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.

Sedentary Accumulation in Midlife and Subsequent SF-36 Physical Functioning and Pain Burden: A Longitudinal Analysis of the 1970 British Cohort Study

Renzheng Zuo
Macao Polytechnic University
Yidan Wang
Nanjing Sport Institute
Deng Wang
Guangzhou Sport University
Tian Hao
Nanjing Sport Institute
Jinlei Nie
Macao Polytechnic University
Qingde Shi
qdshi@mpu.edu.mo

Macao Polytechnic University

Research Article

Keywords: accelerometry, activPAL, sitting patterns, inverse-probability weighting, multiple imputation

Posted Date: September 8th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10742975/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: Sedentary time can be accumulated through many short periods or fewer prolonged bouts, but longitudinal evidence linking these accumulation patterns to subsequent physical functioning and pain in midlife remains limited. We examined whether a greater proportion of waking sitting/lying time accumulated in bouts lasting at least 30 minutes was associated with subsequent physical functioning and pain burden.

Methods: We conducted a longitudinal observational analysis of the 1970 British Cohort Study, linking thigh-worn activPAL data collected at age 46 with health outcomes at age 51. The exposure was the proportion of total waking sitting/lying time accumulated in bouts lasting at least 30 minutes. Outcomes were SF-36 Physical Functioning and pain burden, defined as reverse-scored SF-36 Bodily Pain. Weighted linear regression models incorporated 50 multiple imputations, outcome-specific sequential selection weights, baseline-outcome adjustment, and measured sociodemographic, health, and movement-related covariates. Sensitivity analyses examined alternative adjustment and weighting specifications, outcome bounds, and nonlinear exposure-response relationships.

Results: Approximately 3,820 participants contributed to each outcome analysis. The median long-bout accumulation proportion was 49.15% (interquartile range, 39.93%–58.21%). Per 10-percentage-point higher long-bout proportion, the fully adjusted mean difference was 0.426 points for Physical Functioning (95% confidence interval [CI], −0.096 to 0.948) and −0.099 points for pain burden (95% CI, −0.658 to 0.461). Models accounting for bounded outcome distributions produced similar estimates, and restricted cubic spline analyses showed no clear evidence of nonlinearity. Across comparable sensitivity analyses, estimates ranged from 0.138 to 0.504 points for Physical Functioning and from −0.158 to 0.079 points for pain burden.

Conclusions: In this midlife cohort, a greater proportion of sitting/lying time accumulated in bouts lasting at least 30 minutes was not clearly associated with subsequent SF-36 Physical Functioning or pain burden after accounting for total sitting/lying time and measured covariates. Findings were broadly consistent across alternative model specifications and did not indicate a clear nonlinear association. These results do not establish causality or equivalence.

Introduction

Sedentary behavior is defined as waking behavior performed while sitting, reclining, or lying with low energy expenditure. It is conceptually distinct from insufficient moderate-to-vigorous physical activity[1]. Public health guidance recommends limiting sedentary time and replacing it with physical activity of any intensity[2]. Yet sedentary time is not a single exposure. The same daily volume can be accumulated through many short periods or fewer prolonged, uninterrupted bouts. In a prospective US cohort, both total sedentary time and longer sedentary accumulation were associated with mortality[3]. These findings motivated investigation of accumulation patterns as potential health markers beyond total sedentary volume.

Whether those patterns are associated with later physical functioning is less certain. Among community-dwelling adults aged 65 years or older, fewer very prolonged sedentary bouts were cross-sectionally associated with better performance on selected physical function tests[4]. A cross-sectional BCS70 analysis also related temporal patterns of upright and stepping events to physical function in midlife[5]. In EPIC-Norfolk, movement behaviors were prospectively associated with subsequent walking and chair-stand performance over approximately 6 years[6]. Baseline physical function was also associated with later movement behaviors. Differences in populations, devices, thresholds, and outcomes limit direct comparison with self-reported function during midlife.

Evidence concerning pain is more heterogeneous. In CARDIA, device-estimated sedentary time and activity showed bidirectional associations with knee symptoms in midlife[7]. Several associations attenuated when participants with comorbidities were excluded. A systematic review identified some cross-sectional associations between sedentary behavior and site-specific musculoskeletal pain[8]. However, prospective evidence was insufficient and varied by behavioral domain and pain site. Pain may limit movement, and sedentary contexts differ in their mechanical demands and behavioral setting. Total sedentary time may therefore not capture how sedentary exposure is accumulated.

