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.

Cardiometabolic abnormality burden and the emergence and recovery of health deficits in middle-aged and older adults: a harmonized longitudinal study across three national cohorts

Ruixia Liu
The Second Affiliated Hospital of Anhui Medical University
Xiaotian Wang
The First Affiliated Hospital of Anhui Medical University
Tianrong Pan
The Second Affiliated Hospital of Anhui Medical University
Ping Zhang
efy117405@fy.ahmu.edu.cn
The Second Affiliated Hospital of Anhui Medical University

Research Article

Keywords: cardiometabolic health, chronic disease epidemiology, deficit accumulation, health transitions, population aging, longitudinal cohorts

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11293689/v1

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

Additional Declarations: No competing interests reported.

Abstract

Background

Cardiometabolic abnormalities are common, potentially modifiable public health concerns, but studies of frailty usually examine net deficit burden or incident frailty and cannot show whether associations reflect more new deficits, less recovery, or both. We examined these two transition processes across three national aging cohorts.

Methods

This harmonized longitudinal secondary analysis included 25,590 adults aged 50 years or older from the China Health and Retirement Longitudinal Study, the Health and Retirement Study, and the English Longitudinal Study of Ageing. Baseline burden counted central obesity, high blood pressure, and diagnosed or treated diabetes (0–3). Weighted discrete-time complementary log-log models estimated item-level emergence and recovery across 18 potentially reversible deficits. Cohort-specific estimates were combined by random-effects meta-analysis, and marginal standardization yielded model-based annualized one-year transition probabilities and expected net change.

Results

Emergence analyses included 23,689 participants and 130,915 person-item-interval transition events; recovery analyses included 22,216 participants and 99,789 events. Per additional abnormality, pooled Model A hazard ratios were 1.125 (95% confidence interval 1.073–1.179; I^2=90.9%) for emergence and 0.933 (0.921–0.946; I^2=0%) for recovery. Estimates remained directionally consistent after additional adjustment for interval-start deficit proportion. Across cohorts, three versus zero abnormalities corresponded to a 1.54–1.89-percentage-point higher annualized emergence probability, a 3.38–4.58-point lower recovery probability, and a 2.00–2.27-point more adverse expected one-year change in health-deficit proportion.

Conclusions

Greater cardiometabolic abnormality burden was associated with both more observed deficit emergence and less observed recovery across cohorts in China, the United States, and England. Separating these processes adds dynamic information to aggregate deficit measures and may inform chronic-disease and healthy-aging research. The observational, model-based estimates do not establish causal effects, intervention benefits, or predictive utility.

Background

Frailty reflects diminished reserve and greater vulnerability to stressors, but its longitudinal course is not unidirectional [1–2]. In deficit-accumulation models, a summary score is calculated as the proportion of measured health deficits that are present [3–4]. A rising score may therefore arise because new problems appear, existing problems fail to resolve, or both. Analyses limited to the net score cannot distinguish these pathways.

Transitions at the item level offer a different view of the same process. Among items that are currently absent, a 0-to-1 transition represents deficit emergence; among items that are present, a 1-to-0 transition represents observed recovery. Frailty-state studies have shown that improvement remains possible, although it becomes less frequent with advancing age and higher baseline frailty [5]. Work on binary health attributes has further shown that emergence and recovery have distinct age patterns and jointly determine net deficit accumulation [6]. Because survey-item states can also

fluctuate through temporary illness, threshold crossing, or reporting variation, persistence at a third observation provides an important check on interpretation [7].

Cardiometabolic abnormalities are common targets of chronic-disease prevention and management. Central adiposity, elevated blood pressure, and abnormal glycemia cluster with vascular dysfunction, inflammation, impaired energetics, and functional decline [8–9]. Meta-analyses and prospective cohorts have linked metabolic syndrome, metabolically unhealthy obesity, and related metabolic indices to prevalent frailty, incident frailty, and faster frailty-index progression [10–13]. Recent multicohort studies have related metabolic dysregulation counts and metabolic syndrome to aggregate frailty-index burden and incident frailty [14–15]. Cai and Ren further reported that a fixed-weight cardiometabolic biomarker score combining systolic blood pressure, high-density lipoprotein cholesterol, HbA1c, and C-reactive protein was associated with incident frailty among adults aged 65 years or older in HRS, ELSA, and CHARLS [16]. This cross-cohort evidence strengthens the case for cardiometabolic vulnerability as a correlate of later functional health, but threshold-based incident-frailty outcomes still cannot show whether accumulating burden reflects more new deficits, less recovery from existing deficits, or both.

Studies using categorical frailty states have begun to distinguish deterioration from improvement [17–18]. However, a participant may develop some deficits and recover from others during the same interval while remaining in the same frailty category. To resolve these concurrent processes, we harmonized 18 potentially reversible deficits across CHARLS, HRS, and ELSA and modeled item-level emergence and recovery separately. We examined whether a higher baseline count of central obesity, high blood pressure, and known diabetes was associated with more frequent emergence, less frequent recovery, or both among adults aged 50 years or older; placed the two processes on a common absolute scale; and evaluated robustness to current deficit burden, selective follow-up, response reversals, and objective glycemic information. This approach was intended to clarify the dynamic pathways underlying an established cardiometabolic burden–frailty association rather than to develop a prediction score or estimate an intervention effect.

Methods

Study design and data sources

This secondary analysis used three nationally based longitudinal cohorts of adults in later life: the China Health and Retirement Longitudinal Study (CHARLS), the US Health and Retirement Study (HRS), and the English Longitudinal Study of Ageing (ELSA) [19–21]. Cross-cohort harmonization followed the Gateway to Global Aging Data framework and was checked against original cohort documentation [22]. Cohorts were analyzed separately; participant-level records were not pooled.

CHARLS wave 1 (2011) was the baseline, followed by waves 2–4 (2013, 2015, and 2018), using Harmonized CHARLS Version D and source files needed to verify baseline exposures [23]. HRS baseline was 2006, with biennial follow-up through 2022 using the RAND HRS Longitudinal File 2022, Cross-Wave Tracker, and Harmonized HRS Version D [24–27]; the 2022 observation was early release. ELSA wave 4 (2008–2009) was the baseline and waves 5–10 supplied follow-up; wave 11 lay outside the prespecified window. Data were obtained from UK Data Service Study 5050, 50th edition [28].

Common eligibility required a completed baseline interview, age at least 50 years, and classifiable values for all three cardiometabolic components. The common transition-source population additionally required at least one observed adjacent-wave target-item state pair. Emergence and recovery formed independent risk sets according to the observed start state, with an observed next-wave target state, positive elapsed time, an available positive interval-specific

response weight, and complete model variables. Cohort flow and comparisons between final Model A participants and exposure-complete nonparticipants are reported in Additional file 1: Tables S1–S3.

Baseline cardiometabolic abnormalities

The exposure was a fixed baseline count of central obesity, high blood pressure, and known diabetes. Each component was coded as present or absent, yielding a count from 0 to 3. The same constructs were applied across cohorts using prespecified population-appropriate waist thresholds; source variables and coding are provided in Additional file 1: Table S4.

Central obesity was defined by measured waist circumference of at least 90 cm in men or 85 cm in women in CHARLS and at least 102 cm in men or 88 cm in women in HRS and ELSA [8]. High blood pressure was defined by mean measured systolic pressure at least 140 mmHg, mean diastolic pressure at least 90 mmHg, reported physician diagnosis, or antihypertensive treatment. Known diabetes required a reported physician diagnosis or glucose-lowering treatment; laboratory glycemia was not part of the primary definition. The count was not a complete metabolic-syndrome measure because lipid data could not be aligned with the repeated-wave design. Exposure components and related measurements were excluded from the health-deficit panel to prevent direct exposure-outcome overlap.

