The Journal of Frailty & Aging 15 (2026) 100210

Contents lists available at ScienceDirect

The Journal of Frailty & Aging

journal homepage: www.elsevier.com/locate/tjfa

Original Research

# Physical frailty and functional resilience after hip fracture in older adults: a risk-set matched longitudinal study

Haiyan Hu [a,b,c,*], Fangfang Deng [a,b], Bo Chen [a,b,d,*]

[a] Affiliated Central People's Hospital, China Three Gorges University, Yichang, Hubei, 443003, China
[b] Yichang Central People's Hospital, Yichang, Hubei, 443003, China
[c] Institute of Evidence-Based and Translational Medicine, China Three Gorges University, Yichang, Hubei, 443003, China
[d] Hubei Clinical Research Center for Osteoporotic Fracture, China Three Gorges University, Yichang, Hubei, 443003, China

ARTICLE INFO

*Keywords:*
Physical frailty
Functional resilience
Activities of daily living
Hip fracture
Older adults

ABSTRACT

*Background:* Functional trajectories after hip fracture vary widely; whether physical frailty amplifies excess decline or instead marks poorer functional prognosis remains unclear.
*Objectives:* To distinguish these roles of physical frailty after hip fracture.
*Design:* Population-based cohort study with risk-set matching and a case-only prognostic analysis.
*Setting:* Community-based, China.
*Participants:* Adults aged at least 60 years participating in the China Health and Retirement Longitudinal Study, 2011–2020.
*Measurements:* Hip-fracture cases were matched to up to 4 same-wave fracture-free comparators. Preceding-wave physical frailty was assessed using an adapted Fried score (0–5). Matched models estimated fracture-associated excess activities of daily living (ADL) and instrumental activities of daily living (IADL) change and Fried-score interactions; case-only models estimated the probability of being alive with ADL maintained at or returned to the pre-event level at the next survey.
*Results:* The analysis included 252 cases (mean age at the first fracture-report interview, 70.6 years; 132 [52.4%] women) and 1005 comparator records. Relative to matched comparators, hip fracture was associated with greater ADL and IADL deterioration, with average excess increases of 0.73 difficulties in ADL (95% confidence interval [CI], 0.47 to 0.99) and 0.31 difficulties in IADL (95% CI, 0.09 to 0.53). The ADL contrast did not clearly vary by Fried score (interaction per criterion, 0.03; 95% CI, −0.28 to 0.34). Among cases, the adjusted probability of being alive with ADL maintained at or returned to the pre-event level declined from 72.9% at Fried score 0 to 46.4% at score 3 (risk difference, −26.5 percentage points; 95% CI, −50.3 to −2.8).
*Conclusions:* Physical frailty was associated with poorer functional prognosis, whereas evidence that it modified fracture-associated excess decline was unclear. Functional assessment should span frailty levels; frailty may inform support and follow-up intensity but should not determine rehabilitation access.

## 1. Introduction

Preserving functional independence is central to healthy aging, yet hip fracture is a major threat because disability is common and recovery is often incomplete [1–3]. Previously independent patients frequently develop new activities-of-daily-living (ADL) limitations [4], and subsequent trajectories range from near-baseline function to persistent severe disability [5,6]. This heterogeneity highlights the need to identify pre-event characteristics that distinguish maintenance or return to prior function from persistent loss.

Functional resilience describes maintenance or return toward pre-event function after a health stressor and may arise through resistance to initial decline or recovery after deterioration [7,8]. Physical frailty describes a pre-stressor state of diminished physiological reserve and

*Abbreviations:* ADL, activities of daily living; BMI, body mass index; CES-D, center for epidemiologic studies depression scale; CI, confidence interval; IADL, instrumental activities of daily living; IPCW, inverse probability of censoring weighting; RD, risk difference; RR, risk ratio; SD, standard deviation; SMD, standardized mean difference.

* Corresponding authors at: Affiliated Central People's Hospital, China Three Gorges University, Yichang, Hubei, 443003, China.
*E-mail addresses:* hyan87@gmail.com (H. Hu), chenbo71@sina.com (B. Chen).

https://doi.org/10.1016/j.tjfa.2026.100210
Received 18 July 2026; Received in revised form 29 August 2026; Accepted 2 September 2026
Available online 15 September 2026

vulnerability to disproportionate deterioration [9]. Frailty may therefore inform functional resilience alongside stressor characteristics and other individual and contextual factors [7].

Frailty predicts poorer outcomes after hip fracture [10,11], but much of this evidence comes from case-only clinical studies in which frailty is assessed after fracture and pre-fracture function is reconstructed retrospectively [6,10,11]. Such designs identify prognostic vulnerability but do not distinguish background functional decline from change associated with fracture. A population-based comparator study estimated excess disability associated with hip fracture but did not incorporate pre-event frailty [12]. It therefore remains unclear whether frailty modifies fracture-associated functional decline, predicts poorer subsequent function among cases, or both.

Therefore, using data from the China Health and Retirement Longitudinal Study (CHARLS), we sought to clarify how pre-event physical frailty informs functional vulnerability after hip fracture: whether it marks greater susceptibility to the fracture-associated functional decrement, poorer subsequent ADL function, or both. We addressed this question using two distinct analytical approaches. First, a risk-set matched functional change analysis based on the eligible CHARLS cohort compared participants reporting a first hip fracture with same-wave fracture-free comparators to test whether preceding-wave frailty modified the additional ADL or instrumental activities of daily living (IADL) change associated with hip fracture relative to background change. Second, a case-only prognosis analysis restricted to participants with hip fracture examined whether preceding-wave frailty was associated with a lower probability of being alive with ADL maintained at or returned to the pre-event level at the subsequent survey. Together, these analyses distinguished frailty as a potential modifier of fracture-associated functional change from frailty as a prognostic marker among participants with hip fracture.

## 2. Methods

### 2.1. Study design, data source, and population

CHARLS is an ongoing nationally sampled longitudinal survey of community-based Chinese adults aged 45 years or older and their spouses, selected through a multistage probability-proportional-to-size sampling design [13,14]. Face-to-face computer-assisted interviews are conducted approximately every 2 years. We used the 2011, 2013, 2015, 2018, and 2020 public-use waves. For this study, analyses were restricted to participants aged 60 years or older. Detailed eligibility criteria are described below and in the Supplementary Methods.

The original survey was approved by the Institutional Review Board at Peking University (IRB00001052-11015), and participants provided written informed consent. This secondary study used deidentified public-use data and involved no additional participant contact.

### 2.2. Hip fracture ascertainment and event time

At the 2011 baseline, participants reported whether they had ever sustained a hip fracture; at follow-up interviews, they reported whether one had occurred since the preceding interview. An incident case was indexed at the first follow-up interview reporting hip fracture among participants with no hip fracture reported at baseline. Because the exact clinical fracture date was unavailable, the fracture was known only to have occurred during the interval between the preceding interview and the first fracture-report interview. Case ascertainment therefore required survival and continued participation to that first-report interview.

