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.

Pain-distribution transitions and subsequent basic activities of daily living limitation: a repeated-landmark analysis of the China Health and Retirement Longitudinal Study

Zhenwei Pan
Panyu Hospital of Guangzhou University of Chinese Medicine
Minhua Deng
Guangzhou University of Chinese Medicine
Huimin Xiang
Guangzhou University of Chinese Medicine
Hao Li
Panyu Hospital of Guangzhou University of Chinese Medicine
Junjie He
Panyu Hospital of Guangzhou University of Chinese Medicine
Qipeng Wei
Panyu Hospital of Guangzhou University of Chinese Medicine
Xiaofeng Chen
cxf1159@163.com

Panyu Hospital of Guangzhou University of Chinese Medicine

Article

Keywords: Pain distribution, low back pain, multisite pain, activities of daily living, functional limitation, landmark analysis, CHARLS

Posted Date: September 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10495417/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

Pain distribution may provide useful information about later functional decline in older adults. This study examined whether short-term transitions in pain distribution were associated with new basic activities of daily living (BADL) limitation.

Methods

We used China Health and Retirement Longitudinal Study data from 2011 to 2020. The analysis stacked three repeated-landmark windows. Each window required two observed pain states and BADL independence at the landmark. Pain states were pain-free, non-axial pain only, isolated back or waist pain, and axial multisite pain. We fitted weighted modified Poisson models, estimated standardized risks, and ran prespecified sensitivity analyses.

Results

The primary analysis included 26574 person-intervals from 11947 participants and observed 3983 new BADL events. Persistent axial multisite pain had the strongest association with new BADL limitation compared with stable pain-free status (risk ratio [RR] 3.42, 95% confidence interval [CI] 3.06–3.82). The estimate was 1.65 (95% CI 1.45–1.88) after adjustment for current health and early functional measures. The standardized next-wave risks were 30.6% for persistent axial multisite pain and 9.0% for stable pain-free status.

Conclusions

Persistent axial multisite pain was associated with a high subsequent risk of BADL limitation among middle-aged and older adults who were BADL independent at each landmark. The findings support careful functional assessment in adults with persistent axial multisite pain. The study did not estimate causal effects or validate a clinical prediction rule.

Background

Low back pain and other musculoskeletal pain are common in later life. These symptoms often occur with reduced mobility and loss of independence [1, 2]. A single pain question can miss important differences. Pain distribution, persistence, and change may provide more information than one report of back pain.

Recent studies have described pain classes, pain trajectories, disability trajectories, and movement between latent pain states [3–7]. Those studies show that repeated pain measurement is useful. However, a long-term trajectory can hide a change between two consecutive visits. A final outcome analysis can also retain a healthier subset. In addition, a broad “no axial pain” reference can mix pain-free adults with adults who have pain at other sites.

A repeated-landmark design can address part of this problem. The design repeats the same order at several points: a pain transition, confirmation of BADL independence, and measurement of a new BADL limitation at the

next survey. Landmark methods were first developed for time-updated prognosis, although the present study uses discrete survey waves rather than continuous event times [8]. The design does not remove confounding. It does provide a clear time order and a transparent target population.

This study examined eight transitions among four pain states in three CHARLS windows. The study tested two main hypotheses. First, the study expected persistent axial multisite pain to have the strongest association with subsequent BADL limitation. Second, the study expected spread from isolated axial pain to multisite pain to mark higher risk. The study also explored whether a primary-care facility in the 2011 community changed these associations. The study treated every estimate as an association.

Methods

Data Source and Ethics

CHARLS is a national longitudinal study of Chinese adults aged 45 years or older and their spouses. The cohort profile describes the sampling design and the baseline study [9]. We linked individual data from 2011, 2013, 2015, 2018, and 2020. We also linked the 2011 community questionnaire. The CHARLS user guides describe the released wave files and survey procedures [10, 11].

The original CHARLS study received approval from the Peking University Biomedical Ethics Review Committee under IRB00001052-11015. CHARLS obtained informed consent from participants. The materials available for this blinded draft did not include an institution-specific secondary-analysis determination. This report follows the STROBE guidance for observational research [12].

Repeated-Landmark Design and Estimand

We used three windows. Window 1 used pain in 2011 and 2013, confirmed BADL independence in 2013, and measured BADL in 2015. Window 2 used pain in 2013 and 2015, confirmed independence in 2015, and measured BADL in 2018. Window 3 used pain in 2015 and 2018, confirmed independence in 2018, and measured BADL in 2020. The approximate outcome follow-up periods were 2, 3, and 2 years. Figure 1 shows the design.

The primary estimand was the next-wave BADL risk associated with each adjacent-wave pain transition among adults who were BADL independent at the landmark, were not recorded as dead before the next wave, and had an observed next-wave BADL measure. The censoring model addressed nonresponse among alive eligible intervals under its measured-covariate assumptions. The BADL-or-death sensitivity analysis used a different clinical outcome. It did not estimate the same estimand as the primary analysis.

We stacked eligible intervals and included window indicators in each model. A participant could contribute one, two, or three intervals. The primary variance estimator used baseline community clusters. Every participant retained the same baseline community identifier, so this cluster also grouped the participant’s repeated intervals. A participant-clustered analysis tested the sensitivity of the standard errors.

Eligibility Criteria

The baseline source contained 17,705 records. We retained 17,289 adults who were aged 45 years or older and had a valid 2011 individual weight. Each interval required an observed pain state at the start and landmark waves. Each interval also required a landmark BADL assessment that showed no limitation. We excluded a confirmed death before the outcome wave and then required an observed next-wave BADL assessment. We excluded logically contradictory pain reports. We retained proxy interviews in the primary cohort and excluded them in a sensitivity analysis.

Pain States and Transitions

We used the general pain item and up to 15 harmonized pain sites. The 2011 and 2015 questionnaires used a binary pain screen. The 2013 and 2018 questionnaires used ordered pain responses. The primary definition treated “a little” pain or any higher category as pain. A stricter sensitivity analysis used at least “some” or “somewhat” pain, according to the wave wording. We ignored lower-severity recorded sites when a response fell below the stricter threshold.