In a sensitivity analysis, the glycemic component was expanded to include HbA1c at least 6.5% (48 mmol/mol) [29], using baseline biomarker data from CHARLS, HRS, and ELSA [28, 30–31]. HRS used the provided NHANES-equivalent adjusted HbA1c measure. Primary and HbA1c-enhanced definitions were compared within identical HbA1c-complete transition-source samples and identical outcome-specific analytic samples. Because confirmatory testing was unavailable, a single value at or above the threshold was treated as a glycemic abnormality, not confirmed diabetes.

Harmonized health-deficit panel and transitions

Candidate deficits had to be comparable across cohorts, repeatedly observed, directionally consistent, and plausibly reversible. Diagnosed chronic diseases were excluded because a later negative report might reflect reporting inconsistency rather than recovery. The common panel comprised 18 binary items: five activities of daily living, four instrumental activities, four mobility or physical-function items, four depressive-symptom or sleep items, and fair or poor self-rated health. Source variables, response mappings, and waves are listed in Additional file 1: Table S5. Proxy interviews were retained when the target item was observed; non-administered symptom items remained missing.

For transition models and standardization, the baseline common-panel health-deficit proportion was the number of observed deficits divided by the number of observed panel items and was calculated when at least 10 of 18 items were available. The interval-start current proportion used the corresponding ratio across all common-panel items observed at that wave; the primary state-conditional model required at least one. Sensitivity analyses instead required at least 10 or 14 items and recalculated the proportion after excluding the target item. The panel prioritized reversibility and exact cross-cohort comparability over the item count of a general-purpose frailty index [32–33].

Data were organized by participant, deficit item, and adjacent survey interval; observations were not bridged across skipped waves. Emergence was a transition from 0 to 1 and recovery from 1 to 0, with 0 to 0 and 1 to 1 denoting stable states. These terms refer to changes in observed survey-item states rather than biological repair or disease cure.

A three-wave sensitivity analysis assessed sustained emergence (0 to 1 to 1), sustained recovery (1 to 0 to 0), and the corresponding rapid reversals. Sustained transitions were modeled with modified Poisson regression using both interval lengths, Model A or Model B covariates, item and wave fixed effects, participant-clustered inference, and the

product of the two interval-specific response weights re-truncated at the cohort-specific 1st and 99th percentiles. Rapid reversals were tabulated from the same person-item-sequence dataset.

Covariates, response weighting, and mortality

Covariates included interval-start age; baseline sex, education, marital or partnered status, cohort-specific economic quartile, smoking, alcohol consumption, physical activity, and the 18-item common-panel baseline health-deficit proportion; and survey-wave and deficit-item indicators. Except for age and the Model B interval-start current deficit proportion, participant-level covariates were fixed at baseline. Missing categories were retained for categorical covariates; required continuous variables and outcome states were not imputed.

Model A estimated the baseline-health-deficit-adjusted prospective association without conditioning on time-varying current deficit burden. Model B additionally included the interval-start current common-panel proportion and therefore estimated a state-conditional association. Within each cohort and outcome, Models A and B used identical rows and participants.

To reduce bias from selective follow-up, a stabilized inverse-probability-of-next-wave-response weight conditional on survival was estimated for each eligible adjacent interval [34–35]. All cohorts used the same source frame: baseline respondents aged at least 50 years with all three exposure components classifiable; a target-item state pair was not required. Eligible intervals began with a completed interview and required known next-wave vital and interview status and survival to the planned adjacent wave; response meant completion of that interview. Binomial logistic models included wave as a categorical fixed effect in the numerator and additionally included exposure count, age, current deficit proportion, the prespecified covariates, and a cohort-specific local-panel baseline deficit proportion in the denominator. The local panels (25 items in CHARLS, 28 in HRS, and 22 in ELSA; at least 10 observed) were used only for response modeling; transition and mortality models used the common 18-item baseline proportion, also requiring at least 10 items. Probabilities were bounded at 0.01 and 0.99, and weights were truncated at the cohort-specific 1st and 99th percentiles. Each interval weight was applied to its eligible item rows and was not accumulated across intervals. Diagnostics are reported in Additional file 1: Tables S6–S9.

Death was not coded as non-recovery because the next-wave item state was unobserved. Death before the next wave was instead examined in an independent, unweighted person-interval risk set without item expansion, target-item requirements, or response weighting. Eligibility required the common baseline criteria, an interval-start interview, known next-wave vital status, positive elapsed time, and complete covariates. Cause-specific complementary log-log models used log interval years as an offset, participant-clustered inference, the prespecified covariates, the common 18-item baseline proportion requiring at least 10 items, and wave fixed effects, but not current deficit proportion or item effects [36]. ELSA mortality was identified from interview-status and End-of-Life records; because national linkage was unavailable, its estimate was descriptive and mortality results were not meta-analyzed.

Statistical analysis

Emergence and recovery were modeled separately within each cohort using weighted discrete-time complementary log-log regression. Log elapsed years was an offset for unequal interval length. The fixed baseline abnormality count was the primary continuous exposure; hazard ratios (HRs) represent the multiplicative change in transition hazard per additional abnormality. Models included the prespecified covariates and item and wave fixed effects, with participant-clustered robust standard errors for repeated items and intervals [37].

Cohort-specific log HRs were combined by restricted maximum likelihood random-effects meta-analysis with z-based inference [38–39]. I^2 and τ^2 were reported separately for Models A and B. With only three cohorts, heterogeneity measures were interpreted descriptively and cohort-specific estimates remained central.

Categorical models compared one, two, and three abnormalities with none. A participant-clustered two-degree-of-freedom Wald test assessed departure from a linear log-hazard relationship. Across 12 cohort-outcome-model combinations, nonlinearity analyses were exploratory and not adjusted for multiplicity; category-specific estimates were therefore emphasized.

Absolute estimates were obtained by marginal standardization of the categorical models over a common cohort-specific set of transition-source person-intervals with complete standardization covariates and a positive response weight. Exposure was set in turn to 0 and 3, observed covariates, item, wave, and current item states were retained, and interval length was set to one year. Model-based transition probabilities were expressed per 100 eligible items and averaged using the interval-specific response weights.

On the same target intervals, expected one-year change in observed health-deficit proportion was calculated as predicted emergence across absent items minus predicted recovery across present items, divided by the number of observed items. The response-weighted mean placed both processes on a common deficit-proportion scale. Models A and B used identical within-cohort target sets. These standardized contrasts were descriptive, not intervention effects or mediation analyses.

Percentile 95% confidence intervals for absolute estimates used 500 participant-level cluster bootstrap resamples, with the same resample applied to emergence and recovery to preserve covariance. Each replicate used a one-step estimating-equation approximation; response-weight estimates and truncation limits were held fixed, so intervals did not include uncertainty from estimation of the response weights.

Additional analyses varied construction of the interval-start deficit proportion, omitted each deficit domain in turn, restricted intervals to three years or less, and omitted physical activity in CHARLS. Other analyses examined item-specific associations, maximum event share, sustained transitions, rapid reversals, mortality, and the same-sample HbA1c-enhanced definition. Elapsed time was summarized by person-interval. Analyses used R 4.5.1, two-sided tests, and 95% confidence intervals and followed STROBE [40]. No a priori sample-size calculation was performed because all eligible observations in the prespecified waves were included.