Taken together, these studies have not directly examined whether sedentary accumulation patterns are prospectively associated with subsequent self-reported physical functioning and pain during midlife. Studies of physical function have generally involved older adults, performance-based outcomes, movement-count devices, or cross-sectional analyses[4–6]. Pain studies have focused mainly on site-specific symptoms and provide limited prospective evidence[7, 8]. It therefore remains unclear whether, at fixed total waking sitting/lying time, a higher long-bout accumulation proportion is associated with subsequent self-reported physical functioning and pain during midlife.

The 1970 British Cohort Study (BCS70) provides a national birth-cohort setting for addressing this question. Thigh-worn accelerometry was deployed at age 46, and health was reassessed approximately 5 years later[9, 10]. We examined whether a higher long-bout accumulation proportion was associated with lower age-51 Physical Functioning, the primary outcome, or higher pain burden, the key secondary outcome. At fixed total waking sitting/lying time, a 10-percentage-point higher proportion represents reallocating 10% of sitting/lying time from bouts lasting less than 30 minutes to bouts lasting at least 30 minutes; it does not represent adding sedentary time to the day. The fully adjusted primary model additionally conditioned on device-measurement variables, steps, and measured sociodemographic and health covariates. Sensitivity analyses examined whether the findings depended on that adjustment set, the ratio parameterization, score bounds, or a linear exposure-response form.

Methods

Study Design and Participants

We conducted a longitudinal observational analysis linking BCS70 age-46 main-sweep, age-46 activPAL, and age-51 main-sweep data[11–13]. The exposure was measured at age 46 and outcomes at age 51. The estimand was an adjusted longitudinal between-participant association, not a causal effect or measure of predictive performance.

The target population comprised age-46 participants asked about willingness to wear an activPAL monitor. Stage 1 identified participants with at least 3 archive-defined valid days and a nonmissing exposure. Stage 2 identified device-selected participants with an observed age-51 outcome and was defined separately for each outcome. Inference assumed conditional exchangeability of selection and response, positivity, correctly specified selection models, and missing at random[14, 15]. Counts were rounded to the nearest 10, and cells containing 1 to 9 participants were suppressed.

The age-46 sweep received approval from the South East Coast-Brighton and Sussex Research Ethics Committee (15/LO/1446). The age-51 sweep was approved by the London-Brighton & Sussex Research Ethics Committee (19/LO/1757). Participants provided informed consent. We analyzed deidentified secondary data under UK Data Service licensing and governance conditions.

ActivPAL Exposure

The exposure was the proportion of waking sitting/lying time accumulated in bouts lasting at least 30 minutes[10, 12, 16]. The numerator comprised participant-average sitting/lying hours in 30- to 60-minute and longer-than-60-minute bouts. The denominator was participant-average total waking sitting/lying time. The ratio was analyzed per 10 percentage points. It represented a ratio of participant averages rather than an average of daily ratios.

The 30-minute threshold was retained because it was directly reconstructable from the archived duration bins. It was treated as an operational definition rather than a biological boundary. At fixed total sitting/lying time, a higher proportion represented reallocation from shorter to longer bouts.

We used the archive valid-day classification and required at least 3 valid days. Valid days required at least 10 hours of waking wear, at least 500 steps, and no activity category exceeding 95% of waking time[10, 12]. The first partial monitoring day was excluded. A device-quality-control sensitivity also removed records with more than 24 hours of waking wear and recalculated participant averages.

Outcomes

The primary outcome was age-51 SF-36 Physical Functioning. The secondary outcome was pain burden, defined as 100 minus SF-36 Bodily Pain[17]. Both ranged from 0 to 100. Higher scores indicated better Physical Functioning and greater pain burden, respectively.

Archive-derived scores averaged available recoded items according to cohort documentation and the 1993 SF-36 manual[13, 18]. Primary models used these scores and adjusted for the corresponding age-46 outcome. As a quality check, we independently reconstructed Physical Functioning from at least 5 of 10 items and Bodily Pain from at least 1 of 2 items.18 Reconstructed scores were not substituted into the primary analysis.

Covariates and Follow-up Time

Core models included the corresponding age-46 outcome, age, sex, follow-up time, waking wear time, valid-day count, and total sitting/lying time. Sociodemographic models added education and employment. Fully adjusted models additionally included body mass index, smoking, alcohol use, sleep duration, Malaise score, Equality Act disability, and steps per day.