We defined axial pain as back and/or waist pain. We created four states at each wave: pain-free, non-axial pain only, isolated axial pain, and axial multisite pain. The isolated state required back or waist pain without another painful site. The axial multisite state required back or waist pain plus at least one other painful site.

We mapped every adjacent pair into one of eight transitions. The transitions were stable pain-free; non-axial/no-axial pattern; incident isolated axial pain; persistent isolated axial pain; spread from isolated axial pain to axial multisite pain; incident axial multisite pain; persistent axial multisite pain; and axial-pain improvement or contraction. Supplementary Table S2 gives the complete 4 × 4 map. Supplementary Table S3 gives all 16 observed cells and their events.

Outcomes

The primary outcome was new next-wave BADL limitation. The BADL items covered dressing, bathing, eating, getting into or out of bed, toileting, and controlling urination and defecation. We classified a participant as limited when the participant reported difficulty or a need for help in at least one item. We treated the documented survey skip as no difficulty when all required preceding function items indicated no difficulty.

The sensitivity outcomes were severe BADL limitation, BADL limitation or death, and new next-wave IADL limitation. The severe BADL outcome required a need for help or inability in at least one item. The IADL sensitivity analysis formed a separate endpoint-specific risk set. It required IADL independence at the landmark and then used the five harmonized instrumental activity items at the next wave. The observed IADL analysis count was 25,403 intervals. This count was 1,171 intervals lower than the primary BADL analysis count of 26,574. The two counts differed because the IADL analysis used endpoint-specific entry and response rules.

Covariates

Model 0 included pain transition and window. Model 1 was the primary model. Model 1 added landmark age, sex, education, marital status, rural residence, agricultural hukou, and 2011 body mass index (BMI). Model 2 added cumulative diagnosed chronic-condition count, arthritis, stroke, 10-item Center for Epidemiologic Studies

Depression score, sleep duration, a three-level smoking category, drinking category, fair or poor self-rated health, mobility difficulty count, and landmark IADL difficulty count.

Model 1 estimated a core demographic and socioeconomic adjusted association. Model 2 estimated an association conditional on current health and early functional status. Several Model 2 variables may be intermediate features or early signs of functional decline. Therefore, the two models answer different prognostic questions. Model 2 does not isolate a controlled direct effect.

Baseline Community Primary Care

The exploratory community exposure was the presence of at least one primary-care facility in the 2011 community questionnaire. The facility list included a community health center, a community medical post, a township health clinic or hospital, or a village medical post. We removed repeated community records before linkage.

The exposure described facility presence in the baseline sampling community. The data did not reliably identify later moves. The exposure did not measure quality, staffing, cost, continuity, pain care, or actual use. We did not use travel time as the main modifier because the variable had missing and incomplete reports.

Multiple Imputation

We created 30 completed data sets among 28,403 alive eligible person-intervals. This set included 26,574 intervals with an observed next-wave BADL outcome and 1,829 intervals used only in the response model. The random seeds ranged from 20260718 through 20260747. Each stochastic chain used 15 iterations and posterior sampling. We imputed covariates only. We did not impute pain transition, BADL outcome, death, participant identifier, community identifier, or survey weight.

The imputation matrix included the Model 1 and Model 2 covariates. The matrix also included pain-transition indicators, window indicators, next-wave response, baseline primary-care presence, and the log of the baseline survey weight. The imputation procedure used a continuous conditional model for every incomplete column. It then rounded and bounded categorical values to their prespecified legal codes. It bounded continuous and count variables to their valid ranges. The procedure imputed continuous BMI and bounded it from 10 to 60 kg/m² before analysis. It did not fit separate logistic or multinomial conditional models for categorical variables. Variable-specific models can better preserve category probabilities, so we retained this issue as a limitation [14]. Supplementary Tables S4 and S5 report missingness with their respective denominators and completed-data checks. A complete-case sensitivity analysis examined whether the primary contrast changed when all model covariates were observed.

We pooled coefficients and covariance matrices with Rubin’s rules [13]. The total covariance was

$$T = \bar{U} + (1 + 1/m)B,$$

where $\bar{U}$ was the mean within-imputation covariance, B was the between-imputation covariance, and m equaled 30.

Attrition Weights

We estimated a response model separately in each window among landmark-independent adults who were not recorded as dead before the outcome wave. The numerator was the 2011-survey-weighted response probability in that window. The denominator came from a survey-weighted logistic model. The denominator model included pain transition and every Model 1 and Model 2 covariate.

For participant i in window w, the stabilized censoring weight was

$SW_iw = P(R_iw = 1 \mid W = w) / P(R_iw = 1 \mid A_iw, X_iw, W = w)$.

The program limited the fitted denominator probability to 0.02 through 0.98. The program estimated the response model separately in each imputed data set. The program normalized the 2011 individual survey weight to a mean of one within each window. The final analysis weight was

$CW_iw = SW_iw \times S_i / Sw$.

The primary analysis did not truncate SW. Two sensitivity analyses truncated SW at the 1st and 99th percentiles and at the 0.5th and 99.5th percentiles. This approach follows the general construction of inverse-probability weights [15].

Regression and Standardization

We estimated risk ratios with log-link modified-Poisson regression [16]. The primary models used robust standard errors clustered by baseline community [17]. We used stable pain-free status as the reference. A seven-degree-of-freedom Wald test assessed the transition factor. A 14-degree-of-freedom test assessed the transition-by-window interaction. We used two-sided tests and 95% CIs. We presented exact P values when useful and used P < .001 for smaller values.

We calculated Model 1 standardized risks. Within each completed data set, the program set every interval to each transition in turn and retained that interval's covariates. The program averaged predicted risks with the combined weight. The delta method estimated within-imputation variance. Rubin's rules then pooled each risk and risk difference. When a probability-scale risk interval crossed 0 or 1, we transformed the pooled risk and its standard error to the logit scale and back-transformed the limits. The pooled risks average predictions across 2-, 3-, and 2-year next-wave intervals. They are not fixed two-year risks.

The prespecified direct contrasts compared persistent with incident axial multisite pain and spread with persistent isolated axial pain. A secondary isolated-start model compared spread, persistent isolated pain, and recovery or contraction. A second start-state model separated the landmark states among intervals that began with axial multisite pain. These models helped explain the broad improvement/contraction category.