Use of generative AI-assisted technologies

During preparation of this manuscript, the authors used Doubao to improve language and readability and to assist with reference-format checking and manuscript organization. The authors reviewed and edited all output and take full responsibility for the content of the manuscript.

Results

Cohorts and transition events

The three exposure-complete cohorts included 25,590 adults aged at least 50 years, of whom 23,702 contributed at least one observed adjacent-wave item-state pair. The emergence analyses included 23,689 participants and 130,915 person-item-interval transition events, whereas the recovery analyses included 22,216 participants and 99,789 such events. Detailed cohort flow, exposure distributions, and outcome-specific risk sets are presented in Table 1 and Additional file 1: Fig. S1. Baseline characteristics and comparisons across successive selection stages are reported in

Additional file 1: Tables S1–S3; these comparisons indicated some selection-related differences between final Model A participants and exposure-complete nonparticipants.

Table 1
Baseline sample, exposure distribution, and transition events under common eligibility

Panel A. Participant flow and transition events								
Cohort	**Baseline respondents**	**Age ≥ 50**	**Exposure complete**	**Transition source**	**Emergence risk set**	**Recovery risk set**	Emergence events*	Recovery events*
CHARLS	17708	13466	10559	9755	9743	9417	45161	36437
HRS	18469	17894	6866	6475	6474	6090	45674	33348
ELSA	11050	10749	8165	7472	7472	6709	40080	30004
Panel B. Distribution of cardiometabolic abnormalities, n (%)								
Cohort	**0**		**1**		**2**		**3**	
CHARLS	3603 (36.9%)		3575 (36.6%)		2255 (23.1%)		322 (3.3%)	
HRS	936 (14.5%)		2091 (32.3%)		2528 (39.0%)		920 (14.2%)	
ELSA	1815 (24.3%)		2794 (37.4%)		2435 (32.6%)		428 (5.7%)	

Note. All cohorts used a baseline age threshold of at least 50 years and the same three primary cardiometabolic constructs, with prespecified population-appropriate waist-circumference thresholds. Primary glycemia used physician diagnosis or glucose-lowering treatment; HbA1c was not required. Exposure distributions are shown for participants with at least one observed adjacent-wave item-state pair. Emergence and recovery used outcome-specific risk sets; Models A and B used identical participants and person-item-interval rows within each cohort and outcome. *Events are person-item-interval transition events.

Relative associations and dose-response

In Model A, each additional cardiometabolic abnormality was associated with a higher hazard of deficit emergence (pooled HR, 1.125; 95% CI, 1.073–1.179) and a lower hazard of recovery (pooled HR, 0.933; 95% CI, 0.921–0.946). Cohort-specific estimates were directionally consistent, although the magnitude of the emergence association varied substantially across cohorts, whereas no between-cohort heterogeneity was detected for recovery. Additional adjustment for the interval-start health-deficit proportion in Model B attenuated both associations without changing their direction or statistical support (Table 2).

Table 2
Primary Model A and Model B associations with deficit emergence and recovery

Source	Outcome	Model A HR (95% CI)	Model B HR (95% CI)	Participants, n	Events, n*	P value or heterogeneity (A / B)
CHARLS	Emergence	1.073 (1.049–1.098)	1.063 (1.042–1.085)	9743	45161	< 0.001 / <0.001
HRS	Emergence	1.144 (1.115–1.173)	1.112 (1.089–1.135)	6474	45674	< 0.001 / <0.001
ELSA	Emergence	1.160 (1.130–1.191)	1.133 (1.106–1.160)	7472	40080	< 0.001 / <0.001
Random-effects meta	Emergence	1.125 (1.073–1.179)	1.102 (1.062–1.143)	3 cohorts	—	I^2 = 90.9% / 88.5%; τ^2 = 0.0016 / 0.0009
CHARLS	Recovery	0.931 (0.912–0.951)	0.941 (0.923–0.959)	9417	36437	< 0.001 / <0.001
HRS	Recovery	0.933 (0.911–0.954)	0.953 (0.935–0.972)	6090	33348	< 0.001 / <0.001
ELSA	Recovery	0.937 (0.914–0.961)	0.950 (0.930–0.971)	6709	30004	< 0.001 / <0.001
Random-effects meta	Recovery	0.933 (0.921–0.946)	0.948 (0.937–0.959)	3 cohorts	—	I^2 = 0.0% / 0.0%; τ^2 = 0.0000 / 0.0000

Note. Stabilized interval-specific next-wave response-weighted estimates are shown. Hazard ratios (HRs) are per additional cardiometabolic abnormality. Model A estimates the baseline-health-deficit-adjusted prospective association; Model B additionally adjusts for interval-start observed deficit proportion. Cohort models used participant-clustered robust inference and item and wave fixed effects. Meta-analysis used restricted maximum likelihood random-effects models with z-based inference. P values and heterogeneity statistics are shown as Model A / Model B pairs; I^2 and τ^2 are descriptive because only three cohorts were available. *Events are person-item-interval transition events.

Categorical estimates were monotonic across all cohort, outcome, and model combinations (Fig. 1; Additional file 1: Tables S10 and S11). Nominal departure from log-linearity was observed for CHARLS recovery in both models and for ELSA emergence in Model B. Because these secondary tests were not adjusted for multiplicity, the corresponding category-specific estimates were emphasized rather than interpreting the per-abnormality HR as strictly linear.

Absolute transition probabilities and expected deficit change

Standardized analyses consistently showed a higher one-year emergence probability and a lower recovery probability for three versus zero abnormalities in all three cohorts; the contrasts were smaller after conditioning on the interval-start health-deficit proportion (Fig. 2; Additional file 1: Table S12). When emergence and recovery predictions were applied jointly to the same current item-state patterns, the expected one-year change in observed health-deficit proportion was 2.00-2.27 percentage points more adverse in Model A and 1.64–1.90 percentage points more adverse in Model B for three versus zero abnormalities (Fig. 3; Additional file 1: Table S13). On the 18-item scale, the Model A contrasts corresponded to approximately 0.36–0.41 additional expected net deficit states over one standardized year.

Persistence and robustness analyses

The associations retained the expected direction when a changed deficit state was required to persist through a third observation (Additional file 1: Table S14; Additional file 1: Fig. S2). Among transitions with an eligible third observation, approximately half of emergence transitions and three-fifths of recovery transitions persisted; the remainder were followed predominantly by rapid reversals (Additional file 1: Table S15).

All response-weight models used the same exposure-complete source-frame definition, and the resulting weights were centered near one, with greater dispersion in ELSA (Additional file 1: Tables S8 and S9). Excluding physical activity from the CHARLS models and omitting each deficit domain in turn produced little change in the principal associations (Additional file 1: Tables S16–S18; Additional file 1: Fig. S3). Associations nevertheless varied across individual deficit items in every cohort-outcome-model combination, and several item-specific estimates were null or opposite in direction (Additional file 1: Table S19; Additional file 1: Fig. S4). Thus, the primary estimates represent average associations across the selected 18-item panel.

Requiring more observed items for the interval-start deficit proportion or excluding the target item from that proportion produced negligible changes in Model B estimates (Additional file 1: Tables S20 and S21). Restricting ELSA to intervals of three years or less also left the findings essentially unchanged (Additional file 1: Tables S22 and S23). In the independent mortality analyses, greater cardiometabolic burden was associated with a higher hazard of death before the next wave in CHARLS and HRS, but not in ELSA; the ELSA result remained descriptive because mortality ascertainment was incomplete (Additional file 1: Table S24).

