessed by checking that the iterative-solver objective stabilised well within the 15 iterations across all imputations. Estimates were combined across imputations with Rubin's rules. Complete-case analysis was performed as a comparator. The imputation assumes that covariate data are missing at random (MAR) given the observed covariates; because missingness in smoking and BMI may itself relate to health status, this

assumption is a limitation (see Discussion), and we note that the multiply imputed and complete-case estimates were closely concordant.

Incremental value of the trajectory exposure. To assess whether the joint trajectory adds information beyond simpler exposure definitions, we fitted three additional sets of models under the same specification as the primary analysis (multivariable log-binomial with MICE). First, we substituted a single-timepoint exposure — the 2013 CESD-10 score and the 2013 frailty index, each standardised (per 1 SD) — for the trajectory class, and then entered both the trajectory class and the two single-timepoint terms together. Second, we compared the joint trajectory with two single-phenotype proxies derived from all three waves: an indicator for the upper half of the three-wave mean CESD-10, and the corresponding indicator for the three-wave mean frailty index; these were entered separately and then jointly with the trajectory class. Discriminative performance was summarised as the area under the receiver-operating-characteristic curve (AUC), averaged across imputations, with the covariates included in every model; because only 59 recurrence events occurred, the AUCs are reported descriptively and were not formally compared. Third, to quantify selection bias, we reconstructed the pre-exclusion cohort of 415 participants (participants with a 2018 stroke diagnosis, complete three-wave exposure data and a plausible baseline age) and compared the 394 participants retained with the 21 excluded for an unascertainable 2020 outcome on baseline age, sex, BMI, smoking, CESD-10 and frailty index.

Outcome modelling. Follow-up consisted of a single 2-year interval between the 2018 baseline wave and the 2020 outcome wave, and the exact date of death was not recorded in the available data — only the wave of death. Accordingly, the data provide one binary outcome per participant per endpoint rather than a time-to-event observation, and hazard-based models (Cox proportional hazards, Fine–Gray subdistribution hazards) are not identifiable in this setting. We therefore

estimated risk ratios (RR) for the class 2 versus class 1 contrast with multivariable log-binomial regression, adjusting for age, sex, current smoking, BMI and education. Because log-binomial models can fail to converge when the log-link domain is constrained, robust Poisson regression was used as a corroborating specification, and logistic regression was used to report odds ratios (OR). To account for the competing risk of death when modelling recurrence, a multinomial logit model contrasted recurrence with the composite outcome of event-free survival or death. Absolute risks were summarised as 2-year cumulative incidence by class, with risk differences.

Sensitivity analyses. We repeated the recurrence analysis in the complete-case sample, under robust Poisson, in the subsample with class-membership posterior probability greater than 0.70, and under a logistic specification. We also verified that a recurrence definition restricted to a documented new 2020 diagnosis (that is, excluding the "worsened status" criterion) selected zero events in this cohort, confirming that all recorded recurrences were of the *worsening* type. All analyses were performed in Python 3.13 with statsmodels, scikit-learn and lifelines. Two-sided P values < 0.05 were considered statistically significant. No generative artificial intelligence or AI-assisted technology was used in the preparation of this manuscript.

Results

Cohort characteristics

Of 25,380 merged CHARLS records, 394 stroke survivors met all eligibility criteria and formed the analytic cohort (**Figure 1**). Their mean (SD) age was 65.6 (7.6) years and 201 (51.0%) were men. Slightly fewer than half (177, 44.9%) had completed only primary education or less. The median follow-up was 2 years. Over follow-up, 59 participants (15.0%) had a recurrent stroke and 19 (4.8%) died.

Joint depressive-symptom/frailty trajectories

Group-based trajectory modelling identified two joint trajectory classes (**Figure 2**). Class 1, the *low-rising* class, comprised 222 participants (56.3%). In this class CESD-10 scores were low but nonetheless increased over time (5.46, 5.63 and 7.51 at the 2013, 2015 and 2018 waves), and the frailty index rose from a low baseline (0.061, 0.064 and 0.157). Class 2, the *high-increasing* class, comprised 172 participants (43.7%) and was characterised by markedly higher depressive symptoms at every wave (12.13, 13.50 and 15.08) together with a consistently higher and steeply rising frailty index (0.198, 0.221 and 0.348). Both classes therefore showed upward movement in both phenotypes, but class 1 started from a low level and rose modestly, whereas class 2 started high and deteriorated more steeply — a pattern indicating that depressive symptoms and frailty co-evolve rather than diverge.

Baseline characteristics differed substantially between classes (**Table 1**). Class 1 was 62.6% men and 37.4% women, whereas class 2 was 36.0% men and 64.0% women ($P < 0.001$); in other words, women made up a markedly larger share of the high-increasing class. Participants in class 2 were also more likely to have low educational attainment (56.4% vs 36.0%; $P < 0.001$), and had markedly higher CESD-10 (12.2 vs 5.5; $P < 0.001$) and frailty index (0.200 vs 0.061; $P < 0.001$) values at 2013. Age (66.1 vs 65.2 years; $P = 0.274$), BMI (24.8 vs 25.0 kg/m²; $P = 0.796$) and current smoking (6.4% vs 10.4%; $P = 0.227$) did not differ significantly between classes. Percentages in **Table 1** use the full class-specific denominator; because current smoking was missing for 149 of 394 participants, the smoking percentages are lower than the non-missing-denominator values reported in Additional file 1: Note S2.

Table 1. Baseline characteristics by joint depressive-symptom/frailty trajectory class.

Characteristic	Class 1 low-rising	Class 2 high-	Total (n=394)	P value

	(n=222)	increasing (n=172)		
Age, years, mean (SD)	65.2 (7.7)	66.1 (7.4)	65.6 (7.6)	0.274
Male sex, n (%)	139 (62.6)	62 (36.0)	201 (51.0)	P<0.001
BMI, kg/m2, mean (SD)	25.0 (4.7)	24.8 (3.8)	24.9 (4.3)	0.796
Current smoking, n (%)	23 (10.4)	11 (6.4)	34 (8.6)	0.227
Education: low (illiterate/primary), n (%)	80 (36.0)	97 (56.4)	177 (44.9)	P<0.001
Education: middle (secondary), n (%)	98 (44.1)	66 (38.4)	164 (41.6)	0.294
Education: high (tertiary), n (%)	40 (18.0)	8 (4.7)	48 (12.2)	P<0.001
Female sex, n (%)	83 (37.4)	110 (64.0)	193 (49.0)	P<0.001
CESD-10 at 2013, mean (SD)	5.5 (3.8)	12.2 (6.6)	8.4 (6.2)	P<0.001
Frailty index at 2013, mean (SD)	0.061 (0.054)	0.200 (0.116)	0.121 (0.111)	P<0.001
Stroke recurrence, n (%)	19 (8.6)	40 (23.3)	59 (15.0)	P<0.001
All-cause death, n (%)	7 (3.2)	12 (7.0)	19 (4.8)	0.129

SD, standard deviation; CESD-10, 10-item Center for Epidemiologic Studies Depression scale; BMI, body mass index. Continuous variables compared with Welch t-test, categorical with chi-square test. All categorical percentages (including current smoking) use the full class-specific denominator (class 1, n=222; class 2, n=172; total, n=394). Because current smoking was missing for 149 of 394 participants (37.8%) and BMI for 83 (21.1%), the smoking percentages in this table are lower than the non-missing-denominator percentages reported for the same variable in Additional file 1: Note S2; the full-denominator convention is used throughout the manuscript. Male and female percentages are mutually exclusive and each sums to 100% within a class.