Exploratory Community Interaction

The community analysis combined spread and persistent axial multisite pain as the high-risk pattern. The model used three pain groups: stable pain-free, other transitions, and high-risk pattern. Model 1 added primary-care presence and product terms. We reported the ratio of high-risk-pattern RRs across facility strata. We also reported the additive interaction as R11 − R10 − R01 + R00 [18]. We repeated this test in Window 1 because baseline community context was closest to the exposure time in that window.

Sensitivity and Diagnostic Analyses

The sensitivity analyses removed all analysis weights, used the baseline survey weight only, removed proxy interviews, removed Window 3, used severe BADL, used BADL or death, used the stricter pain threshold, used IADL, and fit each window separately. Other analyses used participant-clustered standard errors and two IPCW truncation rules.

We also resampled 2011 communities 500 times. This bootstrap used the first completed data set and fixed combined weights. Therefore, the bootstrap tested cluster sensitivity only. It did not propagate all multiple-imputation uncertainty. The primary Rubin-pooled analysis remained the main inference. We used Python 3.12.13, pandas 2.2.3, NumPy 2.3.5, SciPy 1.17.0, and scikit-learn 1.8.0.

Results

Sample and Descriptive Results

The stacked data included 26,574 person-intervals from 11,947 participants in 445 baseline communities. We observed 3,983 new BADL events. Window 1 included 9,921 intervals and 1,489 events. Window 2 included 8,273 intervals and 1,064 events. Window 3 included 8,380 intervals and 1,430 events.

The landmark-independent risk sets contained 10,923, 9,136, and 9,031 intervals before the death exclusion. The data recorded 201, 285, and 201 deaths before the next outcome wave. Among the remaining alive intervals, 801, 578, and 450 lacked an observed next-wave BADL measure. Supplementary Table S1 presents the full window flow.

The transition groups were uneven. Persistent isolated axial pain included 90 intervals, and spread included 193 intervals. Persistent axial multisite pain included 1,630 intervals. The largest group was stable pain-free status with 13,479 intervals. Table 1 reports descriptive results.

Table 1
Characteristics of person-intervals by pain transition

Transition	N	Events	Weighted crude risk, %	Age, y	Women, %	Chronic count	CES-D	Mobility count
Stable pain-free	13479	1202	8.6	61.6	43.4	1.42	5.53	1.02
Non-axial/no-axial	4607	781	16.7	62.2	55.6	1.97	8.13	1.85
Incident isolated	731	125	16.1	61.5	49.2	1.55	7.45	1.60
Persistent isolated	90	15	15.9	60.9	48.8	1.67	8.22	1.93
Spread to multisite	193	46	23.3	62.4	57.6	2.25	10.15	2.56
Incident multisite	3094	647	20.5	62.2	59.1	2.48	9.90	2.37
Persistent multisite	1630	562	33.5	62.2	69.5	3.19	12.15	3.05
Improvement/contraction	2750	605	22.0	61.0	59.6	2.32	8.51	1.97

Counts are unweighted. Percentages and means use the combined baseline-survey and censoring weight and average across 30 completed data sets.

Of the 17,289 baseline-eligible adults, 5,342 contributed no primary-analysis interval. These adults differed from contributors on age, residence, hukou, BMI missingness, and pain-state observability. Supplementary Table S10 reports the comparison. These differences show that entry into the analytic risk sets was not random.

Pain Transitions and New BADL Limitation

The transition factor was associated with new BADL limitation in Model 1 (Wald $\chi^2 = 485.12$; 7 df; $P < .001$). Persistent axial multisite pain had the largest Model 1 estimate against stable pain-free status (RR, 3.42; 95% CI, 3.06–3.82). Spread had an RR of 2.53 (95% CI, 1.92–3.32). Incident axial multisite pain had an RR of 2.23 (95% CI, 1.95–2.55). Improvement or contraction had an RR of 2.31 (95% CI, 2.03–2.63). Table 2 presents every model.

Table 2
Risk ratios for new next-wave BADL limitation

Transition	Model 0	Model 1, primary	Model 2
Stable pain-free	1.00 (reference)	1.00 (reference)	1.00 (reference)
Non-axial/no-axial	1.94 (1.73–2.16)	1.78 (1.60–1.99)	1.32 (1.18–1.48)
Incident isolated	1.82 (1.46–2.28)	1.76 (1.45–2.13)	1.39 (1.16–1.67)
Persistent isolated	1.84 (1.15–2.96)	1.82 (1.18–2.80)	1.35 (0.90–2.02)
Spread to multisite	2.78 (2.10–3.68)	2.53 (1.92–3.32)	1.47 (1.12–1.92)
Incident multisite	2.38 (2.07–2.73)	2.23 (1.95–2.55)	1.36 (1.19–1.56)
Persistent multisite	3.87 (3.44–4.35)	3.42 (3.06–3.82)	1.65 (1.45–1.88)
Improvement/contraction	2.54 (2.22–2.90)	2.31 (2.03–2.63)	1.61 (1.42–1.83)

Model 0 includes window. Model 1 adds the core demographic and socioeconomic covariates. Model 2 adds current health and early functional measures. The models use 30 imputations, the combined weight, and community-clustered robust standard errors.

Model 2 attenuated the associations. Persistent axial multisite pain still had the largest estimate (RR, 1.65; 95% CI, 1.45–1.88). The attenuation shows that pain transitions shared prognostic information with current health and early function. It does not show that these variables mediated a causal effect.

Persistent axial multisite pain had a higher risk than incident axial multisite pain (RR, 1.53; 95% CI, 1.34–1.75; P < .001). Spread had a higher point estimate than persistent isolated pain, but the direct comparison was imprecise (RR, 1.39; 95% CI, 0.83–2.32; P = .209). In the isolated-start analysis, spread also had an imprecise comparison with recovery or contraction (RR, 1.27; 95% CI, 0.85–1.88; P = .243). Persistent isolated pain had an RR of 0.90 (95% CI, 0.56–1.46; P = .672) against recovery or contraction. These results did not establish a risk ordering among the small isolated-start paths.