HbA1c-enhanced sensitivity analysis

Within the HbA1c-observed transition-source samples, expanding the glycemic component to include HbA1c of at least 6.5% reclassified only 2.3%-3.8% of participants to a higher abnormality count (Additional file 1: Tables S25 and S26; Additional file 1: Fig. S5). In identical outcome-specific samples, the pooled Model B estimates under the primary and HbA1c-enhanced definitions were nearly unchanged for both emergence and recovery (Additional file 1: Table S27; Additional file 1: Fig. S6).

Discussion

Across three national cohorts using common age, exposure, and outcome definitions, greater baseline cardiometabolic abnormality burden was associated with two complementary transition patterns: more frequent emergence among deficits that were absent and less frequent recovery among deficits that were present. The directions were consistent across China, the United States, and England and remained after conditioning on interval-start deficit burden, although the magnitude of the emergence association varied substantially. When both processes were standardized to the same person-interval item-state patterns, they jointly corresponded to a more adverse model-based annualized expected change in observed health-deficit proportion. Thus, common cardiometabolic abnormalities were associated with both pathways through which overall deficit burden can increase, rather than only with a higher aggregate score.

These findings extend prior reports linking metabolic dysregulation or metabolic syndrome to aggregate frailty-index burden and incident frailty [14–15]. The recent multicohort study by Cai and Ren found that a fixed-weight score based on systolic blood pressure, high-density lipoprotein cholesterol, HbA1c, and C-reactive protein was associated with incident frailty in HRS, ELSA, and CHARLS [16]. The two studies address complementary rather than duplicative questions. That study evaluated a biomarker-based vulnerability score among initially non-frail adults aged 65 years or older, whereas the present analysis used a readily interpretable count of central obesity, high blood pressure, and known diabetes, included adults from age 50 without excluding those above an aggregate frailty threshold, and

modeled each deficit conditional on its starting state. Categorical multistate studies distinguish broad worsening from improvement [17–18], but may still conceal offsetting item-level changes within an unchanged state. By allowing the same participant to contribute both emergent and recovered deficits during one interval, our framework decomposed the net association into the item-level flows that produced it.

The standardized absolute contrasts placed emergence and recovery on a common deficit-proportion scale and clarified the magnitude of their combined association. Raw recovery probabilities cannot be compared directly with emergence probabilities because recovery is evaluated only among deficits already present, whereas emergence is evaluated among deficits currently absent. Applying both models to the same current item-state pattern addressed this denominator difference. Nevertheless, each cohort was standardized to its own analytic target distribution, and survey weights were not used; the estimates therefore are neither direct between-country comparisons nor nationally representative population-attributable effects. Their 18-item equivalents are model-derived annualized expectations rather than directly observed individual changes, and their clinical importance remains to be established.

Model A estimates the baseline-health-deficit-adjusted prospective association, whereas Model B additionally conditions on interval-start deficit burden. Because this proportion is post-baseline, attenuation between models should not be interpreted as mediation. Alternative minimum-item requirements and leave-one-item-out calculations produced nearly identical state-conditional estimates, indicating that the results were not materially dependent on sparse denominators or inclusion of the modeled item.

Emergence associations were weakest in CHARLS and strongest in ELSA, with substantial between-cohort heterogeneity in both models. Because all cohort-specific emergence HRs exceeded 1.0, heterogeneity concerned magnitude rather than direction; pooled estimates should therefore be read alongside cohort-specific estimates. Recovery estimates were closely aligned. Nominal departure from log-linearity was observed for CHARLS recovery in both models and for ELSA emergence in Model B. For CHARLS recovery, the association was concentrated at counts of two and three abnormalities, whereas the largest increment in ELSA Model B emergence occurred between zero and one. The per-abnormality HR is therefore an average log-linear association across categories. Differences in age structure, calendar period, health-care context, exposure composition, and questionnaire implementation may contribute to emergence heterogeneity.

Associations also differed across individual deficits. Item-by-exposure interactions were present in every cohort-outcome-model combination, and several item-specific estimates were null or opposite in direction. However, no leave-one-domain-out estimate reversed the overall direction. Summary HRs therefore describe average associations across the selected 18-item panel rather than a common effect on every deficit, consistent with the multidimensionality of deficit accumulation [41–42].

Of simple transitions followed by a third observation, approximately half of emergence transitions and three-fifths of recovery transitions persisted; the remaining transitions were often followed by rapid reversal. Temporary symptoms, adaptation, threshold crossing, and reporting variability may contribute [7]. Associations retained the expected direction in sustained-transition models from the same Model A sequence risk sets. Recovery therefore denotes an observed survey-item transition rather than disease cure or biological regeneration.

The HbA1c-enhanced sensitivity analysis further supported the robustness of the exposure definition. Within identical HbA1c-complete outcome samples, adding HbA1c of at least 6.5% modestly shifted exposure counts but left emergence and recovery estimates nearly unchanged. However, these findings do not address selection into biomarker collection or potential misclassification from a single HbA1c measurement.

Several biological pathways could plausibly relate cardiometabolic burden to both transition processes. Adiposity and dysglycemia may promote inflammation, impaired muscle quality, fatigue, depression, and mobility limitation, while hypertension and hyperglycemia may add vascular injury [8–9]. Fitness, weight, and diabetes interventions have been associated with slower deficit accumulation in other settings [43–45], but these pathways and treatment effects were not evaluated here. Although waist circumference, blood pressure, and diabetes status are routinely ascertained and are established targets of prevention and management, the present associations do not show that reducing the abnormality count would prevent deficit emergence or improve recovery. Their public health relevance lies instead in identifying two distinguishable outcomes through which cardiometabolic health may relate to functional aging from midlife onward. Future intervention and surveillance studies should determine whether cardiometabolic prevention influences emergence, recovery, or both, and whether transition-specific measures add useful information beyond aggregate frailty indices.

Strengths include common age and exposure-eligibility rules across cohorts, a unified source frame for response weighting, a harmonized repeatedly observed deficit panel, and separate risk sets for emergence and recovery. Exposure components and largely irreversible diagnoses were excluded from the outcome panel. Participant-clustered inference, person-interval standardization, and replication across three independent cohorts further strengthened the analysis. Sustained-transition analyses, same-sample HbA1c comparisons, alternative constructions of current deficit proportion, and restriction to usual follow-up intervals provided complementary robustness checks.

Several limitations define inference. First, this observational analysis remains susceptible to residual confounding, reverse relationships, and exposure misclassification. Abnormalities were measured only at baseline, and later treatment or metabolic changes were not captured. Lipids could not be harmonized. The three-component count assigned equal weight to central obesity, high blood pressure, and known diabetes; component-specific and combination analyses were not performed. Estimates therefore describe overall burden rather than any component's independent contribution. Diagnosis- or treatment-based diabetes may miss hyperglycemia, and the HbA1c comparison applies only to biomarker-observed samples.

Second, the 18-item panel prioritized reversibility and cross-cohort comparability over comprehensive frailty assessment. It excluded cognition, dichotomized severity, and showed marked item-level heterogeneity; the summary HRs describe an average association across this panel rather than an identical association with every deficit. Some interval-start proportions used incomplete item sets, although stricter item thresholds and leave-one-item-out calculations produced nearly identical estimates. Discrete survey waves cannot capture multiple changes occurring within an interval. Some ELSA intervals exceeded three years, but restriction to intervals of three years or less changed estimates minimally. The HRs and one-year probabilities are model-based annualized associations under the assumed time-scaling relationship and do not identify continuous-time forward and reverse transition intensities.