Event time −1 was the last interview before the first fracture report, event time 0 was the first fracture-report interview rather than the clinical fracture date, and event time +1 was the next assigned survey wave. Event times −2 and +2 were retained for supportive pre-report and later estimates when available. First-report waves were 2013, 2015, and 2018. The 2020 wave supplied follow-up outcomes but was not used to identify additional cases because the 2018 wave lacked the complete physical measurements required to construct preceding-wave frailty for cases first reported in 2020.

Potential comparators remained hip-fracture-free through the corresponding first-report wave; participants who subsequently became cases could contribute as comparators only before their own first report. Selected comparator records inherited the case's event-time scale, and observations at or after a comparator's later fracture report were censored. Fig. 1 summarizes participant selection and event-time alignment.

### 2.3. Preceding-wave physical frailty

Physical frailty was operationalized as an adapted Fried phenotype at the preceding survey wave [15–17]. Weakness was dominant-hand grip strength in the lowest quintile within wave, sex, and body mass index (BMI) strata; slowness was average walking-test time in the slowest wave- and sex-specific quintile. Exhaustion required reporting either "everything was an effort" or "could not get going" on at least 3 days in the previous week. Low activity indicated no moderate or light activity lasting at least 10 minutes, and shrinking was BMI $\leq$ 18.5 kg/m$^2$. Thresholds were estimated within each wave among respondents aged 60 years or older with valid measurements. Supplementary Tables S1 and S2 provide item mapping and thresholds.

Observed criteria formed the Fried score (0–5), calculated when at least 4 components were avail-able; a missing fifth component contributed 0, and complete-component scoring was examined in sensitivity analysis. Robust (0), prefrail (1–2), and frail (3–5) categories were used for exact matching and description [15,16]. The continuous score was the main inferential specification; a score-0-versus-$\geq$1 contrast was supportive.

### 2.4. Functional outcomes and indicator of functional resilience

The primary functional outcome was the 6-item ADL difficulty count (0–6), covering dressing, bathing, eating, transferring and room mobility, toileting, and continence [18]. The secondary do-main was the 5-item IADL difficulty count (0–5), covering meal preparation, shopping, money management, medications, and telephone use [19]; the 2011 set assessed household chores rather than telephone use. Each item indicated difficulty or inability to perform the activity, and higher counts indicated greater functional burden. Supportive first-report-wave transitions were worsening, defined as an increase of at least 1 difficulty, and incident difficulty among participants with no pre-event difficulty.

Among cases, maintenance or return to pre-event ADL function was used as an operational indicator of functional resilience [7,8,20]. A favorable outcome required survival to event time +1 with ADL no higher than at event time −1; higher ADL or death was unfavorable, and unknown follow-up remained missing. Deaths recorded by the subsequent assigned survey wave were classified as unfavorable outcomes; exact fracture-to-death timing was unavailable. The indicator combined maintenance without first-report-wave worsening and return after worsening rather than treating both as recovery. In the descriptive decomposition, maintained function required ADL no worse than the pre-event count at both event times 0 and +1; return required first-report-wave worsening followed by an event-time-+1 count no higher than the pre-event count; and persistent loss required a higher count at event time +1. Death and unknown pathways were retained as separate states; an ADL pathway was unknown when survival/function state was unavailable or any of the pre-event, first-report-wave, or subsequent-wave ADL counts needed for classification was missing. The indicator describes the state reached by the subsequent wave, not recovery rate or timing. Full definitions appear in the Supplementary Methods.