Among intervals that began with axial multisite pain, the state at the landmark separated true pain-free recovery from other contractions. Compared with becoming pain-free, a landmark non-axial state had an RR of 1.48 (95% CI, 1.22–1.78), persistent axial multisite pain had an RR of 1.68 (95% CI, 1.40–2.00), and an isolated axial state had an RR of 1.19 (95% CI, 0.90–1.59). Supplementary Table S9 gives the full start-state results.

Standardized Risks and Window Results

The pooled standardized next-wave risk was 9.0% (95% CI, 8.2%–9.7%) for stable pain-free status. The risk was 30.6% (95% CI, 27.9%–33.2%) for persistent axial multisite pain. The risk difference was 21.6 percentage points (95% CI, 19.0–24.3). Spread had a standardized risk of 22.6% (95% CI, 16.7%–28.6%). Table 3 and Fig. 2 present all categories.

Table 3
Model 1 standardized next-wave risks

Transition	Standardized risk, % (95% CI)	Risk difference, percentage points (95% CI)
Stable pain-free	9.0 (8.2–9.7)	0.0 (reference)
Non-axial/no-axial	16.0 (14.6–17.3)	7.0 (5.6–8.5)
Incident isolated	15.7 (12.9–18.6)	6.8 (3.9–9.7)
Persistent isolated	16.3 (9.3–23.3)	7.4 (0.4–14.3)
Spread to multisite	22.6 (16.7–28.6)	13.7 (7.7–19.7)
Incident multisite	20.0 (17.8–22.2)	11.0 (8.7–13.3)
Persistent multisite	30.6 (27.9–33.2)	21.6 (19.0–24.3)
Improvement/contraction	20.7 (18.8–22.6)	11.8 (9.6–13.9)
The pooled risks average the 2-, 3-, and 2-year windows. They are not fixed two-year risks.		

Table 4
Direct contrasts and secondary analyses

Comparison	Analysis	Estimate (95% CI)	P value
Persistent multisite vs stable pain-free	Primary Model 1	3.42 (3.06–3.82)	< .001
Persistent vs incident multisite	Direct contrast	1.53 (1.34–1.75)	< .001
Spread vs persistent isolated	Direct contrast	1.39 (0.83–2.32)	.209
Spread vs recovery/contraction	Isolated-start contrast	1.27 (0.85–1.88)	.243
High-risk pattern × PHC	Ratio of RRs	0.92 (0.72–1.17)	.487
High-risk pattern × PHC	Additive interaction	−3.1 pp (− 8.5 to 2.4)	.271
Persistent multisite vs stable pain-free	New next-wave IADL limitation	2.75 (2.48–3.04)	< .001
PHC means baseline community primary health care. pp means percentage points.			

The window-specific RRs for persistent axial multisite pain were 3.54 (95% CI, 2.93–4.27), 2.97 (95% CI, 2.35–3.76), and 3.62 (95% CI, 3.07–4.28). The transition-by-window test did not show clear heterogeneity ($\chi^2 = 21.49$; 14 df; P = .090). The window-specific standardized risks for persistent axial multisite pain were 31.3% (95% CI, 26.1%–36.5%), 25.8% (95% CI, 20.9%–30.7%), and 33.7% (95% CI, 30.2%–37.1%). Supplementary Table S8 reports the window-specific absolute risks and risk differences.

Exploratory Community Results

Baseline primary-care presence did not show clear modification. The pooled high-risk-pattern ratio of RRs was 0.92 (95% CI, 0.72–1.17; P = .487). The additive interaction was − 3.1 percentage points (95% CI, − 8.5 to 2.4; P = .271). The Window 1 ratio of RRs was 0.79 (95% CI, 0.54–1.17; P = .243). The pooled estimate used all three windows, while the Window 1 estimate used only the 2011–2013–2015 interval. The estimates therefore answer

different timing questions and are not two versions of the same model. Figure 3 and Supplementary Table S11 show the standardized cell risks. The intervals remained wide in the smaller facility-absent cells.

Sensitivity Results

The main contrast remained positive in all planned sensitivity analyses. The estimates were 3.43 without weights, 3.41 with baseline weights only, 3.73 after proxy interviews were removed, and 3.25 after Window 3 was removed. The estimates were 2.76 for severe BADL, 2.93 for BADL or death, and 3.11 with the stricter pain threshold. The new next-wave IADL sensitivity analysis included 25,403 observed intervals and 4,147 events after endpoint-specific IADL eligibility and response rules. Its RR was 2.75 (95% CI, 2.48–3.04). Participant-clustered standard errors gave an RR of 3.42 (95% CI, 3.07–3.81). Both IPCW truncation rules gave an RR of 3.42. The 500-replicate community bootstrap gave a median RR of 3.41 and a percentile interval of 3.04–3.82. The complete-case analysis included 20,403 responding intervals and gave an RR of 3.32 (95% CI, 2.92–3.77). Figure 4 and Supplementary Table S7 report these analyses.

Discussion

Main Findings

This repeated-landmark study found that persistent axial multisite pain was associated with the highest subsequent BADL risk among adults who were still BADL independent. The result was strong on both relative and absolute scales. Persistent axial multisite pain also had a higher risk than incident axial multisite pain. These findings support the persistence hypothesis.

The evidence for spread was weaker. Spread had a clear association against stable pain-free status. However, the spread group contained only 193 intervals. Direct comparisons with persistent isolated pain and recovery or contraction had wide intervals that included no difference. Therefore, the study cannot claim that spread has a worse risk order than every other isolated-axial path.

The broad improvement/contraction group remained at higher risk than stable pain-free status. The start-state analysis showed why a simple “improved” label can mislead. Some participants became pain-free, while others retained non-axial or isolated axial pain. A report without axial pain did not always mean full pain recovery.

Relation to Earlier Work

Earlier studies linked multisite pain with later functional problems and disability paths [2–4]. Other studies used latent classes or long-term trajectories [5–7]. The present study adds a different view. It uses an adjacent-wave transition that clinicians and readers can interpret directly. It also renews the BADL-independent risk set at each landmark.

This design does not replace latent trajectory models. A long-term class describes an overall pattern, while a repeated landmark describes a short sequence tied to the next outcome. The two methods answer different questions. The present results support persistent multisite pain as a strong association marker. The results do not validate a clinical prediction tool.