Third, eligibility required complete exposure classification and observable transition endpoints, and appreciable baseline differences remained between some final Model A participants and exposure-complete nonparticipants. Near-unit weight means and high effective sample sizes indicate limited precision loss from variable weighting among responders but do not demonstrate elimination of attrition bias. Interval-specific response weights addressed measured predictors of next-wave interview response among exposure-complete participants who survived, but not selection into the exposure-complete or biomarker samples or within-interview target-item missingness. Missing categorical covariates used explicit levels rather than multiple imputation, and physical activity was frequently missing in CHARLS. Without survey weights, estimates describe within-cohort associations rather than nationally representative prevalence or population-attributable effects. Sex was included as an adjustment covariate, but sex-specific associations and interactions were not examined. The lower age limit of 50 years means that inference applies to adults in midlife and later life rather than exclusively to populations defined by a conventional older-age threshold.

Fourth, all three abnormalities were uncommon in CHARLS and ELSA, and covariate overlap was not formally quantified. Standardized three-versus-zero contrasts may therefore involve model extrapolation and should be viewed as descriptive rather than intervention effects.

Fifth, death is a terminal competing event. Transition and expected-change estimates are conditional on survival and observation at the next wave, and recovery estimates are additionally conditional on the target deficit being present at interval start. Cause-specific mortality models neither treated death as non-recovery nor corrected transition estimates for competing mortality. Positive mortality associations in CHARLS and HRS indicate possible survivor selection, but these models cannot remove the bias or establish its direction. ELSA mortality ascertainment was also incomplete.

Finally, with only three cohorts, between-cohort variance and heterogeneity were imprecisely estimated, and random-effects confidence intervals used a normal approximation without small-sample correction. Cohort-specific estimates remain central, particularly for emergence, and pooled estimates summarize these cohorts rather than a universal effect.

Conclusions

Higher baseline cardiometabolic abnormality burden was associated with both more frequent emergence and less frequent recovery across a harmonized panel of potentially reversible health deficits in three national cohorts. These item-level processes jointly corresponded to a more adverse model-based annualized expected change in observed deficit proportion. Separating emergence from recovery adds dynamic information to aggregate frailty measures and provides a reproducible outcome framework for chronic-disease and healthy-aging epidemiology. The observational findings do not establish causal effects, intervention benefits, or predictive utility.

Abbreviations

CHARLS
China Health and Retirement Longitudinal Study
CI
confidence interval
ELSA
English Longitudinal Study of Ageing
HbA1c
glycated hemoglobin
HRS
Health and Retirement Study
HR
hazard ratio
NHANES
National Health and Nutrition Examination Survey
STROBE
Strengthening the Reporting of Observational Studies in Epidemiology.

Declarations

Clinical trial number: not applicable.

Ethics approval and consent to participate

The parent studies obtained ethics approval and informed consent from participants as described in their cohort documentation. CHARLS survey and anthropometric procedures were approved by the Peking University Biomedical Ethics Review Committee (IRB00001052-11015), and biomarker collection was approved under IRB00001052-11014. HRS was approved by the University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board (HUM00061128), with additional written consent for physical measurements and biomarker collection. ELSA waves 4–10 received approval from the relevant UK research ethics committees (07/H0716/48, 09/H0505/124, 11/SC/0374, 13/SC/0532, 15/SC/0526, 17/SC/0588, and 21/SC/0030). The present study used deidentified secondary data without participant contact or intervention. The Ethics Committee of the Second Affiliated Hospital of Anhui Medical University determined that this secondary analysis was exempt from further ethics review (YJ-YX2026-311); no additional participant consent was required for the present analysis. The study was conducted in accordance with the Declaration of Helsinki and applicable institutional requirements.

Consent for publication

Not applicable. The manuscript contains no identifiable individual-level information, images, or videos.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Anhui Provincial Department of Education 2025 Key Scientific Research Project for Higher Education Institutions (grant number 2025AHGXZK30664). The funder had no role in the design of the study; collection, analysis, or interpretation of data; preparation of the manuscript; or decision to submit the manuscript for publication.

Author Contribution

RL conceived the study, curated the data, contributed to the methodology and formal analysis, validated the results, prepared the visualizations, and drafted the manuscript. XW contributed resources, software, formal analysis, and visualization. TP contributed to the investigation, methodology, project administration, supervision, and revision of the manuscript. PZ acquired funding and contributed to the investigation, methodology, project administration, supervision, and revision of the manuscript. All authors read and approved the final manuscript.

Acknowledgement

We thank the participants and study teams of CHARLS, HRS, and ELSA. We thank the China Center for Economic Research, National School of Development, Peking University, for providing the CHARLS data. This analysis uses data or information from the Harmonized CHARLS dataset and Codebook, Version D as of June 2021, and the Harmonized HRS dataset and Codebook, Version D as of August 2023, developed by the Gateway to Global Aging Data [22–23,27]. Development of these harmonized datasets was funded by the National Institute on Aging (R01 AG030153, RC2 AG036619, and R03 AG043052). This document used CHARLS Waves 1 through 4 as of June 2021. CHARLS is supported by Peking University, the National Natural Science Foundation of China, the National Institute on Aging, and the World Bank. HRS is sponsored by the National Institute on Aging (U01AG009740) and conducted by the University

of Michigan. The RAND HRS Longitudinal File 2022 (V1) was produced by the RAND Center for the Study of Aging with funding from the National Institute on Aging (U01AG009740 and R01AG073289) and the Social Security Administration. Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Award Number R01AG017644. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. ELSA is funded by the NIHR Policy Research Programme (HEI) 198_1074_03. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Data Availability

The CHARLS data used in this study, including Harmonized CHARLS Version D, are available to registered users through the CHARLS and Gateway to Global Aging Data repositories [23]. HRS public-use data are available to registered users through the HRS data portal [24–27]; access to the 2006 Biomarker Data requires supplemental registration and approval under the HRS Sensitive Health Data process. ELSA data are available to registered users through the UK Data Service (Study Number 5050, 50th Edition) [28]. The applicable data-use agreements do not permit redistribution of participant-level data by the authors. Analysis code and non-disclosive aggregate outputs are available from the corresponding author on reasonable request, subject to the respective data-use agreements.