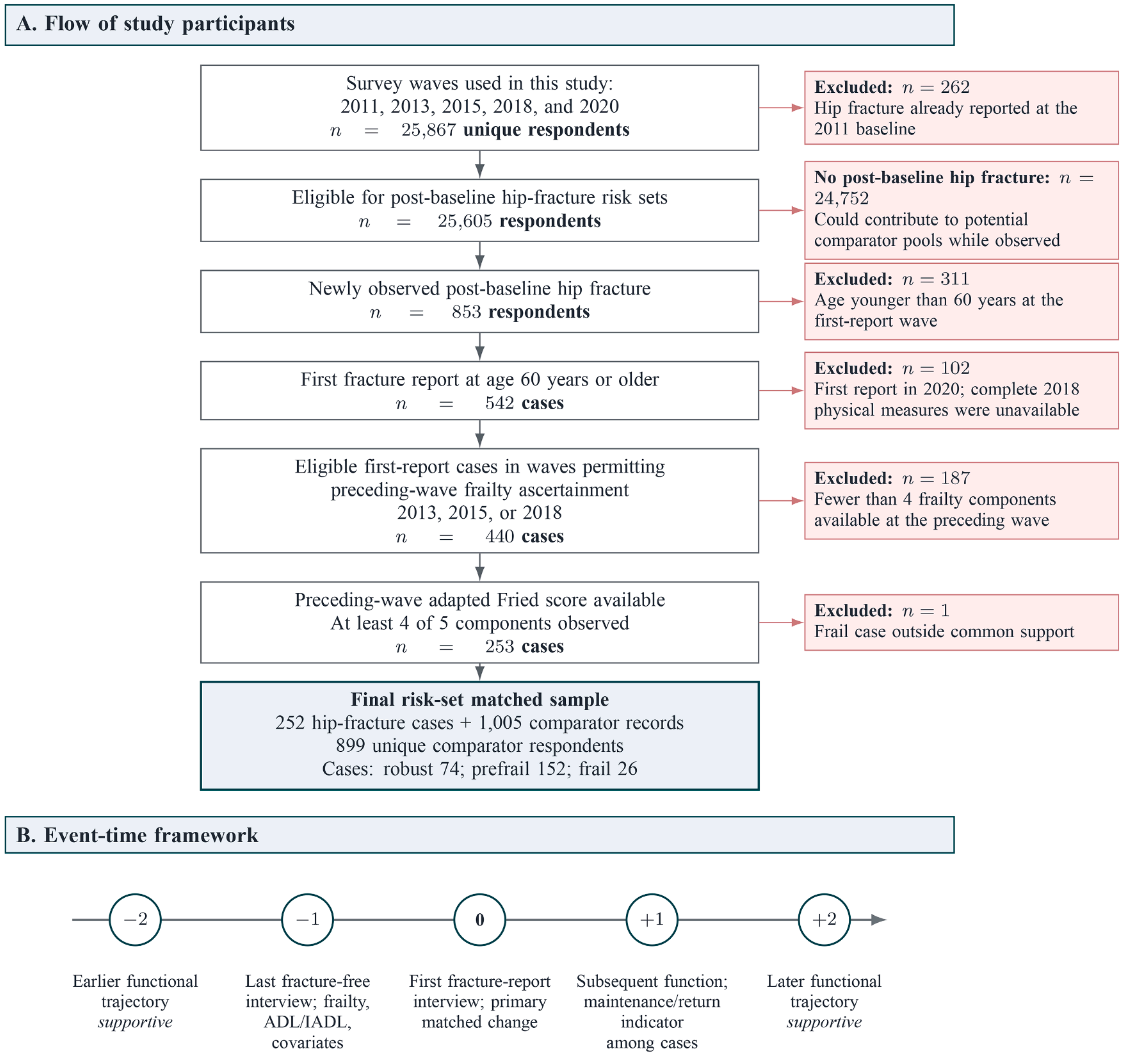


**Fig. 1.** Flow of study participants and event-time framework. *Note.* Panel A traces the source cohort through the post-baseline hip-fracture case frame, age and first-report-wave restrictions, preceding-wave frailty availability, and risk-set matching. Panel B distinguishes the unknown fracture interval from the interviews that define event time. Event time −1 is the last fracture-free interview, when frailty, ADL/IADL, and matching covariates were measured; a hip fracture occurred at an unknown date between −1 and 0. Event time 0 is the first interview reporting that fracture and supplies the primary matched-change comparison. Event time +1 is the next assigned survey wave and supplies subsequent outcomes; the case-only ADL indicator requires survival with the ADL count no higher than at event time −1. Event times −2 and +2 contribute supportive earlier and later trajectory assessments when available. Potential comparators were respondents still fracture-free through the corresponding first-report wave; future cases could contribute before their own report. Event time is indexed by survey waves, approximately 2–3 years apart, rather than by a fixed interval from the clinical fracture date. Frailty was classified as robust (score 0), prefrail (1–2), or frail (3–5). ADL, activities of daily living; IADL, instrumental activities of daily living.

### *2.5. Risk-set matching*

Matching covariates were measured at event time −1 and selected a priori to align demographic, social, medical, and functional characteristics related to hip-fracture risk or background functional change [3,4,6,10,21,22]. Before propensity-score estimation, missing matching covariates were filled with the wave-specific median or mode and accompanied by missingness indicators; frailty and outcomes were not imputed. Within each first-report wave, the propensity-score model included Fried score (0–5), age (years), education (less than lower secondary, upper secondary or vocational training, or tertiary), residence (rural vs urban), partnered status (married or partnered vs not partnered), chronic-condition count (0–9), any inpatient hospitalization in the previous year, and pre-event ADL (0–6) and IADL (0–5) difficulty counts. Fried score, age, chronic-condition count, ADL, and IADL were modeled as continuous linear terms; education was a 3-level factor, and residence, partnered status, and hospitalization were binary factors.

After propensity-score estimation within each first-report wave, the candidate pool was restricted to exact strata of sex and frailty status. Within each stratum, cases were matched with replacement to up to 4 nearest comparator observations on the propensity-score logit within a 0.2-standard deviation (SD) caliper [23,24]. Cases outside common support or without an eligible comparator inside the caliper were excluded. Comparator weights summed to 1 within each set; respondents could contribute to multiple eligible sets before their own fracture. We assessed overlap, reuse, and standardized mean differences (SMDs), prespecifying $|SMD| < 0.10$ as acceptable. Full selection, weighting, and covariate specifications appear in the Supplementary Methods.

### *2.6. Statistical analysis*

#### *2.6.1. Analysis framework*

We evaluated 2 distinct interpretations of frailty after hip fracture.

The matched analysis tested whether Fried score marked differential susceptibility to the fracture-associated ADL and IADL decrement. The case-only prognostic analysis tested whether Fried score marked a lower probability of subsequent maintenance or return to pre-event ADL function. Juxtaposing these estimands distinguished effect modification from prognosis.

*2.6.2. Matched functional change*

Weighted linear event-time models included matched-set fixed effects and interactions among case status, event time, and mean-centered Fried score. With event time −1 as reference, the primary estimands were the case-comparator difference in ADL change at event time 0, evaluated at the weighted mean Fried score, and the change in this contrast per additional Fried criterion. IADL and event time +1 were secondary. Two-way cluster-robust standard errors accounted for matched set and respondent. Weighted linear probability models estimated supportive matched risk differences for worsening and incident difficulty. Heterogeneity was assessed from the continuous-score interaction, not group-specific significance. This design aligned cases and comparators over the same survey interval and excluded already exposed respondents from the risk set [25–27].

*2.6.3. Case-only prognosis*

Logistic models related Fried score to the ADL maintenance-or-return indicator. The fully adjusted model included pre-event ADL and IADL, age, sex, chronic-condition count, first-report wave, an 8-item Center for Epidemiologic Studies Depression Scale (CES-D) score excluding the 2 symptoms used for Fried exhaustion, and cognition combining immediate and delayed recall, time orientation, serial 7s, and figure drawing (range, 0–21). These covariates represented functional headroom and demographic, somatic, affective, and cognitive contributors to subsequent function. Missing expanded covariates were handled using multiple imputation as detailed in the Supplementary Methods. Marginal standardization yielded score-specific probabilities and risk differences, combined using Rubin's rules. A score-0-versus-≥1 specification, complete-case analysis, survivor-conditioned analysis, IADL indicator, and descriptive ADL pathways were supportive.

*2.6.4. Additional and sensitivity analyses*

Additional analyses examined selection, sex, follow-up observability, improvement relative to the maximum available opportunity, and an older-age subgroup comprising cases aged at least 75 years at the first fracture-report interview. The older-age analysis retained the original matched sets and was interpreted only if all displayed post-match absolute SMDs remained below 0.10. Sensitivity analyses varied frailty-component requirements, matching specifications, outcome models, the IADL item set, inclusion of 2020 outcomes, and assumptions concerning missingness and survival. Full specifications are provided in the Supplementary Methods. Analyses used R 4.5.3 and followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidance [28].

## 3. Results

*3.1. Study population and analytic samples*

Among 440 age-eligible hip-fracture cases in the 2013, 2015, and 2018 risk sets, the mean first-report-wave age was 71.1 years and 239 (54.3%) were women. Of these, 253 had an adapted Fried score at the preceding wave; wave-specific availability is reported in Supplementary Table S3. Be-fore matching, 74 were robust, 152 prefrail, and 27 frail. One frail case was outside common sup-port, leaving 252 cases and 1,005 matched comparator records representing 899 unique respondents (Fig. 1A). In the final case sample, mean age was 68.3 years at the preceding frailty assessment and 70.6 years at the first fracture-report interview; 132 of 252 participants (52.4%) were women. Detailed age distributions across selection stages are reported in Supplementary Table S14.