Interpretation

Persistent axial multisite pain may co-occur with greater pain burden, multimorbidity, depressed mood, sleep problems, reduced activity, and early mobility loss. Model 2 included many of these measures. Its lower estimates show that these measures carried overlapping risk information. The analysis cannot separate confounding, mediation, measurement overlap, and reverse influence from early decline.

The pooled absolute risks require care. The three windows lasted about 2, 3, and 2 years. The pooled 30.6% risk is an average next-wave risk over those unequal periods. It is not a fixed two-year risk. The window-specific estimates and the interaction test provide the time-scale check.

Clinical and Community Meaning

The results identify a subgroup with elevated subsequent BADL risk in this association analysis. Clinicians may consider pain distribution together with current mobility, mood, sleep, chronic disease, and IADL status. However, the study did not test whether pain transitions improved discrimination, calibration, Brier score, or net benefit. The study also did not show that changing pain would prevent BADL limitation.

The community analysis was exploratory. The data did not show clear evidence that baseline facility presence changed the association. This result does not show that primary care is ineffective. Facility presence was a coarse measure. It did not measure care quality, staffing, access, use, or pain treatment. It also did not track residential moves.

Strengths

The study used a large national cohort and five individual survey waves. The design separated truly pain-free adults from adults with non-axial pain. It also used clear time ordering and renewed BADL-independent risk sets. We reported RRs, standardized risks, and risk differences. We disclosed all 16 state cells and used start-state models to explain the broad contraction group. We also tested participant clustering, community resampling, weight truncation, alternative outcomes, a stricter pain threshold, and death in a composite outcome.

Limitations

The study relied on self-reported pain and function. The pain questions differed across waves. The 2011 and 2015 screens were binary, while the 2013 and 2018 screens were ordered. The stricter threshold analysis reduced but did not remove the risk of wave-related state reclassification. Supplementary Table S12 shows that the threshold changed several transition counts.

The study conditioned on BADL independence at each landmark. This choice defined the target population, but it removed adults who had already developed BADL limitation. The primary analysis also described alive participants with an observed next-wave outcome. The BADL-or-death analysis used a broader clinical outcome. A Fine–Gray model was not used because BADL onset time was not observed between survey waves. The available data did not support continuous-time precision.

Selection remained possible. A total of 5,342 baseline-eligible adults contributed no interval. The IPCW model addressed next-wave nonresponse among alive eligible intervals. It did not repair missing pain states before a risk set was formed. The contributor comparison showed meaningful differences. The findings may therefore understate or overstate associations in less observed and less healthy adults.

The multiple-imputation program used continuous conditional models and post-imputation rounding for categorical variables. This method may not preserve category probabilities as well as variable-specific logistic or multinomial models. BMI was missing in 16.5% of the 26,574 outcome-observed intervals and 17.3% of the 28,403 alive eligible intervals. The complete-case analysis provides one check, but it cannot remove this limitation.

The primary standard errors used baseline community clusters. The participant-clustered sensitivity was similar. The study did not reproduce every original survey stratum and primary sampling unit. We used the 2011 individual weight and the baseline community identifier. Therefore, the results should not be read as a full reproduction of every CHARLS complex-survey feature.

The community bootstrap used one completed data set and fixed weights. It did not include all imputation uncertainty. The bootstrap served as a diagnostic only. The persistent isolated and spread groups were small. The direct comparisons for these paths had limited precision. Residual confounding and reverse influence from early functional decline also remained possible.

Conclusions

Persistent axial multisite pain was associated with a high risk of new next-wave BADL limitation among BADL-independent middle-aged and older Chinese adults. The association remained across many checks. The evidence for an adverse ordering of spread was not conclusive. Baseline community primary-care presence did not show clear modification in the exploratory analysis. Future studies should test external prediction performance and interventions before using these transitions as a clinical decision rule.

Abbreviations

BADL
basic activities of daily living
BMI
body mass index
CES-D
Center for Epidemiologic Studies Depression Scale
CHARLS
China Health and Retirement Longitudinal Study
CI
confidence interval
IADL
instrumental activities of daily living
IPCW
inverse-probability-of-censoring weight
PHC
primary health care
RR
risk ratio
STROBE

Strengthening the Reporting of Observational Studies in Epidemiology.

Declarations

Ethics approval and consent to participate

The original CHARLS study was approved by the Peking University Biomedical Ethics Review Committee (IRB00001052-11015). CHARLS obtained informed consent from participants.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by the Panyu District Science and Technology Plan Medical and Health Project (2025-Z04-72), the Traditional Chinese Medicine Research Project of Guangdong Provincial Bureau of Traditional Chinese Medicine (20261311), and the Guangzhou Third-level Famous Traditional Chinese Medicine Studio Construction Project (Professor Feng Huang).

Author Contribution

ZWP and XFC conceived the study. ZWP, MHD, HMX, HL, JJH and QPW collected and checked the data. ZWP performed the statistical analysis and drafted the manuscript. XFC supervised the study and revised the manuscript critically. All authors read and approved the final manuscript.

Acknowledgements

The authors thank the CHARLS study team and all participants.

Data Availability

The datasets analysed during the current study are available in the China Health and Retirement Longitudinal Study (CHARLS) repository, https://charls.charlsdata.com/. Access to CHARLS data requires user registration and compliance with the CHARLS data-use agreement. The authors did not generate new participant-level datasets during this study.

References

1. Hartvigsen, J. et al. What low back pain is and why we need to pay attention. *Lancet* **391**, 2356–2367. 10.1016/S0140-6736(18)30480-X (2018).
2. Butera, K. A., Fox, E. J. & George, S. Z. The impact of multisite pain on functional outcomes in older adults. *J. Pain Res.* **12**, 1115–1125. 10.2147/JPR.S192755 (2019).