These models represented alternative conditioning specifications rather than a hierarchy in which additional adjustment necessarily reduced bias[19]. The analytical roles of covariates are summarized in Supplementary Table S1 and Figure S4.

Follow-up was calculated from the midpoints of available interview or monitoring months. The monitoring month defined baseline when it occurred from 90 days before to 365 days after the age-46 interview month. Otherwise, the interview month was used. A sensitivity analysis used the interview month for all participants.

Statistical Analysis

We generated 50 imputed data sets using chained bootstrap predictive mean matching in R (10 cycles, 5 donors, seed 20260724)[20–23]. Baseline covariates were imputed within the target population (rounded N = 7440). Follow-up time was imputed only among participants with at least 1 observed age-51 outcome. Outcomes were not imputed.

Predictors included demographic and baseline variables, selection and response indicators, age-51 linkage, and observed age-51 outcomes. Follow-up-time imputation also included device variables and age at follow-up. Supplementary Table S3 reports the implementation and diagnostic specifications. Diagnostics included chain traces, observed-versus-imputed summaries, fraction of missing information, Monte Carlo error, and relative efficiency[24, 25]. A sensitivity analysis used type-specific logistic or multinomial imputation for categorical variables (seed 20260813).

Two logistic selection models were estimated within each imputed data set. Stage 1 modeled device selection using baseline demographic, health, and behavioral predictors. Stage 2 modeled outcome response among device-selected participants and additionally included the exposure and device-measurement variables. Full formulas and reference categories are provided in Supplementary Table S5.

The sequential weight was the inverse product of the Stage 1 and outcome-specific Stage 2 probabilities[14]. Within each outcome analysis, weights were normalized to mean 1, truncated at the 1st and 99th percentiles, and renormalized[15]. Stage-specific covariate balance was assessed against the corresponding risk set[26]. The archive age-51 self-completion weight was evaluated as a sensitivity analysis because it targeted a broader survey population.

Weighted linear regression estimated the adjusted mean difference in each age-51 outcome per 10-percentage-point higher exposure. HC3 covariance estimates treated the realized weights as fixed within each imputation. Coefficients and variances were combined using Rubin's rules[21, 27].

A parametric sensitivity analysis propagated uncertainty in the fitted selection models using 200 coefficient draws per imputation. This procedure was not a full nonparametric bootstrap. Holm correction was restricted to the 2 fully adjusted primary outcome tests[28]. Other analyses were reported without multiplicity adjustment.

Selection and weighting sensitivities included complete cases, unweighted imputation, alternative weight truncation, a consenting-participant target, and archive weights. Measurement and modeling sensitivities examined device quality control, alternative timing, reduced adjustment, absolute long-bout time, and a long-/short-bout partition.

Fractional-logit models estimated marginal score differences for bounded outcomes[29]. Weighted quasibinomial-logit models estimated marginal risk differences for any Physical Functioning limitation and any pain burden. Four-knot restricted cubic splines used the 5th, 35th, 65th, and 95th exposure percentiles[30, 31]. Nonlinear terms were jointly tested using Rubin total covariance[21]. Influence diagnostics and prespecified exclusion refits were repeated across all imputations.

Software and Data Handling

Preprocessing used Python 3.12.13 with pandas 2.2.3 and NumPy 2.3.5. Models used R 4.5.2 with sandwich 3.1-1, MASS 7.3–65, and nnet 7.3–20. Reporting followed STROBE recommendations[32].

Results

Participant Flow

The age-46 main sweep included approximately 8580 participants. The activPAL willingness question was administered to approximately 7440 participants, approximately 6560 consented, and about 5260 met archive device selection. Approximately 4600 linked to the age-51 main sweep. Each outcome-specific sample rounded to 3820. Approximately 3810 participants had both outcomes observed; the Physical-Functioning-only and pain-burden-only cells each contained fewer than 10 participants and were suppressed. Approximately 1440 device-selected participants had neither outcome observed (Fig. 1).

Figure 1. Outcome-Specific Response Risk Sets

Stage 1 identifies archive device selection within the target population. Stage 2 branches directly from all device-selected participants to the 2 outcome-response samples and combines age-51 nonlinkage with absence of the relevant outcome. Linkage is displayed only as a descriptive annotation. Counts are rounded to the nearest 10; cells containing 1 to 9 participants are suppressed. Exact exclusion counts should not be inferred by subtracting rounded totals.