Cases and comparators were closely aligned on pre-event characteristics after matching (Table 1). The largest absolute SMD was 0.055 for pre-event ADL count; all other displayed differences were smaller. Frailty category and sex were identical by design, and the mean adapted Fried score was 1.13 among cases and 1.14 among comparators. The full balance assessment and propensity-score overlap are shown in Supplementary Table S4 and Supplementary Figs. S1 and S2; outcome observability is reported in Supplementary Table S5.

*3.2. Hip-fracture-associated functional change across frailty levels*

At the primary event time 0, ADL difficulty rose more among cases than among matched comparators: the average matched change contrast was 0.73 difficulties (95% CI, 0.47 to 0.99; $P < .001$). The corresponding secondary IADL contrast was 0.31 difficulties (95% CI, 0.09 to 0.53; $P = .006$) (Table 2).

For the primary ADL outcome, the event time 0 matched change contrast changed by 0.03 difficulties per additional Fried criterion (95% CI, −0.28 to 0.34; $P = .854$). The corresponding secondary IADL interaction was 0.02 difficulties (95% CI, −0.25 to 0.29; $P = .901$) (Table 2; Supplementary Fig. S6). These near-zero interaction estimates provided no clear evidence that higher Fried scores marked greater susceptibility to the fracture-associated increment; their confidence intervals,

**Table 1**
Pre-event characteristics after risk-set matching.

| Characteristic | Hip-fracture cases | Matched comparators | Weighted absolute SMD |
|---|---|---|---|
| Age at preceding survey wave, y | 68.3 (7.7) | 68.3 (7.6) | 0.007 |
| Women | 132 (52.4%) | 525 (52.2%) | < .001 |
| **Education level, *n* (%)** | | | |
| Less than lower secondary | 244 (96.8%) | 981 (97.6%) | 0.044 |
| Upper secondary or vocational training | 6 (2.4%) | 18 (1.8%) | 0.037 |
| Tertiary education | 2 (0.8%) | 6 (0.6%) | 0.024 |
| Rural residence | 208 (84.2%) | 837 (85.5%) | 0.035 |
| Partnered | 190 (75.4%) | 752 (74.8%) | 0.015 |
| Chronic condition count | 1.5 (1.3) | 1.5 (1.3) | 0.036 |
| Hospitalized in prior year | 46 (18.3%) | 197 (19.6%) | 0.035 |
| Pre-event ADL difficulty count | 0.7 (1.3) | 0.7 (1.2) | 0.055 |
| Pre-event IADL difficulty count | 0.7 (1.2) | 0.7 (1.1) | 0.052 |
| Adapted Fried score, mean (SD) | 1.13 (0.98) | 1.14 (0.99) | 0.007 |
| **Adapted Fried frailty category, *n* (%)** | | | |
| Robust (score 0) | 74 (29.4%) | 296 (29.4%) | < .001 |
| Prefrail (scores 1–2) | 152 (60.3%) | 607 (60.3%) | < .001 |
| Frail (scores 3–5) | 26 (10.3%) | 102 (10.1%) | < .001 |

*Note.* Values are mean (SD) or number (percentage). All characteristics, including age, were measured at the survey wave preceding the case's first-report wave. At the first-report wave, cases had a mean age of 70.6 years. The table includes 252 cases and 1,005 matched comparator records; 899 comparator respondents were unique because reuse was allowed. Comparator counts and percentages are unweighted descriptions of matched records; SMDs use the matched analysis weights. Education is displayed using the same 3-level factor used in propensity-score estimation; its SMD is shown separately for each indicator. Because sex and frailty category were exact-matching variables, their weighted absolute SMDs are exactly 0.000 rather than rounded representations of residual imbalance. Percentages use observed denominators and may not sum to the full sample when a displayed covariate was missing; matching used filled values and missingness indicators. Partnered status comprised married or partnered respondents. SD indicates standardized deviation; SMD indicates standardized mean difference; ADL, activities of daily living; IADL, instrumental activities of daily living.

**Table 2**
Fracture-associated functional change and modification by preceding-wave frailty.

| Outcome | Average excess change (95% CI) | *P value* | Modification per Fried criterion (95% CI) | *P value* |
|---|---|---|---|---|
| ADL difficulty count | 0.73 (0.47 to 0.99) | < .001 | 0.03 (−0.28 to 0.34) | .854 |
| IADL difficulty count | 0.31 (0.09 to 0.53) | .006 | 0.02 (−0.25 to 0.29) | .901 |

*Note.* Event time −1 was the reference, and event time 0 was the first fracture-report interview. Average excess change is the case-comparator difference in change at the weighted mean Fried score (1.13); modification is the change in this contrast per additional Fried criterion. Positive values indicate greater increases in difficulty among cases. Models included matched-set fixed effects and used two-way cluster-robust confidence intervals by matched set and respondent. ADL indicates activities of daily living; IADL, instrumental activities of daily living.

however, did not establish identical contrasts at every score.

Supporting analyses extended these findings across time and onto binary clinical transitions (Supplementary Figs. S3 and S4). Complete average matched contrasts at event times −2, 0, +1, and +2 are reported in Supplementary Table S6. At event time 0, cases had a 24.8-percentage-point higher risk of ADL worsening than matched comparators (95% CI, 17.5 to 32.1). Among participants without pre-event ADL difficulty, the risk of incident ADL difficulty was 26.0 percentage points higher (95% CI, 15.9 to 36.2; Supplementary Table S7).

### 3.3. Preceding-wave frailty and maintenance or return to pre-event ADL function

For the case-only ADL indicator of functional resilience, the fully adjusted probability of being alive at event time +1 with ADL maintained at or returned to the pre-event count was 72.9% (95% CI, 63.7% to 82.1%) at Fried score 0 and 46.4% (95% CI, 28.9% to 64.0%) at score 3. The principal score-3-versus-0 difference was −26.5 percentage points (95% CI, −50.3 to −2.8). Supportive score-1-versus-0 and score-2-versus-0 differences were −8.1 percentage points (95% CI, −15.0 to −1.3) and −17.1 percentage points (95% CI, −32.4 to −1.9), respectively. The secondary IADL maintenance-or-return estimate was less precise: the score-3-versus-0 difference was −15.5 per-centage points (95% CI, −38.9 to 7.8) (Fig. 2; Supplementary Table S9).

The descriptive ADL pathway decomposition separated the 2 favorable response patterns combined in the binary indicator. Function was maintained in 56.8% of cases with no Fried criterion and 36.0% of cases with at least 1 criterion; persistent functional loss occurred in 18.9% and 33.1%, respectively. Return after first-report-wave worsening was 17.6% and 17.4%. By the next assigned survey wave after the first fracture-report interview, 13 of 252 cases (5.2%) were classified as dead and 16 (6.3%) had an unknown ADL pathway (Fig. 3; Supplementary Tables S10 and S16).