Joint trajectories and 2-year outcomes

Class 2 membership was associated with a substantially higher risk of stroke recurrence (**Table 2**). In the primary multivariable log-binomial analysis with MICE, the RR for recurrence comparing class 2 with class 1 was 3.09 (95% CI 1.83–5.23; P < 0.001). The corresponding

univariable 2-year cumulative incidences were 23.3% in class 2 and 8.6% in class 1, a risk difference of 14.7 percentage points (**Figure 3**).

Table 2. Joint trajectory class and 2-year stroke recurrence and mortality.

Outcome	Model	RR (95% CI)	P value
Stroke recurrence (primary)	Multivariable log-binomial, MICE (m=10)	3.09 (1.83–5.23)	P<0.001
All-cause death	Multivariable log-binomial, MICE (m=10)	2.71 (1.07–6.83)	0.035
All-cause death	Univariable log-binomial, MICE (m=10)	2.21 (0.89–5.50)	0.087
Stroke recurrence (discrete-time competing risk)	Multinomial logit vs event-free survival	5.60 (0.53–59.69)	0.153

RR, risk ratio; CI, confidence interval; MICE, multiple imputation by chained equations. Adjusted for age, sex, current smoking, BMI and education level. Follow-up was a single 2-year interval, so risk ratios (cumulative incidence ratios) are reported rather than hazard ratios.

For all-cause death — an exploratory secondary endpoint — the multivariable log-binomial RR was 2.71 (95% CI 1.07–6.83; P = 0.035), whereas the univariable estimate was 2.21 (95% CI 0.89–5.50; P = 0.087). Only 19 deaths occurred in the whole cohort (12 in class 2 and 7 in class 1), so this estimate is imprecise: the confidence interval is wide and approaches the null in the unadjusted analysis. Absolute 2-year mortality was 7.0% in class 2 versus 3.2% in class 1 (risk difference 3.8 percentage points). Because of the small number of events, the mortality findings are reported for completeness and should be treated as hypothesis-generating rather than confirmatory.

When recurrence was modelled as a discrete-time competing-risk outcome with a multinomial logit specification, the point estimate for class 2 was 5.60 (95% CI 0.53–59.69; P = 0.153). This estimate is accompanied by an extremely wide confidence interval and should be interpreted with

caution: only five deaths occurred among participants with complete covariate data, so the death category was weakly identified.

Incremental value relative to simpler exposure definitions

Because a single-timepoint or single-phenotype exposure would be far easier to apply in practice than a three-wave joint trajectory, we asked whether the trajectory classification carries information beyond these simpler alternatives (**Table 3**). Used alone, the 2013 single-timepoint CESD-10 score was only weakly associated with recurrence (RR 1.23 per 1 SD, 95% CI 0.99–1.53; P = 0.067) and the 2013 frailty index moderately so (RR 1.37 per 1 SD, 95% CI 1.14–1.66; P = 0.001). When the two single-timepoint terms were added to a model that already contained the trajectory class, the class 2 versus class 1 contrast remained (RR 2.19, 95% CI 1.13–4.25; P = 0.020) while both single-timepoint terms became non-significant (CESD-10 RR 1.14, 95% CI 0.90–1.44; frailty index RR 1.16, 95% CI 0.93–1.44). The symmetry of this result implies that the trajectory exposure captures most of the prognostic information, but it does not fully separate the *slope* of the trajectory from the 2018 level, and the attenuation of the class contrast from 3.09 to 2.19 confirms that the two exposures overlap substantially.

Table 3. Incremental value of the joint trajectory versus simpler exposures.

Analysis	N	RR (95% CI)	P value
Recurrence: joint trajectory only (class 2 vs 1; reference for the single-timepoint block below)	394	3.09 (1.83–5.23)	P<0.001
Recurrence: 2013 single-timepoint CESD-10 only (per 1 SD)	394	1.23 (0.99–1.53)	0.067
Recurrence: 2013 single-timepoint frailty index only (per 1 SD)	394	1.37 (1.14–1.66)	0.001
Recurrence: joint trajectory, further adjusted for 2013 single timepoint (class 2 vs 1)	394	2.19 (1.13–4.25)	0.020
Recurrence: 2013 CESD-10, further adjusted for joint	394	1.14 (0.90–1.44)	0.276

trajectory (per 1 SD)			
Recurrence: 2013 frailty index, further adjusted for joint trajectory (per 1 SD)	394	1.16 (0.93–1.44)	0.198
Recurrence: high CESD-10 level only (3-wave mean, top half)	394	2.33 (1.37–3.95)	0.002
Recurrence: high frailty-index level only (3-wave mean, top half)	394	2.52 (1.48–4.31)	P<0.001
Recurrence: joint trajectory only (class 2 vs 1; reference for the single-phenotype block below)	394	3.09 (1.83–5.23)	P<0.001
Recurrence: joint trajectory, adjusted for both single-phenotype indicators (class 2 vs 1)	394	2.22 (0.95–5.19)	0.064
Discrimination (AUC): joint trajectory + covariates	394	AUC 0.696	—
Discrimination (AUC): 2013 single timepoint + covariates	394	AUC 0.677	—
Discrimination (AUC): joint trajectory + 2013 single timepoint + covariates	394	AUC 0.701	—

RR, risk ratio; CI, confidence interval; AUC, area under the receiver-operating-characteristic curve; SD, standard deviation. All models include age, sex, current smoking, BMI and education, and are fitted with multiple imputation (m = 10) and pooled by Rubin's rules. Continuous single-timepoint exposures are standardised (per 1 SD). Single-phenotype indicators contrast the top half of the three-wave mean of that phenotype with the bottom half. AUC values are averaged across imputations and are descriptive only; no formal test of difference between AUCs is reported because of the small number of events (59).