3. Zhu, C. et al. Patterns of pain in middle-aged and older adults in China: a latent class analysis. *Pain Ther.* **14**, 1809–1831. 10.1007/s40122-025-00775-x (2025).
4. Liu, P. et al. Association between multisite musculoskeletal pain and disability trajectories among community-dwelling older adults. *Aging Clin. Exp. Res.* **36**, 115. 10.1007/s40520-024-02764-0 (2024).
5. Xiang, Q. et al. Pain trajectories in relation to incident functional limitation among older adults: a prospective cohort study. *J. Nutr. Health Aging*. **29**, 100704. 10.1016/j.jnha.2025.100704 (2025).
6. Cheng, F. et al. Longitudinal associations of chronic pain severity trajectories and number of pain site trajectories with risk of limitations in ability in daily activities: evidence from two 10-year prospective cohort studies. *J. Nutr. Health Aging*. **27**, 1003–1013. 10.1007/s12603-023-2008-x (2023).
7. Liu, J. et al. Latent pain class identification, longitudinal transitions, and machine learning prediction of incident low back pain in middle-aged and older Chinese adults. *BMC Med. Inf. Decis. Mak.* **26**, 107. 10.1186/s12911-026-03460-x (2026).
8. van Houwelingen, H. C. Dynamic prediction by landmarking in event history analysis. *Scand. J. Stat.* **34**, 70–85. 10.1111/j.1467-9469.2006.00529.x (2007).
9. Zhao, Y., Hu, Y., Smith, J. P., Strauss, J. & Yang, G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). *Int. J. Epidemiol.* **43**, 61–68. 10.1093/ije/dys203 (2014).
10. China Health and Retirement Longitudinal Study. CHARLS 2018 user's guide. Peking University. https://charls.pku.edu.cn/en/data/User2018.pdf. Accessed 18 Jul 2026.
11. China Health and Retirement Longitudinal Study. CHARLS 2020 user guide. Peking University. (2023). https://charls.charlsdata.com/Public/ashelf/public/uploads/document/2020-charls-wave5/application/CHARLS_2020_User_Guide_English.pdf. Accessed 18 Jul 2026.
12. von Elm, E. et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement. *PLoS Med.* **4**, e296. 10.1371/journal.pmed.0040296 (2007).
13. Rubin, D. B. *Multiple imputation for nonresponse in surveys* (Wiley, 1987). 10.1002/9780470316696
14. White, I. R., Royston, P. & Wood, A. M. Multiple imputation using chained equations: issues and guidance for practice. *Stat. Med.* **30**, 377–399. 10.1002/sim.4067 (2011).
15. Cole, S. R. & Hernan, M. A. Constructing inverse probability weights for marginal structural models. *Am. J. Epidemiol.* **168**, 656–664. 10.1093/aje/kwn164 (2008).
16. Zou, G. A modified Poisson regression approach to prospective studies with binary data. *Am. J. Epidemiol.* **159**, 702–706. 10.1093/aje/kwh090 (2004).
17. MacKinnon, J. G., Nielsen, M. O. & Webb, M. D. Cluster-robust inference: a guide to empirical practice. *J. Econom*. **232**, 272–299. 10.1016/j.jeconom.2022.04.001 (2023).
18. Knol, M. J. & VanderWeele, T. J. Recommendations for presenting analyses of effect modification and interaction. *Int. J. Epidemiol.* **41**, 514–520. 10.1093/ije/dyr218 (2012).

Figures

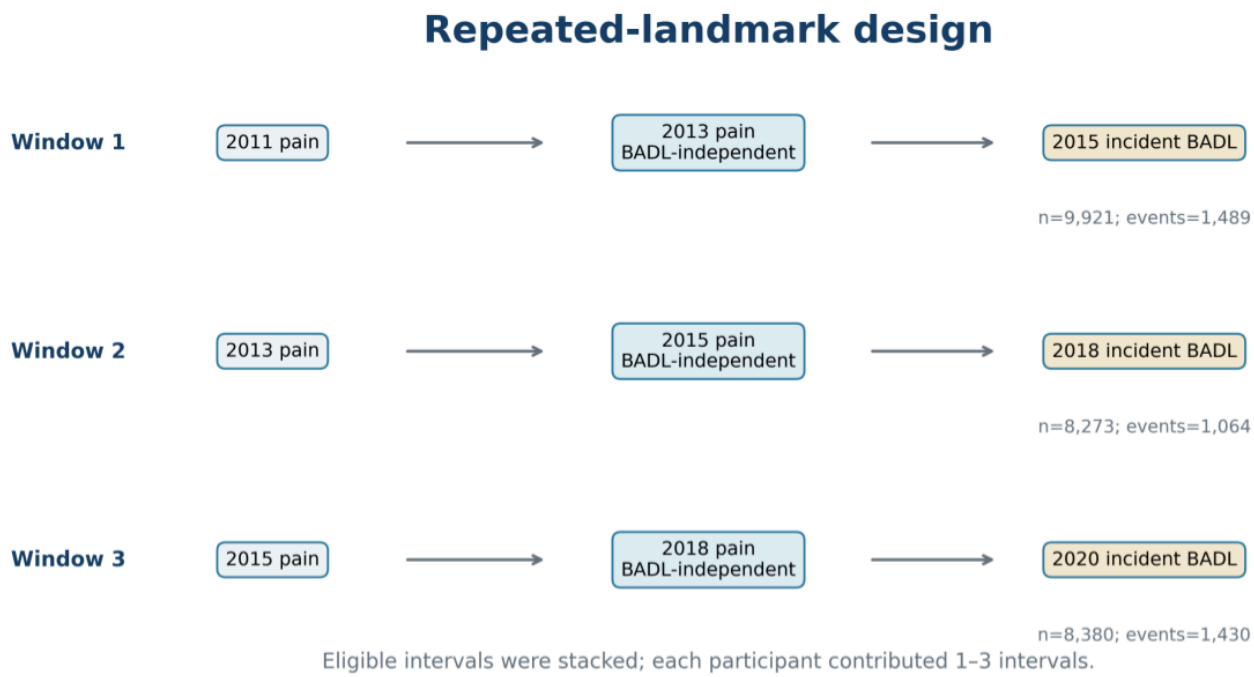


Figure 1

Repeated-landmark design. The figure shows the two pain assessments, landmark BADL independence, next-wave outcome, and sample size for each window.