Among 121 cases eligible for the maximum-potential-improvement analysis (alive with observed function and first-report-wave ADL count > 0), the descriptive bootstrap median difference for cases with at least 1 Fried criterion versus none was −55 percentage points (95% CI, −75 to 50). Group-specific ADL distributions and supportive IADL results are reported in Supplementary Table S16.

The binary supportive analysis compared no criterion with at least 1 criterion. Probabilities of maintaining or returning to pre-event ADL function were 74.4% among robust cases and 58.9% among cases with at least 1 criterion, an adjusted difference of −15.5 percentage points (95% CI,−30.3 to −0.7). Adjustment for pre-event function, cognition, depressive symptoms, comorbidity, age, sex, and first-report wave attenuated this difference by 16.6% relative to the minimally adjusted estimate (Supplementary Table S8).

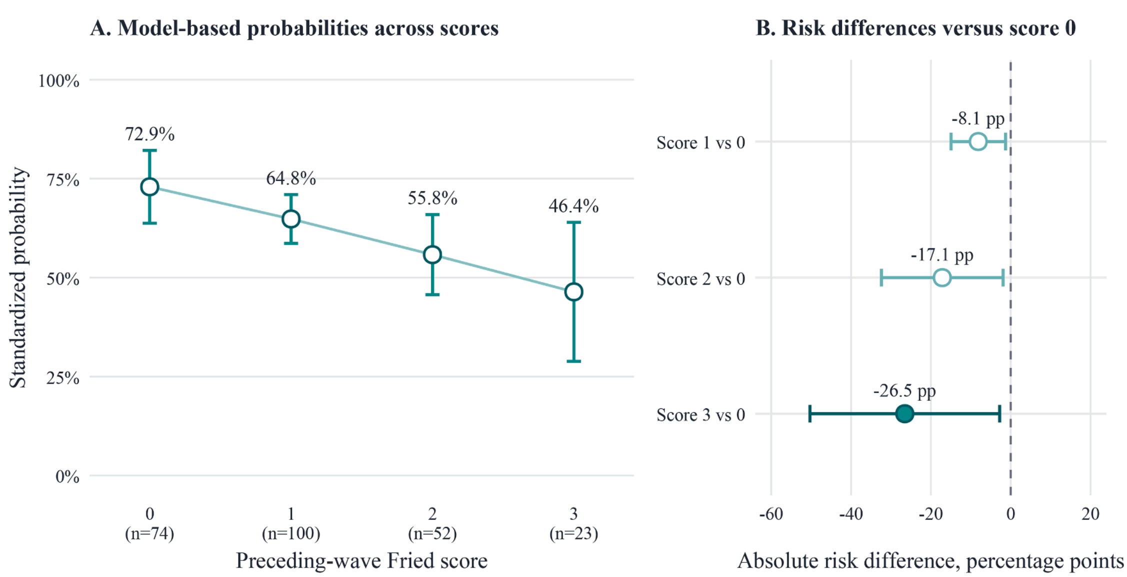


**Fig. 2.** Model-standardized probability of maintaining or returning to pre-event ADL function across preceding-wave Fried scores. *Note.* Panel A shows multiply imputed standardized probabilities at observed integer Fried scores 0–3 among hip-fracture cases; the x-axis gives the number of cases at each score. Lines connect model-based predictions for visual guidance; they do not represent observed trajectories or establish an empirical dose–response relationship. Score 4 was not shown because only 3 cases had that score, and no case had score 5. Panel B shows model-standardized absolute risk differences for Fried scores 1, 2, and 3 relative to score 0. The score-3-versus-0 difference was the principal contrast; intermediate contrasts are supportive. The plotted outcome was being alive at event time +1 with ADL no worse than at event time −1. It combines maintenance and return after first-report-wave worsening; Fig. 3 displays these pathways separately. The model adjusted for both pre-event functional counts, age, sex, first-report-wave indicators, chronic conditions, CES-D8, and cognition. Death was unfavorable, unknown follow-up remained missing, and intervals are Rubin-pooled 95% CIs using HC3 variance within each of 40 imputations.