The single-phenotype proxies pointed the same way. Defining exposure only by a high three-wave mean CESD-10 gave RR 2.33 (95% CI 1.37–3.95; P = 0.002) and by a high three-wave mean frailty index gave RR 2.52 (95% CI 1.48–4.31; P < 0.001), both lower than the joint trajectory estimate of 3.09 (95% CI 1.83–5.23). When the joint trajectory was adjusted for both single-phenotype indicators simultaneously, its estimate was RR 2.22 (95% CI 0.95–5.19; P = 0.064) — attenuated and no longer statistically significant — so the joint classification and the

two phenotype-level indicators explain overlapping, though not identical, variance. Discriminative performance was similar across the three exposure definitions (AUC 0.696 for the joint trajectory, 0.677 for the 2013 single timepoint and 0.701 for the two combined); with only 59 events these differences are descriptive and not formally testable. Taken together, the joint trajectory does not demonstrably out-perform simpler exposures on discrimination, but it identifies a risk contrast at least as large as either single phenotype and remains the only definition in which both phenotypes are represented simultaneously.

Selection and comparison with excluded participants

Of the 415 participants who met the cohort definition up to and including the outcome-ascertainment step (2018 stroke diagnosis, complete three-wave exposure data, plausible baseline age), 394 had an ascertainable 2020 outcome and 21 did not. The excluded participants were somewhat older (mean 68.8 vs 65.6 years; P = 0.064) and had somewhat higher 2018 depressive-symptom scores (CESD-10 13.2 vs 10.9; P = 0.067) and frailty-index values (0.28 vs 0.24; P = 0.184). Sex, BMI and current smoking were closely similar between the two groups (all P > 0.39; Additional file 1: Table S10). The pattern suggests that exclusion was not random with respect to health: participants lost to outcome ascertainment tended to be older and more symptomatic, which is consistent with attrition through death or severe illness. Because these characteristics are themselves associated with recurrence, the retained cohort is plausibly *healthier* than the full eligible population, and the trajectory–recurrence association observed here may therefore underestimate the contrast that would be seen if all eligible participants could be followed. This directional argument is heuristic and cannot be verified in the present data.

Sensitivity analyses

The association between class 2 membership and stroke recurrence was robust across all prespecified sensitivity analyses (**Table 4**, **Figure 4**). In the complete-case sample (n = 188, 31 recurrence events), the multivariable log-binomial RR was 2.66 (95% CI 1.22–5.80; P = 0.014); robust Poisson gave 2.70 (95% CI 1.26–5.79; P = 0.011); and logistic regression gave an OR of 3.29 (95% CI 1.32–8.19; P = 0.010). Restricting the analysis to participants whose class-membership posterior probability exceeded 0.70 (n = 182, 30 events) yielded an RR of 2.59 (95% CI 1.18–5.67; P = 0.017). As prespecified, the estimates were also unchanged when the multiply imputed and complete-case analyses were compared, indicating that the association is not an artefact of the missing-data mechanism. The model-selection statistics underlying the two-class solution are shown in **Figure 5**.

Table 4. Sensitivity analyses for trajectory class and 2-year outcomes.

Analysis	N	Events	RR / OR (95% CI)	P value
Main: complete case, log-binomial	188	31	2.66 (1.22–5.80)	0.014
Main: multiple imputation (m=10), log-binomial	394	59	3.09 (1.83–5.23)	P<0.001
Complete case, robust Poisson	188	31	2.70 (1.26–5.79)	0.011
Subsample with posterior probability >0.70	182	30	2.59 (1.18–5.67)	0.017
Complete case, logistic regression (OR)	188	31	3.29 (1.32–8.19)	0.010
All-cause death, univariable (crude), log-binomial, full cohort (exploratory)	394	19	2.21 (0.89–5.50)	0.087

RR, risk ratio; OR, odds ratio; CI, confidence interval. All recurrence rows estimate the class 2 vs class 1 contrast. The prespecified composite recurrence definition (worsened stroke status or a new 2020 diagnosis) was numerically identical to the primary definition in every specification, because eligibility required a physician-diagnosed stroke already recorded at the 2018 baseline;

it is therefore not listed as a separate row. The mortality row is unadjusted (crude) and based on very few events; it is reported for completeness and must not be interpreted as confirmatory.

Discussion

In this nationally representative cohort of older Chinese stroke survivors, the two most commonly co-occurring health states in this population — depressive symptoms and frailty — resolved into two distinct joint trajectory classes. A little over two in five survivors followed a high-increasing trajectory in which both phenotypes were elevated at baseline and continued to worsen, and this trajectory was associated with approximately threefold risk of stroke recurrence within two years. The recurrence association was stable across complete-case and multiply imputed analyses, across log-binomial, robust Poisson and logistic specifications, and after restricting to participants with confident class membership. The mortality association was directionally consistent after adjustment, but it rested on only 19 deaths and should be regarded as exploratory. Two further qualifications temper the interpretation and are developed below and in the limitations: the trajectory contrast weakened from 3.09 to 2.19 once a single baseline timepoint was accounted for, and the joint trajectory did not out-perform simpler exposure definitions on discrimination — so the findings are best regarded as hypothesis-generating rather than as establishing a predictive instrument.

These findings extend the literature in three ways. First, they move beyond single-phenotype, single-timepoint exposure definitions. Previous work has established that frailty accelerates around stroke (6) and that cognitive and functional decline precede the event (7), and separately that depression and disability are bidirectionally coupled (8, 9, 10). Our results suggest that these two processes may be better understood as a single co-evolving trajectory: in class 2 the CESD-10 and frailty-index curves rose roughly in parallel, and it was this joint pattern that marked the

highest-risk group. That said, our comparison against simpler exposures (Table 3) requires a cautious reading: the joint trajectory did not clearly out-perform either a single 2013 timepoint or a three-wave mean of either phenotype on discrimination (AUC 0.696 vs 0.677 vs 0.701), and its estimate was attenuated when the single-timepoint terms were included. We therefore do not claim that the *joint trajectory shape itself* is a demonstrated independent driver; the data are equally compatible with the simpler reading that survivors with both high depressive symptoms and high frailty at any time are at elevated risk. We advance the co-evolution interpretation as a hypothesis — one consistent with the notion of a shared disabling pathway in which mood and physiological reserve reinforce one another — rather than as an established mechanism. Formal mediation or joint-model analyses, which were not possible here, would be required to test it.

Second, the study addresses a well-documented gap in outcome selection. The psychosocial literature around stroke is dominated by soft endpoints such as symptom scores, ADL function and cognitive test performance (5, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17). Studies reporting hard endpoints have generally examined metabolic exposures instead — estimated glucose disposal rate (18), remnant cholesterol inflammatory index (19) and HDL-C (20) — rather than peri-stroke trajectories, and the single recurrence-prediction model we identified was internally validated only (1). By relating joint trajectories to recurrence and death, our analysis supplies evidence that two widely prevalent population-level health states carry prognostic information for outcomes that matter to patients, clinicians and health services alike.