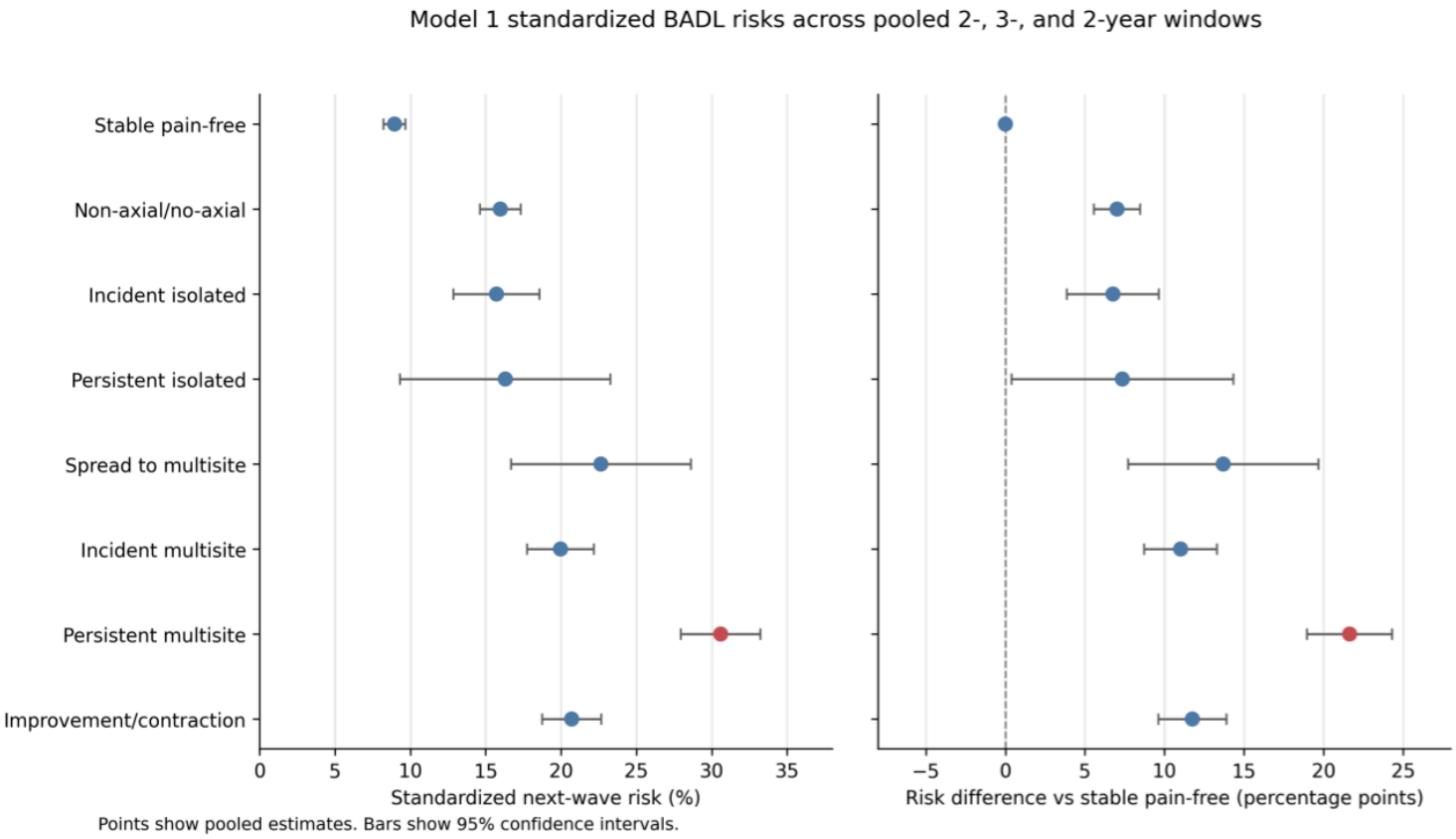


Figure 2

Standardized BADL risk by pain transition

Model 1 standardized risk of subsequent BADL limitation by pain transition. The points show pooled standardized next-wave risks. The bars show 95% CIs. The pooled scale averages the 2-, 3-, and 2-year windows.