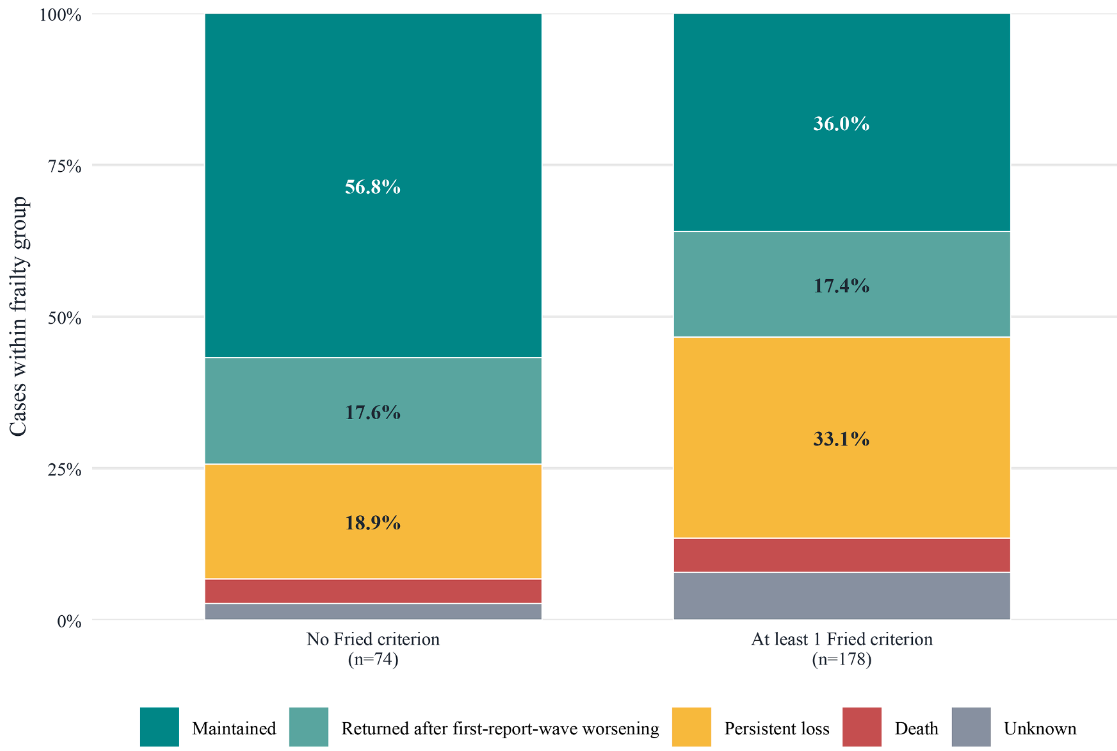


**Fig. 3.** Subsequent ADL functional pathways by preceding-wave Fried criteria. *Note.* Bars partition all hip-fracture cases into 5 mutually exclusive subsequent ADL pathways: maintained function, return after first-report-wave worsening, persistent functional loss, death, or unknown pathway. The first 2 path-ways form the favorable binary indicator used in Fig. 2. Groups compare cases with no Fried criterion and those with at least 1 criterion. Across groups, 13 of 252 cases were classified as dead by the subsequent assigned survey wave and 16 had an unknown ADL pathway because survival/function state was unavailable or at least 1 ADL count needed for pathway classification was missing. These survey-wave states are not fixed-time postoperative mortality or loss-to-follow-up rates. Percentages use all cases within each group; detailed definitions and Wilson 95% CIs are reported in Supplementary Table S10. ADL, activities of daily living; IADL, instrumental activities of daily living.

### *3.4. Sensitivity analyses*

The analysis restricted to cases aged at least 75 years did not meet the prespecified post-match balance criterion (76 cases; maximum absolute SMD, 1.015) and was not interpreted. Among women, the event-time 0 ADL average contrast was 0.91 difficulties (95% CI, 0.52 to 1.29), and the Fried-score interaction was −0.16 per criterion (95% CI, −0.57 to 0.26; Supplementary Table S15).

When 2020 outcomes were excluded, the event-time +1 ADL average contrast was 0.28 difficulties (95% CI, −0.03 to 0.59), and the interaction was −0.25 per criterion (95% CI, −0.59 to 0.09). With IPCW for event-time +1 observability, the corresponding estimates were 0.30 (95% CI, 0.04 to 0.56) and −0.02 (95% CI, −0.32 to 0.28), respectively (Supplementary Tables S12 and S13; Supplementary Fig. S5).

In the survivor-conditioned analysis of 224 cases alive with observed subsequent function, the adjusted ADL maintenance-or-return risk difference for cases with at least 1 Fried criterion versus none was −15.0 percentage points (95% CI, −29.9 to −0.2; Supplementary Table S9). Remaining prespecified sensitivity results are reported in Supplementary Tables S8, S11, S12, S13, and S15.

## 4. Discussion

### *4.1. Principal findings*

This study examined community-dwelling Chinese adults aged 60 years or older participating in CHARLS. Compared with contemporaneous fracture-free respondents, participants reporting hip fracture experienced greater ADL and IADL deterioration, but preceding-wave frailty did not clearly identify a larger fracture-associated decrement. Among participants with hip fracture, greater preceding-wave frailty was associated with a lower likelihood of maintaining or returning to prior ADL function. Taken together, these findings suggest that pre-event frailty is more informative as a prognostic marker of post-fracture functional outcome than as a modifier of the excess functional decline attributable to hip fracture itself. Clinically, this distinction supports functional assessment across the frailty spectrum, with frailty informing anticipatory support, rehabilitation planning, and follow-up intensity rather than defining recovery potential.

### *4.2. Frailty, hip fracture, and functional resilience: distinguishing prognosis from effect modification*

Physical frailty and functional resilience describe related but distinct stages of a stress response: frailty reflects pre-event vulnerability, whereas resilience is inferred from subsequent functional response [7, 9]. The matched analysis showed no clear frailty gradient in fracture-associated excess ADL or IADL deterioration. In the case-only analysis, however, higher Fried scores were associated with a lower probability of being alive with ADL maintained at or returned to the pre-event level at the next survey. Thus, frailty showed a prognostic association without clear evidence that it amplified the comparator-referenced decrement. Near-zero interactions do not establish identical effects across frailty levels, particularly because high-score estimates were imprecise.

A reserve-based interpretation is consistent with this pattern. Frailty may summarize pre-event physiological and functional vulnerability relevant to reaching a favorable subsequent state with-out necessarily producing a larger comparator-referenced change during the fracture-report interval. Because the analyses address different estimands and periods, the case-only association cannot establish the mechanism underlying the near-zero interaction. Attenuation after adjustment suggests that Fried frailty shares prognostic information with pre-event function, cognition, depressive symptoms, and comorbidity; incremental prediction and causal decomposition were not evaluated.

We operationalized functional resilience as a clinically meaningful subsequent state—being alive with ADL maintained at or returned to the pre-event level—rather than as the process or speed of recovery. Descriptively, frailty-group differences chiefly reflected less maintenance and more persistent loss; return after worsening occurred in both groups, but the analysis did not test equivalent recovery. The study therefore identifies variation in the likelihood of reaching a favorable functional state after a first fracture report. More closely timed clinical studies are needed to distinguish resistance from recovery and identify their determinants.

### 4.3. Comparison with previous studies

Clinically timed hip-fracture cohorts have shown heterogeneous and often incomplete functional recovery, with disability persisting for many patients [4–6,29]. These cohorts characterize recovery trajectories but generally do not include nonfracture comparators needed to estimate excess fracture-associated change. Population-based comparisons demonstrate substantial excess disability relative to controls, although pre-event differences in health and function can complicate interpretation [12]. Together, these literatures leave unresolved whether pre-event frailty amplifies the fracture-associated functional decrement or predicts the functional state subsequently reached.

Hip-fracture frailty studies have largely addressed prognosis, linking frailty with poorer functional and perioperative outcomes within case cohorts [10,11]; studies outside fracture settings like-wise associate frailty with disability onset and recovery [17,30,31]. The two analytical perspectives used here—risk-set matched functional change and case-only subsequent ADL prognosis—separate effect modification from prognostic stratification. Recent matched Health and Retirement Study (HRS) analyses also show that event-centered patterns differ across health domains: cognitive vulnerability was evident before the fracture-report index, whereas depressive symptoms increased mainly around the reporting interval [32,33]. Across these studies, pre-event vulnerability, event-associated disruption, and subsequent adaptation appear to contribute differently to cognitive, affective, and functional outcomes.

The CHARLS analysis provides community-based Chinese data with prospectively observed preceding-wave frailty and function [13]. At the first fracture-report interview, cases had a mean age of 70.6 years and 52.4% were women, contrasting with many hospital cohorts of patients in their 80s with a stronger female predominance [5,34]. Analytic inclusion also depended on survival and continued participation to that interview and availability of preceding-wave frailty measures. These characteristics limit direct transportability to consecutive hospitalized patients while positioning the survey-based evidence as a complement to clinically timed cohorts that capture acute treatment, recovery, and mortality.