Third, the robustness profile is worth emphasising. The recurrence estimate varied between 2.19 and 3.29 across specifications, and the confidence interval excluded 1.0 in every primary and sensitivity analysis (though not after simultaneous adjustment for both single-phenotype indicators). The concordance of the multiply imputed and complete-case estimates argues that the

association is not an artefact of the missing-data mechanism, and its stability across the restricted (high-posterior) subsample argues that it is not driven by uncertain class assignment. At the same time, we must apply to our own work the same criticism we level at others: this study provides no external validation. All estimates derive from a single sampling frame (CHARLS), a single questionnaire battery and a single national population, and we did not split the data into derivation and validation subsets, nor did we test the trajectory classification in an independent cohort. The internal consistency across specifications demonstrates that the result is not fragile to analytic choices, but it cannot establish that the pattern generalises beyond China or beyond CHARLS respondents.

The absolute risks of recurrence are clinically informative. A 2-year recurrence risk of 23.3% in the high-increasing class, against 8.6% in the low-rising class, represents a risk difference of nearly 15 percentage points. If these findings are replicated, joint trajectory status could be examined as a candidate criterion for prioritising older survivors for intensified follow-up, and for integrated programmes that address mood and physical reserve together rather than in separate service silos. Because depressive symptoms and frailty are both at least partly modifiable, unlike age or stroke subtype, this is a potentially actionable stratification at the level of chronic disease programmes; we stress, however, that the present design supports only the generation of that hypothesis. Our incremental-value analysis (Table 3) shows that the trajectory contrast shrinks from 3.09 to 2.19 once the 2018 single-timepoint terms are adjusted for, so the analysis does not establish that the *trajectory shape* provides benefit over a simpler single-visit assessment, and any claim of practical superiority would require a prospective comparison. The corresponding mortality contrast (7.0% vs 3.2%) points in the same direction but is based on too few events to support a firm conclusion.

Several limitations should be acknowledged. First, the temporal framing requires care. The trajectory exposure was constructed from CESD-10 and frailty-index measurements collected at the 2013, 2015 and 2018 waves, whereas the baseline stroke diagnosis was recorded at the 2018 wave; the exposure therefore captures psychosocial and functional change *before and at* stroke rather than strictly after it, and we refer to it throughout as a peri-stroke trajectory. Participants had to survive to the 2018 wave with a recorded stroke diagnosis and complete three-wave exposure data, so the cohort is left-truncated at 2018 and subject to survivorship bias: individuals who died or were lost between 2013 and 2018, including those with the most severe post-stroke phenotypes, are not represented. This would tend to attenuate contrast between trajectory classes and to bias estimates towards the null, but the direction cannot be assumed with certainty. A related mechanism operates at the final step, and we were able to quantify it: of the 415 participants eligible at the outcome-ascertainment step, the 21 excluded for an unrecorded 2020 outcome were older and more depressed than the 394 retained (see Results), so the analytic sample is likely to be somewhat healthier than the full eligible population. Second, CHARLS does not record stroke subtype (ischaemic or haemorrhagic), lesion location, stroke severity (for example the NIH Stroke Scale), the exact date of stroke onset, or secondary prevention medication such as antiplatelet agents or statins; these omissions have been noted by others using the same data (4) and mean that residual confounding by stroke severity and treatment cannot be excluded. Third, the outcome definition rests on self-reported worsening rather than adjudicated recurrence. The `da002_8_` item asks whether the doctor-diagnosed disease is better, worse or the same as at the previous interview; it therefore conflates a genuine recurrent event with a change in the respondent's global appraisal, and it may capture both false positives (non-vascular worsening attributed to stroke) and false negatives (a silent or mildly symptomatic recurrence reported as "the same"). Misclassification of this kind is most plausibly non-differential with

respect to the trajectory class, which would bias the association towards the null, but differential recall — for instance, more depressed survivors reporting worsening more readily — would bias it away from the null, so the direction cannot be assumed; a sensitivity analysis restricted to a documented new 2020 diagnosis returned zero events (Additional file 1: Table S9), so the primary estimate depends entirely on the worsening criterion and its validity remains the single most important assumption of the analysis. Fourth, the imputation of missing covariates assumed that they were missing at random given the observed covariates. Smoking and BMI were missing for a substantial minority, and if their missingness depended on unmeasured health status — that is, if the data were missing not at random (MNAR) — the imputed values would be biased; the close agreement between the multiply imputed and complete-case estimates is reassuring but does not exclude this possibility. Fifth, follow-up comprised a single 2-year interval and the exact date of death was unavailable, so we could not estimate time-to-event hazards; the reported measures are therefore risk ratios over a fixed interval, not hazard ratios, and Cox and Fine–Gray models are not identifiable in this dataset. Consequently the multinomial competing-risk estimate is imprecise and rests on only a few deaths in the complete-case stratum, and the mortality analysis itself, based on 19 deaths, should be regarded as exploratory rather than confirmatory. Finally, the joint trajectory did not demonstrably out-perform simpler exposure definitions on discrimination (Table 3), so the added analytic complexity of a three-wave joint trajectory model may not be justified for predictive purposes; its value, if any, lies in description and hypothesis generation rather than in risk stratification. The cohort is also restricted to Chinese adults aged 45 years and older and to participants with complete three-wave data, and the findings may not generalise to other populations or to survivors with different data-availability patterns; external validation in an independent cohort — ideally one drawn from a different sampling frame — remains necessary.

Conclusions

Among older Chinese stroke survivors, a joint trajectory characterised by high and increasing depressive symptoms and frailty identified a group at approximately threefold risk of stroke recurrence over two years. This association was robust across analytic strategies, but it should be treated as hypothesis-generating: the analysis could not separate the trajectory *shape* from the 2018 level, and it did not demonstrate that the joint trajectory out-performs simpler single-phenotype exposures on discrimination. Mortality was higher in this group in the same direction, but the small number of deaths precludes a firm conclusion. Two widely prevalent and at least partly modifiable health states therefore appear to evolve together in this population, which suggests that mood and physiological reserve should be monitored jointly rather than separately in older survivors. Whether joint trajectory assessment can improve population-level chronic disease management requires prospective testing and external validation before it can inform practice.

List of abbreviations

ADL: activities of daily living; AUC: area under the receiver-operating-characteristic curve; BIC: Bayesian information criterion; BMI: body mass index; CESD-10: 10-item Center for Epidemiologic Studies Depression scale; CHARLS: China Health and Retirement Longitudinal Study; CI: confidence interval; FI: frailty index; GBTM: group-based trajectory modelling; MAR: missing at random; MICE: multiple imputation by chained equations; MNAR: missing not at random; OR: odds ratio; PSD: post-stroke depression; RR: risk ratio; SD: standard deviation.

Declarations

Ethics approval and consent to participate: The CHARLS survey was approved by the Biomedical Ethics Review Committee of Peking University (IRB00001052-11015), and all participants provided written informed consent. This study is a secondary analysis of de-identified, publicly available CHARLS data; it involved no new data collection and no contact with participants. As the data are de-identified and publicly released, the analysis was determined to be exempt from further institutional review board approval at the authors' institution; no separate local approval number was therefore issued.