References

1. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet. 2013;381:752–62. https://doi.org/10.1016/S0140-6736(12)62167-9.
2. Kim DH, Rockwood K. Frailty in older adults. N Engl J Med. 2024;391:538–48. https://doi.org/10.1056/NEJMra2301292.
3. Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating a frailty index. BMC Geriatr. 2008;8:24. https://doi.org/10.1186/1471-2318-8-24.
4. Theou O, Haviva C, Wallace L, Searle SD, Rockwood K. How to construct a frailty index from an existing dataset in 10 steps. Age Ageing. 2023;52:afad221. https://doi.org/10.1093/ageing/afad221.
5. Kojima G, Taniguchi Y, Iliffe S, Jivraj S, Walters K. Transitions between frailty states among community-dwelling older people: a systematic review and meta-analysis. Ageing Res Rev. 2019;50:81–8. https://doi.org/10.1016/j.arr.2019.01.010.
6. Farrell S, Kane AE, Bisset E, Howlett SE, Rutenberg AD. Measurements of damage and repair of binary health attributes in aging mice and humans reveal that robustness and resilience decrease with age, operate over broad timescales, and are affected differently by interventions. Elife. 2022;11:e77632. https://doi.org/10.7554/eLife.77632.
7. Stolz E, Schultz A, Hoogendijk EO, Theou O, Rockwood K. Short-term frailty index fluctuations in older adults: noise or signal? J Gerontol Biol Sci Med Sci. 2025;80:glae262. https://doi.org/10.1093/gerona/glae262.
8. Neeland IJ, Lim S, Tchernof A, Gastaldelli A, Rangaswami J, Ndumele CE, et al. Metabolic syndrome. Nat Rev Dis Primers. 2024;10:77. https://doi.org/10.1038/s41572-024-00563-5.
9. Mishra M, Wu J, Kane AE, Howlett SE. The intersection of frailty and metabolism. Cell Metab. 2024;36:893–911. https://doi.org/10.1016/j.cmet.2024.03.012.
10. Dao HHH, Burns MJ, Kha R, Chow CK, Nguyen TN. The relationship between metabolic syndrome and frailty in older people: a systematic review and meta-analysis. Geriatr (Basel). 2022;7:76. https://doi.org/10.3390/geriatrics7040076.

11. He D, Qiu Y, Yan M, Zhou T, Cheng Z, Li J, et al. Associations of metabolic heterogeneity of obesity with frailty progression: results from two prospective cohorts. J Cachexia Sarcopenia Muscle. 2023;14:632–41. https://doi.org/10.1002/jcsm.13169.
12. McCarthy K, Laird E, O'Halloran AM, Fallon P, Romero-Ortuño R, Kenny RA. Association between metabolic syndrome and risk of both prevalent and incident frailty in older adults: findings from The Irish Longitudinal Study on Ageing (TILDA). Exp Gerontol. 2023;172:112056. https://doi.org/10.1016/j.exger.2022.112056.
13. Zeng P, Li M, Cao J, Zeng L, Jiang C, Lin F. Association of metabolic syndrome severity with frailty progression among Chinese middle and old-aged adults: a longitudinal study. Cardiovasc Diabetol. 2024;23:302. https://doi.org/10.1186/s12933-024-02379-9.
14. Sun W, Lang XL, Yi K. Association between metabolic dysregulation and frailty in four prospective cohorts. J Health Popul Nutr. 2026;45:120. https://doi.org/10.1186/s41043-026-01290-1.
15. Zhai J, Zheng H, Jiang Y, Jiang P. Frailty in relation to metabolic syndrome: cross-sectional and longitudinal insights from the CHARLS and ELSA cohorts. Exp Gerontol. 2026;221:113195. https://doi.org/10.1016/j.exger.2026.113195.
16. Cai Z, Ren Z. Cardiometabolic vulnerability and incident frailty across three aging cohorts: a multicohort longitudinal study of HRS, ELSA, and CHARLS. BMC Geriatr. 2026. https://doi.org/10.1186/s12877-026-08417-3.
17. Njouonkep Sime SK, Irié BBAS, Jules E, Fezeu KL, Pujos-Guillot E, Comte B. Metabolic syndrome as a determinant of bidirectional transitions between frailty states: evidence from the Whitehall II Study. J Cachexia Sarcopenia Muscle. 2026;17:e70341. https://doi.org/10.1002/jcsm.70341.
18. Lv B, Li X, Lu Q, Kong X, Liu B, Wu Y, et al. Association between obesity and the transitions of frailty based on multi-state Markov model: a prospective study. Arch Gerontol Geriatr. 2026;141:106074. https://doi.org/10.1016/j.archger.2025.106074.
19. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol. 2014;43:61–8. https://doi.org/10.1093/ije/dys203.
20. Sonnega A, Faul JD, Ofstedal MB, Langa KM, Phillips JW, Weir DR. Cohort profile: the Health and Retirement Study (HRS). Int J Epidemiol. 2014;43:576–85. https://doi.org/10.1093/ije/dyu067.
21. Steptoe A, Breeze E, Banks J, Nazroo J. Cohort profile: the English Longitudinal Study of Ageing. Int J Epidemiol. 2013;42:1640–8. https://doi.org/10.1093/ije/dys168.
22. Lee J, Phillips D, Wilkens J, Gateway to Global Aging Data Team. J Gerontol B Psychol Sci Soc Sci. 2021;76:S5–16. https://doi.org/10.1093/geronb/gbab050. Gateway to Global Aging Data: resources for cross-national comparisons of family, social environment, and healthy aging.
23. Phillips D, Green H, Petrosyan S, Shao K, Wilkens J, Lee J, Harmonized CHARLS, Version D. (2011–2018), June 2021. Gateway to Global Aging Data, University of Southern California. 2021. https://g2aging.org/ Accessed 17 Jul 2026.
24. Health and Retirement Study. Cross-Wave Tracker File (Final 2022 V1.0). University of Michigan. 2025. https://doi.org/10.7826/OZGE2932
25. RAND Center for the Study of Aging. RAND HRS Longitudinal File 2022 (V1). RAND Center for the Study of Aging. 2025. https://hrsdata.isr.umich.edu/data-products/rand-hrs-longitudinal-file-2022 Accessed 17 Jul 2026.
26. Health and Retirement Study. RAND HRS Longitudinal File 2022 (V1) [dataset]. University of Michigan. 2025. https://hrsdata.isr.umich.edu/data-products/rand-hrs-longitudinal-file-2022. Accessed 17 July 2026.
27. Wang Y, Cole A, Green H, Wilkens J, Phillips D, Lee J, Harmonized HRS, Version D. (1992–2021). Gateway to Global Aging Data. 2023. https://doi.org/10.25549/4smz-hp46