### 4.4. Clinical implications: interpreting frailty within comprehensive geriatric assessment

Comprehensive geriatric assessment integrates medical, psychological, functional, and social in-formation to inform individualized and coordinated care [35]. In hip-fracture care, it is commonly embedded within orthogeriatric and multidisciplinary rehabilitation pathways [36]. The present findings clarify the place of pre-event physical frailty within multidomain prognostic assessment.

Fracture-associated ADL and IADL deterioration occurred without a clear gradient across Fried scores, supporting direct functional assessment across frailty levels (Table 2). Higher preceding-wave Fried scores were also associated with a lower probability of being alive with ADL maintained at or returned to the pre-event level (Fig. 2; Supplementary Table S9). Frailty may therefore provide prognostic context for rehabilitation needs, discharge support, caregiver involvement, and follow-up, alongside direct assessment of function.

Frailty is most informative when considered alongside cognition, depressive symptoms, comorbidity, and social circumstances. Return after observed worsening occurred in both frailty groups; the descriptive pathway analysis leaves between-group recovery differences uncertain (Fig. 3; Supplementary Table S10). Frailty alone therefore provides insufficient grounds for judging recovery potential. Prospective clinical studies can assess the added predictive value of multidomain assessment, identify useful decision thresholds, and evaluate frailty-informed care pathways.

### 4.5. Strengths and limitations

Key strengths were the prospective ordering of preceding-wave frailty, directly observed pre-fracture ADL and IADL, repeated within-person functional measurements, and well-balanced same-wave risk-set comparators. Together, these features supported two complementary estimands: comparator-referenced excess functional change and case-only maintenance or return to pre-event ADL function. Retaining death as an unfavorable state also avoided restricting the primary prognostic outcome to subsequent survivors.

Hip fracture was self-reported and dated only to the interval between surveys. Event time 0 represented the first fracture-report interview rather than the clinical fracture date; variation in elapsed time and reporting error could attenuate event-aligned contrasts or obscure effect modification. Fractures followed by death before a reporting interview could not be identified, whereas deaths after the first-report interview were retained in the prognostic outcome. The design therefore cannot characterize acute postoperative change, fixed-time mortality from the fracture date, or the timing of functional recovery.

Selection limits transportability. Of 440 age-eligible first-report cases, 253 had calculable preceding-wave Fried scores and 252 were matched; only 26 were classified as frail. The analytic cases were younger and included a lower proportion of women than typical hospital cohorts. Restricting the original matched sets to cases aged at least 75 years failed the prespecified balance criterion. The findings therefore apply most directly to survey-observed participants with preceding-wave participation and physical measurements, rather than to consecutive, predominantly female hospitalized patients in their 80s.

ADL and IADL are bounded counts, so scale position constrains both further deterioration and numerical improvement. The conditional maximum-potential-improvement measure made one form of headroom explicit but required survival, observed function, and first-report-wave difficulty greater than 0; it did not estimate recovery rate or the proportion of fracture-attributable loss regained. The adapted Fried phenotype used survey-specific thresholds and a modified shrinking definition, and sparse high scores limited precision. Multiple imputation assumed data were missing at random conditional on included variables, while residual bias may remain from informative missingness, unobserved pre-report deaths, and unmeasured fracture severity, treatment, rehabilitation, medications, nutrition, delirium, and social support.

## 5. Conclusions

In this community-based cohort of older adults, hip fracture was associated with marked functional deterioration and persistent ADL impairment. Greater pre-fracture frailty was associated with poorer maintenance and return to pre-event ADL function, but not with a clearly larger excess functional decline attributable to hip fracture itself. These findings suggest that frailty is more in-formative as a prognostic marker of post-fracture functional vulnerability than as a modifier of the fracture-associated decrement. Frailty should therefore inform post-fracture risk stratification and follow-up intensity, but it should not be used alone to determine rehabilitation access or recovery potential.

## Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the authors used OpenAI ChatGPT 5.5 to assist with manuscript formatting and language editing. The authors reviewed and edited all outputs and take full responsibility for the content of the manuscript.

## Ethics

The original survey protocol was approved by the Institutional Review Board at Peking University (IRB00001052-11015), and participants provided written informed consent. This study used deidentified public-use data.

## Funding and sponsor role

This study was supported by the National Clinical Key Specialty Construction Program of China. The funder had no role in study design; data collection, analysis, or interpretation; manuscript preparation; or the decision to submit.

## CRediT authorship contribution statement

**Haiyan Hu:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Fangfang Deng:** Writing – original draft, Data curation. **Bo Chen:** Writing – review & editing, Supervision, Resources, Funding acquisition, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. This statement is made on behalf of all authors.

## Acknowledgements

We are grateful to the participants of the China Health and Retirement Longitudinal Study for generously contributing their time and information, and to the investigators and field staff whose work made this study possible.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tjfa.2026.100210.

## Data availability

The public-use China Health and Retirement Longitudinal Study data are available from the official study repository at https://charls.pku.edu.cn/.

## References

[1] Chang Y, Huang Y, Li R, Gui L. Association between diabetes of different durations and hip fracture in middle-aged and older people: a stratified cohort study from CHARLS 2011-2020. BMC Public Health 2025;25(1):691. https://doi.org/10.1186/s12889-025-21923-0.

[2] Jing X, Hou S, Li Y. Association of average cumulative sensory impairment with falls and hip fractures in older adults: evidence from a longitudinal study. J Nutr Health Aging 2026;30(8):100895. https://doi.org/10.1016/j.jnha.2026.100895.

[3] Zhang X, Xu D, Yang Y, Zeng Y, Yu H, Fang T, et al. Intrinsic capacity and risk of hip fracture in older adults: evidence from five multinational aging cohorts. J Nutr Health Aging 2026;30(2):100773. https://doi.org/10.1016/j.jnha.2026.100773.

[4] Ouellet JA, Ouellet GM, Romegialli AM, Hirsch M, Berardi L, Ramsey CM, et al. Functional outcomes after hip fracture in independent community-dwelling patients. J Am Geriatr Soc 2019;67(7):1386–92. https://doi.org/10.1111/jgs.15870.

[5] Aarden JJ, van der Esch M, Engelbert RHH, van der Schaaf M, de Rooij SE, Buurman BM. Hip fractures in older patients: trajectories of disability after surgery. J Nutr Health Aging 2017;21(7):837–42. https://doi.org/10.1007/s12603-016-0830-y.

[6] Dakhil S, Saltvedt I, Benth JS, Thingstad P, Watne LO, Bruun Wyller T, et al. Longitudinal trajectories of functional recovery after hip fracture. PLoS One 2023;18(3):e0283551. https://doi.org/10.1371/journal.pone.0283551.

[7] Whitson HE, Duan-Porter W, Schmader KE, Morey MC, Cohen HJ, Colon-Emeric CS. Physcal resilience in older adults: systematic review and development of an emerging construct. J Gerontol A Biol Sci Med Sci 2016;71(4):489–95. https://doi.org/10.1093/gerona/glv202.