Consent for publication: Not applicable.

Availability of data and materials: The datasets analysed during the current study are publicly available from the China Health and Retirement Longitudinal Study (CHARLS) website, http://charls.pku.edu.cn/ (accessed 15 September 2026); access requires registration and acceptance of the CHARLS data user agreement. The design of the CHARLS cohort is described in the accompanying cohort profile (22). The analysis code (data cleaning, trajectory modelling, imputation, regression and table and figure generation) is available from the corresponding authors on reasonable request.

Competing interests: The authors declare that they have no competing interests.

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Authors' contributions: PY, LL and CZ conceived and designed the study and contributed to data acquisition. JC performed the statistical analysis and the trajectory modelling. JC and YT interpreted the data and drafted the manuscript. PY and LL supervised the study and critically revised the manuscript for important intellectual content. All authors read and approved the final

manuscript and agree to be accountable for all aspects of the work. PY and LL contributed equally to this work and share first authorship. YT and CZ contributed equally to this work and share corresponding authorship.

Acknowledgements: The authors thank the China Health and Retirement Longitudinal Study (CHARLS) research team and all the respondents for making the data publicly available. The authors also thank the staff of the Rehabilitation Medicine Center, Fujian Provincial Governmental Hospital, for their support of this work.

Authors' information: Not applicable.

References

1. Jin Z, Gao W, Yu T, Guo F, Shi Q, Yu S, et al. Development and Internal Validation of a Model Predicting the Risk of Recurrent Stroke for Middle-Aged and Elderly Patients: A Retrospective Cohort Study. World Neurosurg. 2022;168:e418-e431. doi:10.1016/j.wneu.2022.10.049

2. Cai Q, Qian M, Chen M. Association between socioeconomic status and post-stroke depression in middle-aged and older adults: results from the China health and retirement longitudinal study. BMC Public Health. 2024;24(1):1007. doi:10.1186/s12889-024-18503-z

3. Xiao W, Liu Y, Huang J, Huang LA, Bian Y, Zou G. Analysis of factors associated with depressive symptoms in stroke patients based on a national cross-sectional study. Sci Rep. 2024;14(1):9268. doi:10.1038/s41598-024-59837-3

4. Shi Y, Liu C, Sun X, Li D, Wang S, Zhu X, et al. Analysis of the current status and associated factors of depressive symptoms in Chinese middle-aged and elderly stroke patients-based on CHARLS data. Front Psychiatry. 2025;16:1602287. doi:10.3389/fpsyt.2025.1602287

5. Zhuang Z, Ji L, Chen Y, Chen E, Mo F, Zhou J, et al. Developing a predictive model for classifying high-risk groups of post-stroke depressive symptoms based on the health ecology model and Harvard Cancer Index. Psychol Health Med. 2026;31(7):1823-1841. doi:10.1080/13548506.2025.2601066

6. Wang C, Qian X, Zhao X, Jin M, Jia C, Jia F. Frailty index changes before and after stroke: Findings from four cohorts covering 18 countries. Int J Stroke. 2026;21(6):871-878. doi:10.1177/17474930251396655

7. Cai Y, Duan C, Wen S, Ling Z, Huang G, Zou J, et al. Progression of cognitive and functional decline before and after stroke: evidence from four prospective cohort studies across 23 countries. J Neurol. 2025;272(12):779. doi:10.1007/s00415-025-13511-4

8. Lin F, Zhou M. Establishment and validation of a model for predicting depression risk in stroke patients. BMC Psychiatry. 2025;25(1):668. doi:10.1186/s12888-025-06885-2

9. Lin F, Zhou M. Development and Validation of a Risk Prediction Model for Depression in Patients with Stroke. Arch Clin Neuropsychol. 2025;40(6):1082-1090. doi:10.1093/arclin/acaf021

10. Ren H, Wang W, Tao L, Song X, Bai B, Wang M, et al. Development and validation of a prediction model for activities of daily living dysfunction among stroke survivors: insights from the CHARLS cohort. J Neuroeng Rehabil. 2026;23(1). doi:10.1186/s12984-026-02063-x

11. Zhang Q, Gao X, Huang J, Xie Q, Zhang Y. Association of pre-stroke frailty and health-related factors with post-stroke functional independence among community-dwelling Chinese older adults. J Stroke Cerebrovasc Dis. 2023;32(6):107130. doi:10.1016/j.jstrokecerebrovasdis.2023.107130

12. Lin B, Chen N, Chen X, Jia J, Xiang J, Zhang T, et al. Predicting post-stroke depression in middle-aged and older adults in china using obesity-related index: A cohort study. J Stroke Cerebrovasc Dis. 2025;34(12):108468. doi:10.1016/j.jstrokecerebrovasdis.2025.108468

13. Zhao Y, Zeng D, Yu H, Zheng Y, Peng T, Duan X, et al. Blood Biomarker-Based Machine Learning Model for Predicting Cognitive Impairment in Stroke Patients. World Neurosurg. 2025;201:124276. doi:10.1016/j.wneu.2025.124276

14. Zuo S, Liu N, Wang J, Li J, Zhu X, Jia Y. Development and validation of a prediction model for long-term cognitive frailty risk in stroke patients based on CHARLS data. PLoS One. 2026;21(3):e0340715. doi:10.1371/journal.pone.0340715

15. Huang X, Tang Z, Xiong T. Associations of pre-stroke function disability and post-stroke cognitive impairment among older adults in China. Sci Rep. 2026;16(1). doi:10.1038/s41598-026-39018-0

16. Fang Z, Chen J, Deng J. Association Between Social Activities and Depression in Middle-Aged and Older Stroke Patients. Int J Behav Med. 2026. doi:10.1007/s12529-026-10438-6

17. Zhang H, Liu W, Hao X, Zhao X, Li K. Patterns of Social Participation Among Survivors of Stroke in China and the Relationship With Personal and Environmental Factors: A Latent Class Analysis. West J Nurs Res. 2026;48(3):255-263. doi:10.1177/01939459251405055

18. Li D, Guo Y, Cui F, Cheng Y, Yan S, Yu T, et al. The association between insulin resistance assessed by estimated glucose disposal rate and stroke prevalence and mortality in non-diabetic people: evidence from two prospective cohorts. Diabetol Metab Syndr. 2025;17(1):390. doi:10.1186/s13098-025-01956-6

19. Chen H, Wu JY, Yan H, Gao J, Li C, Xie JH, et al. Association of Remnant Cholesterol Inflammatory Index with Stroke, Heart Disease and All-Cause Mortality Across Cardiovascular-

Kidney-Metabolic Syndrome Stages 0-3: A National Cohort Study. Nutrients. 2026;18(2). doi:10.3390/nu18020205

20. Liu Z, Liang L, Du X, He H. Association of HDL-C Levels with Stroke and All-cause Mortality: A Prospective Cohort Study from the CHARLS. Brain Behav. 2026;16(3):e71179. doi:10.1002/brb3.71179

21. Lin F, Liu N. Development and validation of a risk prediction model for activities of daily living dysfunction in stroke survivors. Front Neurol. 2025;16:1529724. doi:10.3389/fneur.2025.1529724

22. 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. doi:10.1093/ije/dyt203

Figure legends

Figure 1. Flow chart of participant selection from the CHARLS 2013–2020 waves.