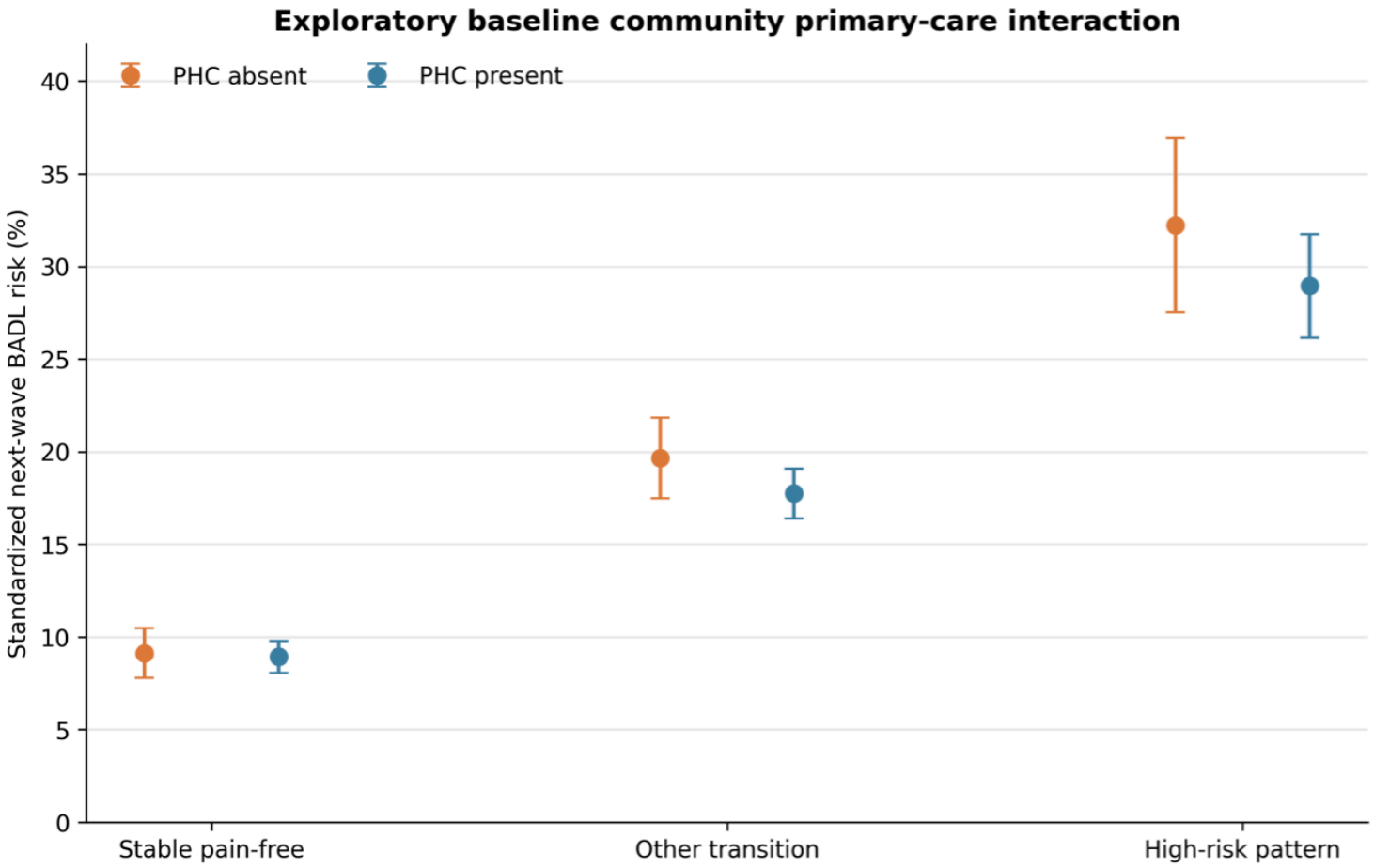


Figure 3

Standardized BADL risk by pain group and baseline PHC

Exploratory standardized BADL risk by pain group and baseline primary-care facility presence. The high-risk group combines spread and persistent axial multisite pain. The bars show 95% CIs. PHC means primary health care.

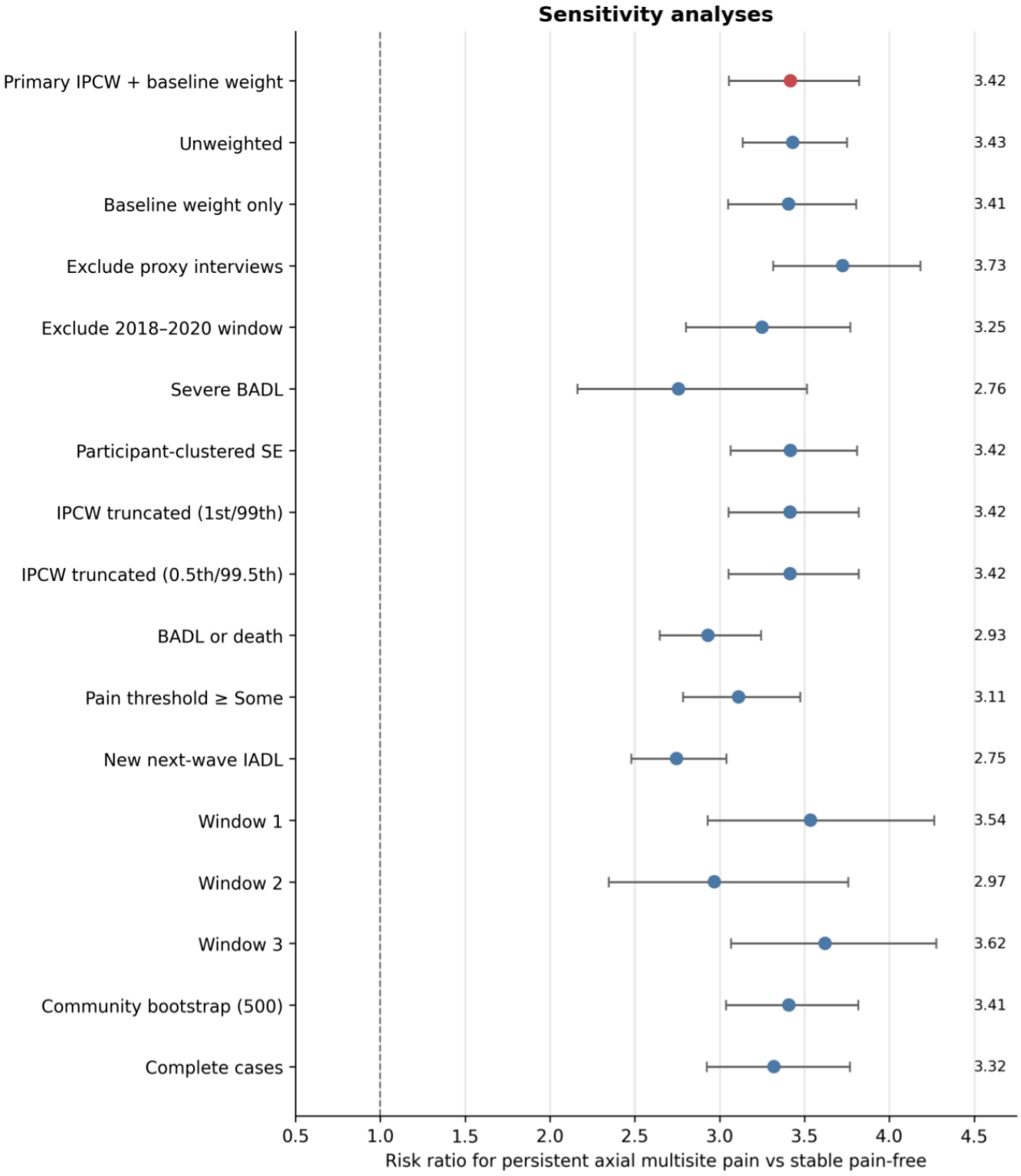


Figure 4

Sensitivity analyses of the primary contrast

Sensitivity analyses for persistent axial multisite pain versus stable pain-free status. The points show RRs. The bars show 95% CIs. The community bootstrap uses percentile limits from 500 resamples of baseline communities. The IADL entry uses a separate endpoint-specific risk set. It required landmark IADL independence and represents new next-wave IADL limitation.

Supplementary Files

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

- OnlineSupplement.docx