28. Banks J, Cribb J, Coughlin K, Di Gessa G, Kapadia D, Lloyd L et al. English Longitudinal Study of Ageing: Waves 0–11, 1998–2024. 50th ed. UK Data Service. 2026. https://doi.org/10.5255/UKDA-SN-5050-37
29. American Diabetes Association Professional Practice Committee for Diabetes. 2. Diagnosis and classification of diabetes: Standards of Care in Diabetes-2026. Diabetes Care. 2026;49(Suppl 1):S27-49. https://doi.org/10.2337/dc26-S002
30. Zhao Y, Crimmins E, Hu P, Hu Y, Ge T, Kim JK. China Health and Retirement Longitudinal Study 2011–2012 National Baseline Blood Data Users' Guide. Beijing: China Center for Economic Research, National School of Development, Peking University; 2014.
31. Health and Retirement Study. 2006 Biomarker Data (Final V1.0). University of Michigan. 2013. https://doi.org/10.7826/UMWV1829
32. Hanlon P, Welsh SA, Evans NR. Constructing a quality frailty index: you get out what you put in. Age Ageing. 2024;53:afad248. https://doi.org/10.1093/ageing/afad248.
33. Seligman B, Ward M, Ferranna M, Bloom DE, Kenny RA, Orkaby AR. A harmonized frailty index using global aging data. J Gerontol Biol Sci Med Sci. 2026;81:glaf217. https://doi.org/10.1093/gerona/glaf217.
34. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. Am J Epidemiol. 2008;168:656–64. https://doi.org/10.1093/aje/kwn164.
35. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med. 2015;34:3661–79. https://doi.org/10.1002/sim.6607.
36. Andersen PK, Geskus RB, de Witte T, Putter H. Competing risks in epidemiology: possibilities and pitfalls. Int J Epidemiol. 2012;41:861–70. https://doi.org/10.1093/ije/dyr213.
37. Cameron AC, Miller DL. A practitioner's guide to cluster-robust inference. J Hum Resour. 2015;50. https://doi.org/10.3368/jhr.50.2.317. :317 – 72.
38. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21:1539–58. https://doi.org/10.1002/sim.1186.
39. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw. 2010;36:1–48. https://doi.org/10.18637/jss.v036.i03.
40. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med. 2007;147:573–7. https://doi.org/10.7326/0003-4819-147-8-200710160-00010.
41. Johnson L, Guthrie B, Kelly PAT, Anand A, Marshall A, Seth S. Frailty or frailties: exploring frailty index subdimensions in the English Longitudinal Study of Ageing. J Epidemiol Community Health. 2024;78:609–15. https://doi.org/10.1136/jech-2023-221829.
42. Johnson L, Guthrie B, Anand A, Marshall A, Seth S. Accelerated deficit accumulation in frailty and associations with adverse outcomes: a longitudinal population data analysis. Lancet Healthy Longev. 2026;7:100848. https://doi.org/10.1016/j.lanhl.2026.100848.
43. Evans JK, Usoh CO, Simpson FR, Espinoza S, Hazuda H, Pandey A, et al. Long-term impact of a 10-year intensive lifestyle intervention on a deficit accumulation frailty index: Action for Health in Diabetes trial. J Gerontol Biol Sci Med Sci. 2023;78:2119–26. https://doi.org/10.1093/gerona/glad088.
44. Olson K, Houston DK, Ross J, Wing RR, Simpson FR, Pandey A, et al. Associations that cardiorespiratory fitness and body mass index loss have with deficit accumulation frailty. Med Sci Sports Exerc. 2024;56:717–24. https://doi.org/10.1249/MSS.0000000000003353.

45. Park CM, Thanapluetiwong S, Chen X, Oh G, Ko D, Kim DH. Sodium-glucose cotransporter 2 inhibitors, glucagon-like peptide 1 receptor agonists, and frailty progression in older adults with type 2 diabetes. Diabetes Care. 2026;49:147–51. https://doi.org/10.2337/dc25-1031.

Figures

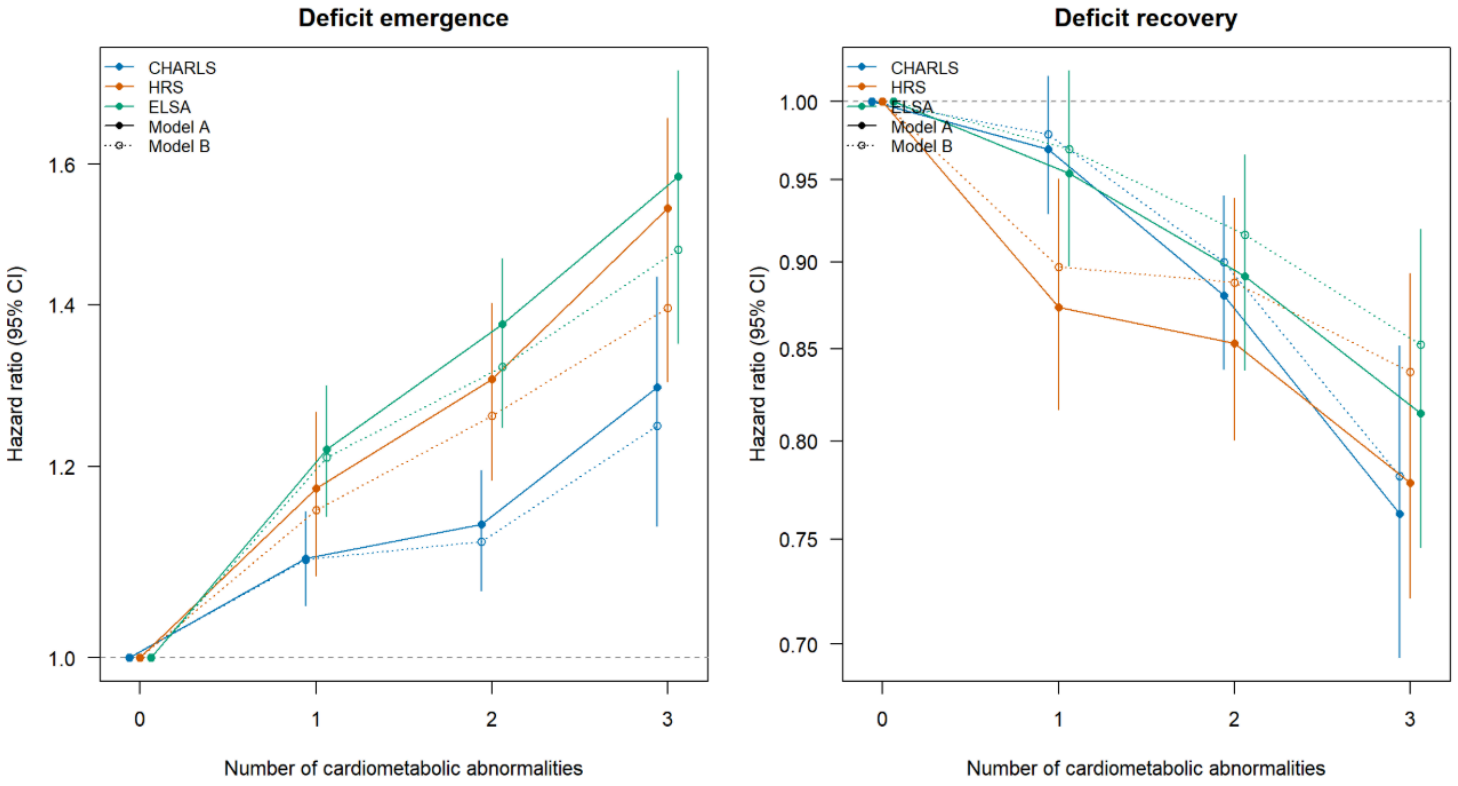


Figure 1

Categorical dose-response for zero to three baseline cardiometabolic abnormalities. The left panel shows emergence and the right panel recovery; zero abnormalities is the reference category. Solid lines denote Model A and dotted lines denote Model B; whiskers are 95% confidence intervals. All point-estimate sequences were monotonic. Nominal departure from log-linearity was observed for CHARLS recovery in Model A (P = 0.047) and Model B (P = 0.020), and for ELSA emergence in Model B (P = 0.032); category-specific estimates are therefore emphasized for these model-outcome combinations.