Participant selection from the merged CHARLS 2013–2020 records to the final analytic cohort of 394 stroke survivors. Numbers at each step are shown; the corresponding tabulation is given in Additional file 1: Table S1.

Figure 2. Estimated joint trajectories of depressive symptoms and frailty, by latent class.

Estimated joint trajectories of depressive symptoms (CESD-10) and frailty (frailty index) from 2013 to 2018, by latent class. Class 1 (low-rising, n = 222) and class 2 (high-increasing, n = 172); both classes show rising trajectories, differing in starting level and rate of increase. Points are class-specific estimated means at each wave; error bars indicate 95% confidence intervals. CESD-10, 10-item Center for Epidemiologic Studies Depression scale.

Figure 3. Two-year cumulative incidence of stroke recurrence and death, by trajectory class.

Cumulative incidence of stroke recurrence and all-cause death over the 2-year follow-up, by joint trajectory class. Bars show the proportion of participants experiencing each outcome within each class, with 95% confidence intervals.

Figure 4. Risk estimates for the high-increasing trajectory across analytic specifications.

Forest plot of risk estimates for the high-increasing joint trajectory (class 2 vs class 1). The upper five rows show stroke recurrence in the primary analysis (multiple imputation, adjusted) and in the four sensitivity analyses (complete case, robust Poisson, class-membership posterior probability > 0.70, and logistic specification); the bottom row shows the adjusted estimate for all-cause death. The logistic analysis reports an odds ratio.

Figure 5. Model-selection statistics for the group-based joint trajectory models.

Model-selection statistics for the group-based trajectory models. Panel A shows the Bayesian information criterion across solutions specifying one to four classes; panel B shows the size of the smallest class and the mean posterior probability of class membership. The two-class solution was selected.

Additional files

Additional file 1. Supplementary tables and notes.

File format: .docx (Microsoft Word).

Title of data: Supplementary tables and notes.

Description of data: Supplementary Tables S1–S12 and Supplementary Notes S1–S2, comprising the participant-selection steps, model-selection statistics, class-specific estimated means, absolute 2-year risks, full multivariable model coefficients under multiple imputation, complete-case

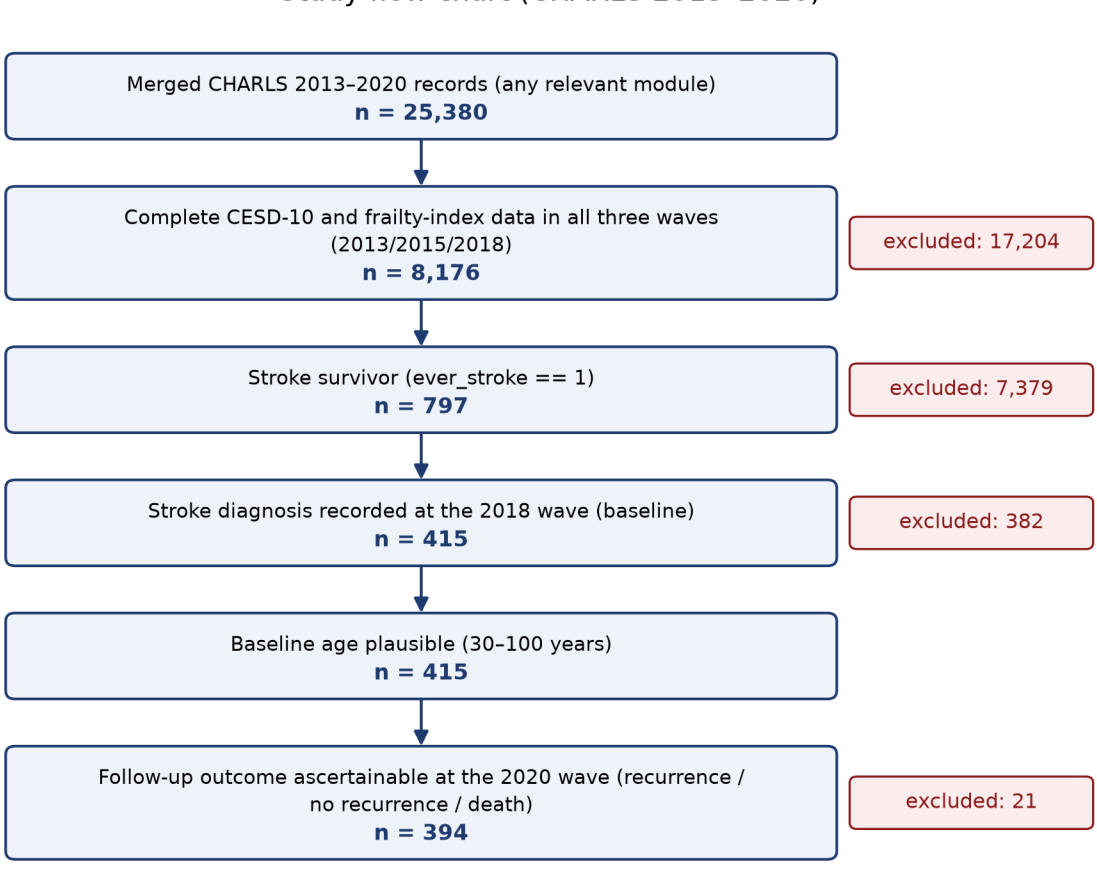
models, the discrete-time competing-risk analysis, software versions, the number of events captured by each recurrence definition, the comparison of retained and excluded participants, the incremental-value models, and the multiple-imputation specification.