[8] Xu J, Wang Y, Muniz-Terrera G, Xue QL, Wu C. Physical resilience after hip fracture: unpacking the roles of resistance and recovery. J Am Geriatr Soc 2025;73(11):3416–24. https://doi.org/10.1111/jgs.70088.

[9] Clegg A, Young J, Iliffe S, Olde Rikkert MGM, Rockwood K. Frailty in older people. Lancet 2013;381(9868):752–62. https://doi.org/10.1016/S0140-6736(12)62167-9.

[10] Xu BY, Yan S, Low LL, Vasanwala FF, Low SG. Predictors of poor functional outcomes and mortality in patients with hip fracture: a systematic review. BMC Musculoskelet Disord 2019;20(1):568. https://doi.org/10.1186/s12891-019-2950-0.

[11] Song Y, Wu Z, Huo H, Zhao P. The impact of frailty on adverse outcomes in geriatric hip fracture patients: a systematic review and meta-analysis. Front Public Health 2022;10:890652. https://doi.org/10.3389/fpubh.2022.890652.

[12] Leibson CL, Tosteson ANA, Gabriel SE, Ransom JE, Melton III LJ. Mortality, disability, and nursing home use for persons with and without hip fracture: a population-based study. J Am Geriatr Soc 2002;50(10):1644–50. https://doi.org/10.1046/j.1532-5415.2002.50455.x.

[13] 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. https://doi.org/10.1093/ije/dys203.

[14] Zhao Y, Strauss J, Yang G, Giles J, Hu PP, Hu Y, et al. China health and retirement longitudinal study: 2011–2012 national baseline user's guide. Beijing, China: National School of Development, Peking University,; 2013. Available from, https://charls.pku.edu.cn/en/data/User2011.pdf.

[15] Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001;56(3):M146–57. https://doi.org/10.1093/gerona/56.3.M146.

[16] Wu C, Smit E, Xue QL, Odden MC. Prevalence and correlates of frailty among community-dwelling Chinese older adults: the China health and retirement longitudinal study. J Gerontol A Biol Sci Med Sci 2018;73(1):102–8. https://doi.org/10.1093/gerona/glx098.

[17] Zhang Q, Zhao X, Liu H, Ding H. Frailty as a predictor of future falls and disability: a four-year follow-up study of Chinese older adults. BMC Geriatr 2020;20(1):388. https://doi.org/10.1186/s12877-020-01798-z.

[18] Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged: The index of ADL: a standardized measure of biological and psychosocial function. JAMA 1963;185(12):914–9. https://doi.org/10.1001/jama.1963.03060120024016.

[19] Lawton MP, Brody EM. Assessment of older people: Self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179–86. https://doi.org/10.1093/geront/9.3_Part_1.179.

[20] Colon-Emeric C, Whitson HE, Pieper CF, Sloane R, Orwig D, Huffman KM, et al. Resiliency groups following hip fracture in older adults. J Am Geriatr Soc 2019;67(12):2519–27. https://doi.org/10.1111/jgs.16152.

[21] Zhong BX, Zhong HL, Zhou GQ, Xu WQ, Lu Y, Zhao Q. Physical performance and risk of hip fracture in community-dwelling older people in China: A 4-year longitudinal cohort study. Maturitas 2021;146:26–33. https://doi.org/10.1016/j.maturitas.2021.01.003.

[22] Magaziner J, Simonsick EM, Kashner TM, Hebel JR, Kenzora JE. Predictors of functional recovery one year following hospital discharge for hip fracture: a prospective study. J Gerontol 1990;45(3):M101–7. https://doi.org/10.1093/geronj/45.3.M101.

[23] Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res 2011;46(3):399–424. https://doi.org/10.1080/00273171.2011.568786.

[24] Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treat-ment weighting using the propensity score to estimate causal treatment effects in observational studies. Stat Med 2015;34(28):3661–79. https://doi.org/10.1002/sim.6607.

[25] Hernan MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. Am J Epidemiol 2016;183(8):758–64. https://doi.org/10.1093/aje/kwv254.

[26] Sun L, Abraham S. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. J Econom 2021;225(2):175–99. https://doi.org/10.1016/j.jeconom.2020.09.006.

[27] Callaway B, Sant'Anna PHC. Difference-in-differences with multiple time periods. J Econom 2021;225(2):200–30. https://doi.org/10.1016/j.jeconom.2020.12.001.

[28] von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007;370(9596): 1453–7. https://doi.org/10.1016/S0140-6736(07)61602-X.

[29] Tseng MY, Shyu YIL, Liang J. Functional recovery of older hip-fracture patients after inter-disciplinary intervention follows three distinct trajectories. Gerontologist 2012;52(6):833–42. https://doi.org/10.1093/geront/gns058.

[30] Xu W, Li YX, Hu Y, Wu C. Association of frailty with recovery from disability among community-dwelling Chinese older adults: China health and retirement longitudinal study. BMC Geriatr 2020;20(1):119. https://doi.org/10.1186/s12877-020-01519-6.

[31] Wu C, Kim DH, Xue QL, Lee DSH, Varadhan R, Odden MC. Association of frailty with recovery from disability among community-dwelling older adults: results from two large U.S. cohorts. J Gerontol A Biol Sci Med Sci 2019;74(4):575–81. https://doi.org/10.1093/gerona/gly080.

[32] Hu H, Deng F, Chen Z, Chen B. Cognitive trajectories before and after geriatric hip fracture: a matched longitudinal analysis of the health and retirement study, 1996–2016. BMC Geriatr 2026;26:1012. https://doi.org/10.1186/s12877-026-07865-1.

[33] Hu H, Shou K. Depressive symptom trajectories around incident hip fracture in older adults: an event-centered matched analysis from the Health and Retirement Study. Maturitas 2026;211:109032. https://doi.org/10.1016/j.maturitas.2026.109032.

[34] Inoue T, Misu S, Tanaka T, Kakehi T, Kakiuchi M, Chuman Y, et al. Frailty defined by 19 items as a predictor of short-term functional recovery in patients with hip fracture. Injury 2019;50(12):2272–6. https://doi.org/10.1016/j.injury.2019.10.011.

[35] Parker SG, McCue P, Phelps K, McCleod A, Arora S, Nockels K, et al. What is comprehensive geriatric assessment (CGA)? An umbrella review. Age Ageing 2018; 47(1):149–55. https://doi.org/10.1093/ageing/afx166.

[36] National Institute for Health, Care Excellence, Hip fracture: management, national institute for health and care excellence, 2011, Published June 22, 2011, updated January 6, 2023, Accessed August 26, 2026, Available from: https://www.nice.org.uk/guidance/cg124.