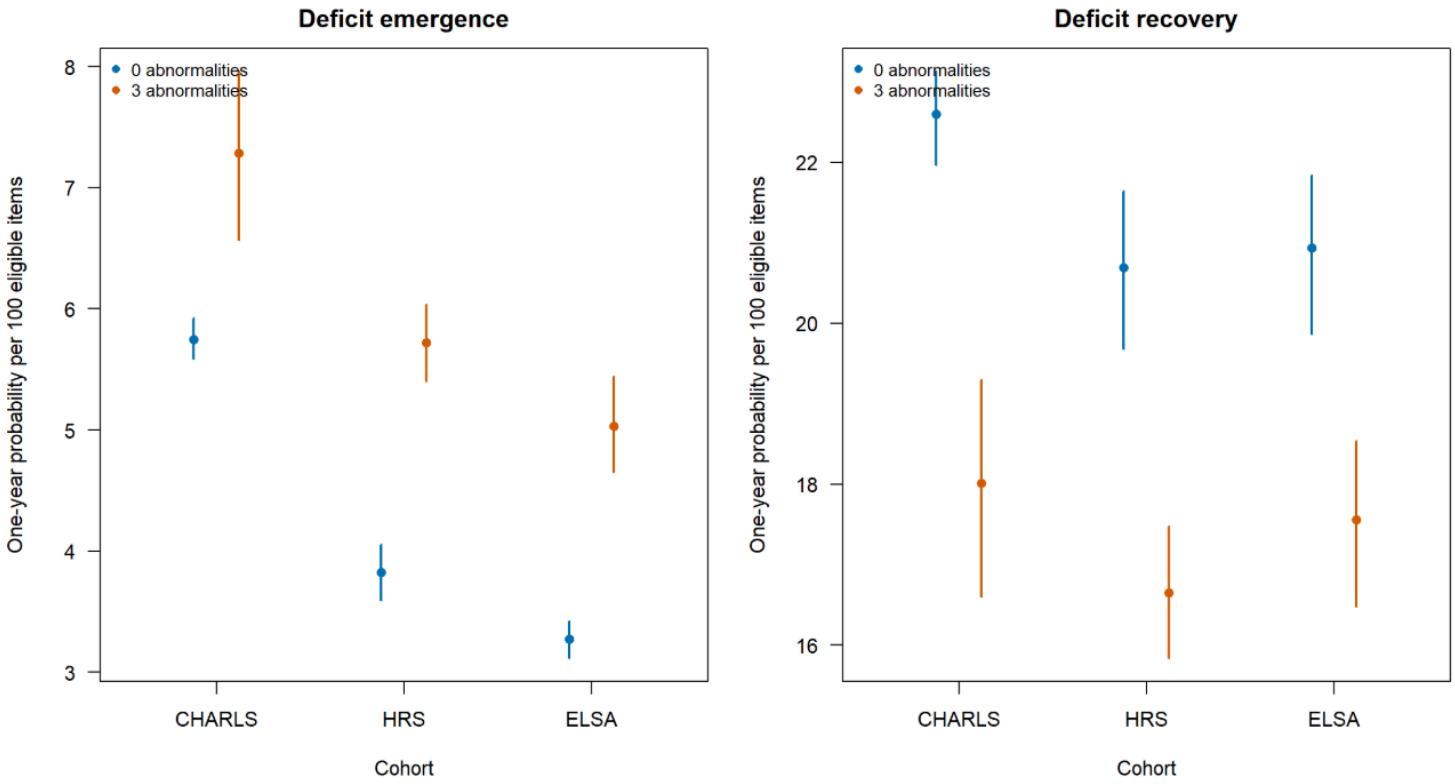


Figure 2

Standardized one-year probabilities of health-deficit emergence and recovery. The left panel shows emergence and the right panel recovery. Values show Model A probabilities per 100 eligible items for zero and three baseline cardiometabolic abnormalities, standardized over the common cohort-specific transition-source person-interval target set with complete standardization covariates and an available positive next-wave response weight; whiskers are 95% confidence intervals from 500 participant-level bootstrap resamples. These are transition probabilities, not item-year rates.

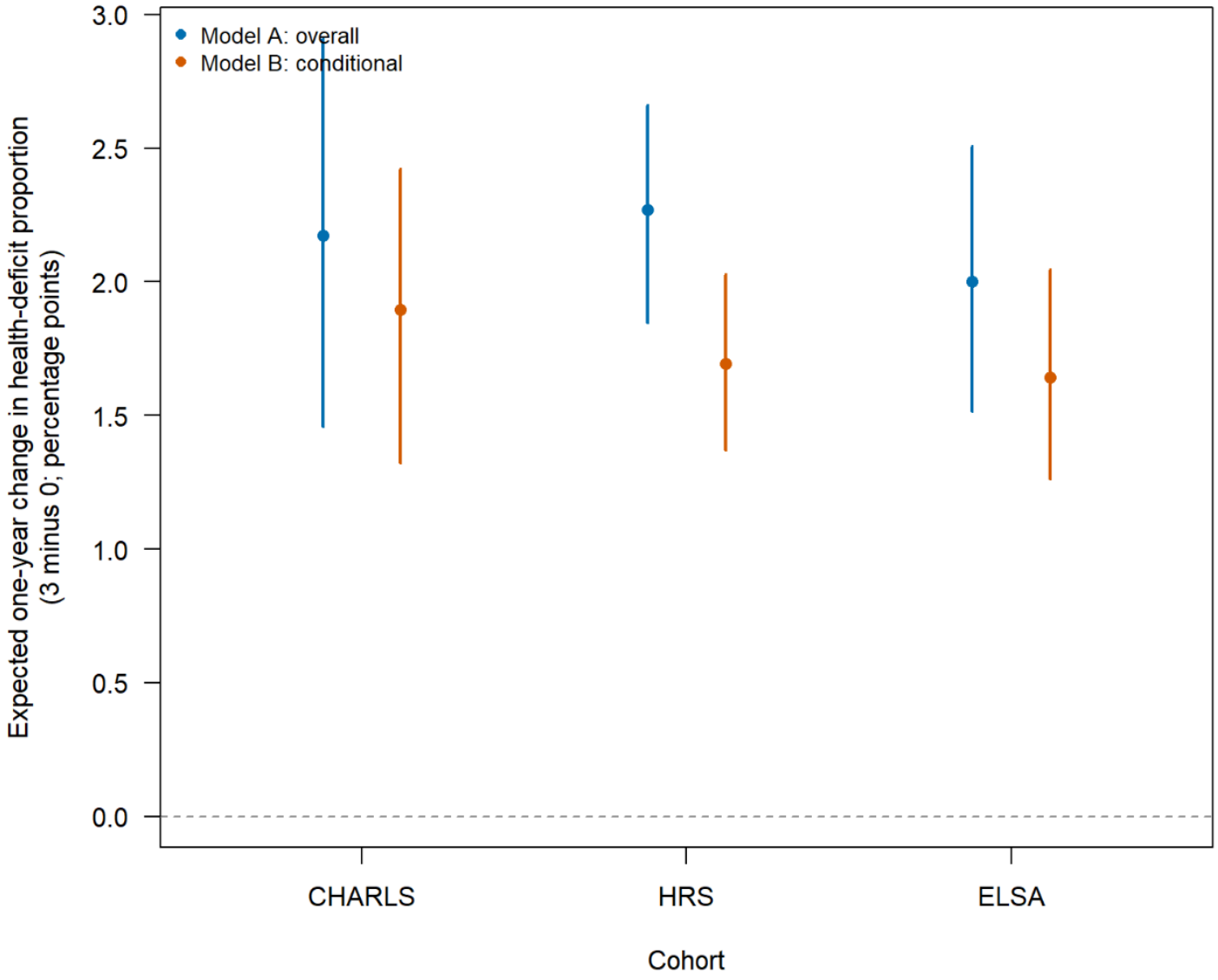


Figure 3

Expected one-year health-deficit change for three versus zero abnormalities. Points show Model A and Model B differences in percentage points; whiskers are 95% confidence intervals from 500 participant-level bootstrap resamples. Expected change was standardized within a common cohort-specific transition-source person-interval target, defined by at least one observed adjacent-wave item-state pair and complete standardization covariates with an available positive next-wave response weight; emergence and recovery predictions were applied jointly to the same observed current item-state pattern.

Supplementary Files

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

- Additionalfile1.docx