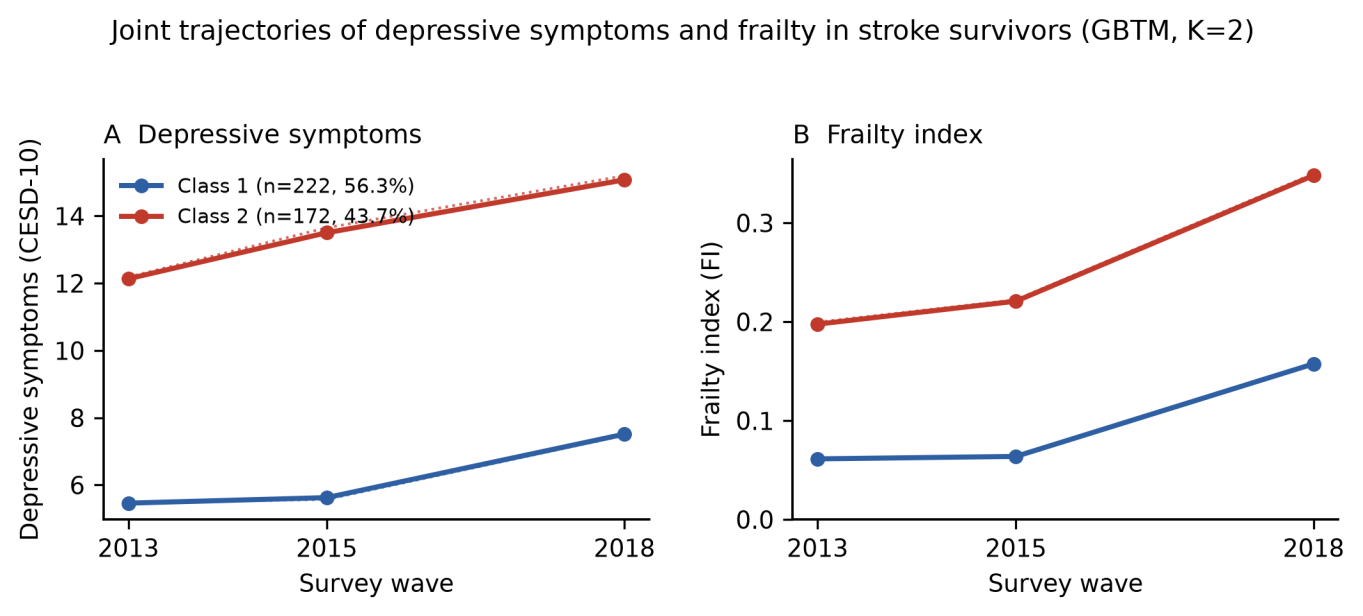
Additional file 2. STROBE checklist.

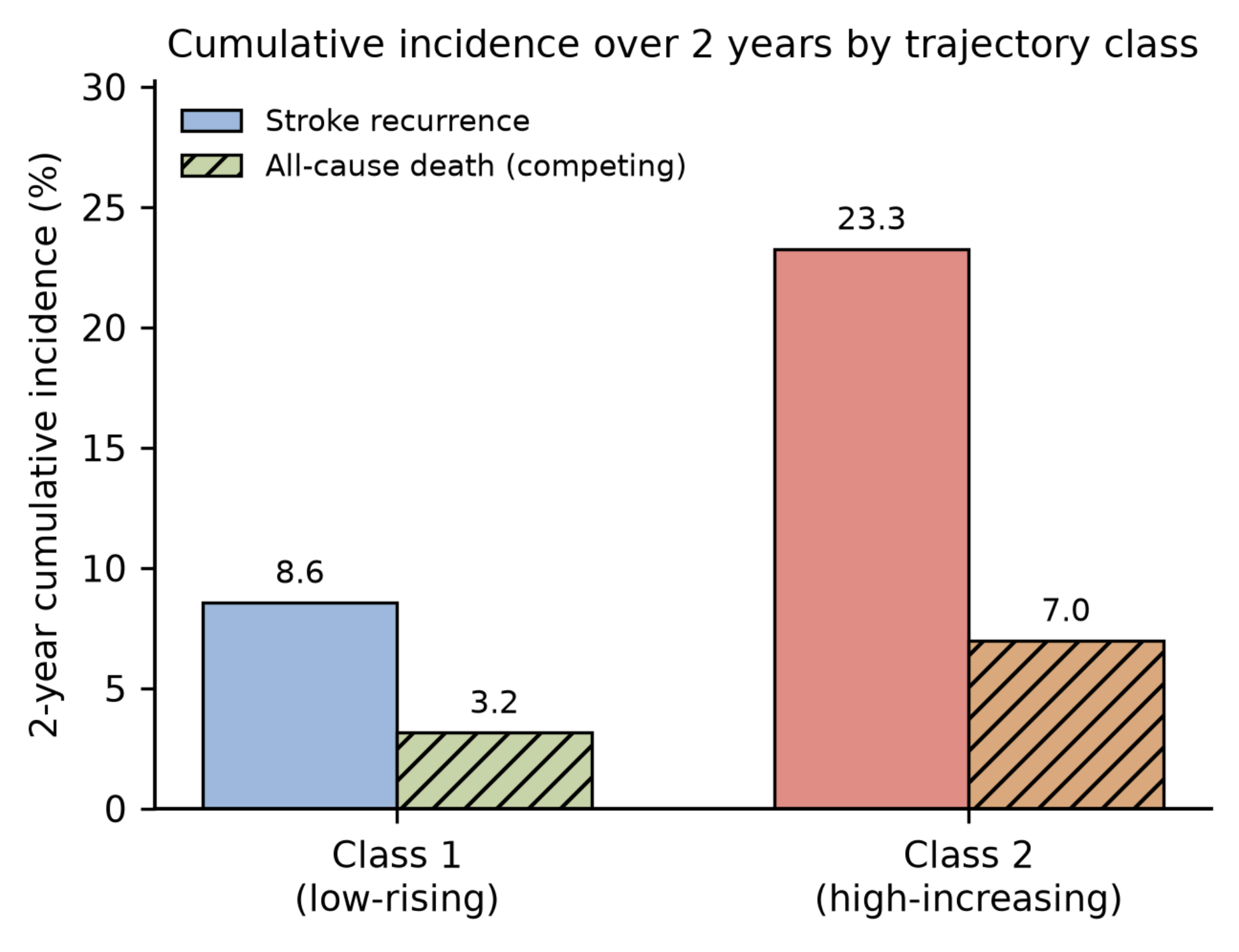
File format: .docx (Microsoft Word).

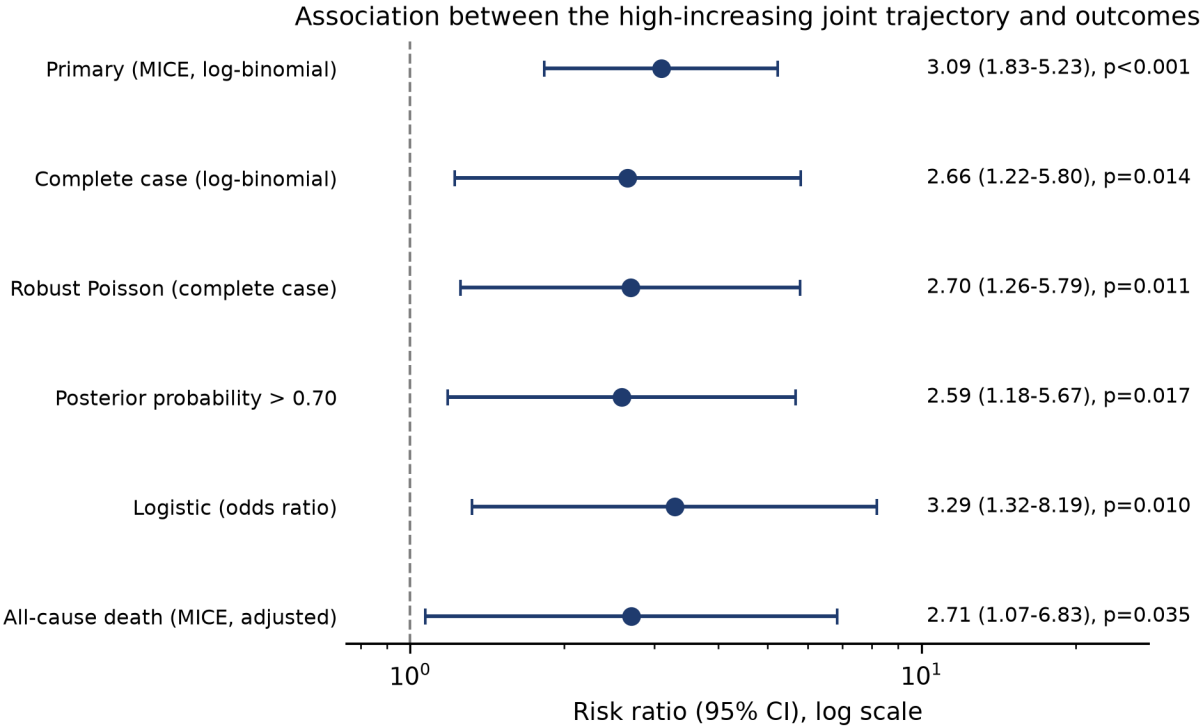
Title of data: STROBE Statement — checklist of items for cohort studies.

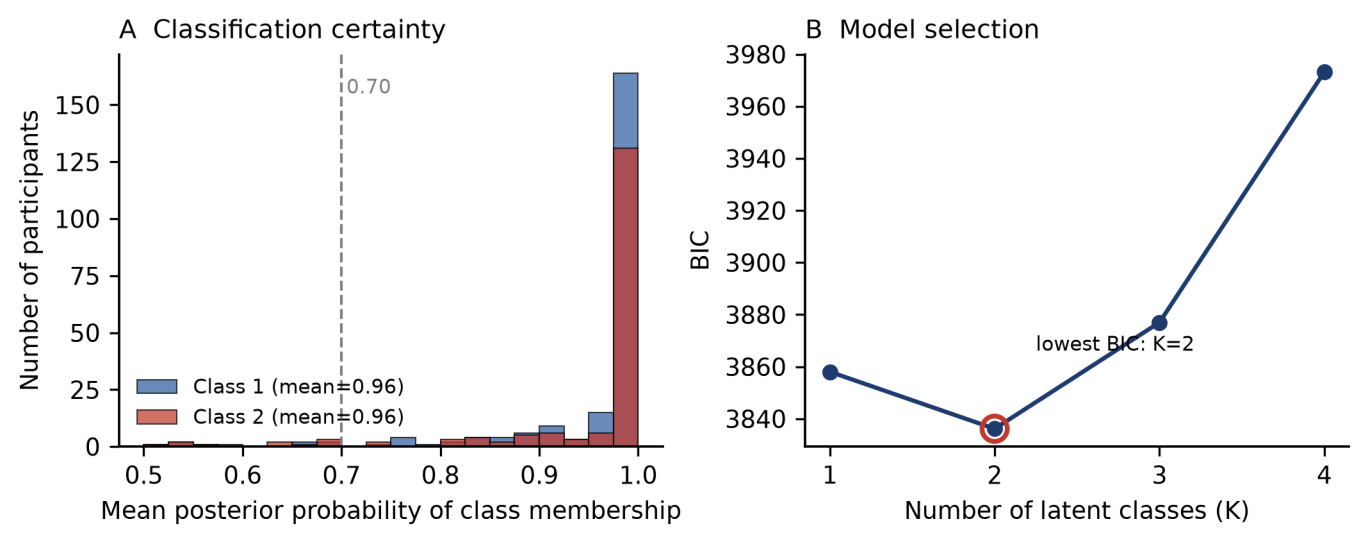
Description of data: Completed STROBE checklist for cohort studies (items 1–22) indicating the manuscript location of each item.


Study flow chart (CHARLS 2013–2020)
Merged CHARLS 2013–2020 records (any relevant module)
n = 25,380
Complete CESD-10 and frailty-index data in all three waves (2013/2015/2018)
n = 8,176
excluded: 17,204
Stroke survivor (ever_stroke == 1)
n = 797
excluded: 7,379
Stroke diagnosis recorded at the 2018 wave (baseline)
n = 415
excluded: 382
Baseline age plausible (30–100 years)
n = 415
Follow-up outcome ascertainable at the 2020 wave (recurrence / no recurrence / death)
n = 394
excluded: 21


Joint trajectories of depressive symptoms and frailty in stroke survivors (GBTM, K=2)
A Depressive symptoms
Class 1 (n=222, 56.3%)
Class 2 (n=172, 43.7%)
Depressive symptoms (CESD-10)
14
12
10
8
6
2013
2015
2018
Survey wave
B Frailty index
Frailty index (FI)
0.3
0.2
0.1
0.0
2013
2015
2018
Survey wave


Cumulative incidence over 2 years by trajectory class
Stroke recurrence
All-cause death (competing)
2-year cumulative incidence (%)
30
25
20
15
10
5
0
8.6
3.2
23.3
7.0
Class 1
(low-rising)
Class 2
(high-increasing)


Association between the high-increasing joint trajectory and outcomes
Primary (MICE, log-binomial)
3.09 (1.83-5.23), p<0.001
Complete case (log-binomial)
2.66 (1.22-5.80), p=0.014
Robust Poisson (complete case)
2.70 (1.26-5.79), p=0.011
Posterior probability > 0.70
2.59 (1.18-5.67), p=0.017
Logistic (odds ratio)
3.29 (1.32-8.19), p=0.010
All-cause death (MICE, adjusted)
2.71 (1.07-6.83), p=0.035
10^0
10^1
Risk ratio (95% CI), log scale


A Classification certainty
0.70
Number of participants
150
125
100
75
50
25
0
Class 1 (mean=0.96)
Class 2 (mean=0.96)
0.5
0.6
0.7
0.8
0.9
1.0
Mean posterior probability of class membership
B Model selection
BIC
3980
3960
3940
3920
3900
3880
3860
3840
lowest BIC: K=2
1
2
3
4
Number of latent classes (K)

Supplementary Files

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

- Additionalfile1.docx
- Additionalfile2STROBE.docx