Descriptive Characteristics

Table 1 describes the outcome-specific samples, with full distributions in Supplementary Table S7 and Figure S1. Among all device-selected participants, the long-bout proportion had a median of 49.15% (interquartile range, 39.93%–58.21%; 5th–95th percentiles, 26.91%–69.55%; observed range, 3.12%–87.23%). Age-51 Physical Functioning had a mean of 89.66 (SD, 18.79) and median of 95 (interquartile range, 90–100); 47.3% scored 100 and 0.6% scored 0. Pain burden had a mean of 21.59 (SD, 22.12) and median of 10 (interquartile range, 10–32); 24.7% scored 0 and 1.2% scored 100.

Table 1
Observed Characteristics and Distributions of the Outcome-Specific Analysis Samples

Characteristic	Physical Functioning sample	Pain-burden sample
Analysis sample		
Participants	3,820	3,820
Demographic characteristics		
Age at baseline, y	47.24 (0.66)	47.24 (0.66)
Sex, male	1,690 (44.3%)	1,690 (44.3%)
Sex, female	2,130 (55.7%)	2,130 (55.7%)
Socioeconomic characteristics		
Education, lower	2,020 (53.5%)	2,020 (53.6%)
Education, intermediate	590 (15.5%)	590 (15.5%)
Education, degree	1,170 (31.0%)	1,170 (30.9%)
Employment, employed	3,460 (90.8%)	3,460 (90.9%)
Employment, unemployed or training	50 (1.4%)	50 (1.4%)
Employment, sick or disabled	90 (2.3%)	90 (2.3%)
Employment, other	210 (5.5%)	210 (5.5%)
Device-derived exposure		
Long-bout accumulation proportion, % of waking sitting/lying time	48.94 (12.80)	48.95 (12.80)
Device characteristics		
Waking wear time, h/d	15.86 (1.17)	15.86 (1.17)
Archive-defined valid days, d	6.49 (1.08)	6.49 (1.08)
Total waking sitting/lying time, h/d	9.27 (1.88)	9.27 (1.88)
Steps, steps/d	9,555 (3,560)	9,555 (3,559)
Baseline health		

Note. Values are mean (SD), median (IQR), observed range, or n (%), as indicated. Participant counts are rounded to the nearest 10; cells containing 1 to 9 participants are suppressed. Percentages use exact unsuppressed denominators and may not be reproduced by dividing displayed rounded counts. The 2 outcome-specific samples overlap substantially but are not identical. Descriptive values use observed data; variable-specific missingness is reported in Supplementary Table S2. SF-36 indicates 36-Item Short Form Health Survey.

Characteristic	Physical Functioning sample	Pain-burden sample
SF-36 Physical Functioning, points	89.93 (18.82)	89.96 (18.78)
Pain burden, points	19.96 (21.87)	19.95 (21.87)
Body mass index, kg/m²	28.05 (5.34)	28.04 (5.34)
Sleep duration, h/d	6.83 (1.10)	6.83 (1.10)
Malaise score, points	1.58 (1.96)	1.58 (1.96)
Equality Act disability, no	3,230 (84.7%)	3,230 (84.7%)
Equality Act disability, yes	590 (15.3%)	590 (15.3%)
Health behaviour		
Smoking, never	1,980 (51.9%)	1,980 (51.9%)
Smoking, former	1,250 (32.8%)	1,260 (32.9%)
Smoking, current	580 (15.2%)	580 (15.2%)
Alcohol use, abstainer	340 (8.9%)	340 (8.9%)
Alcohol use, lower risk	2,660 (69.8%)	2,660 (69.8%)
Alcohol use, higher risk	810 (21.3%)	810 (21.3%)
Exercise frequency, d/wk	3.04 (2.38)	3.04 (2.38)
Follow-up and outcomes		
Follow-up time, y	4.93 (0.88)	4.93 (0.88)
Age-51 SF-36 Physical Functioning, points	89.66 (18.79)	89.68 (18.77)
Age-51 pain burden, points	21.59 (22.13)	21.59 (22.12)
Age-51 Physical Functioning at ceiling, score = 100	1,810 (47.3%)	1,800 (47.3%)
Long-bout accumulation proportion, median (IQR), %	49.17 (39.91–58.11)	49.17 (39.91–58.11)
Long-bout accumulation proportion, observed range, %	3.12–85.99	3.12–85.99
Age-51 Physical Functioning, median (IQR)	95 (90–100)	95 (90–100)

Note. Values are mean (SD), median (IQR), observed range, or n (%), as indicated. Participant counts are rounded to the nearest 10; cells containing 1 to 9 participants are suppressed. Percentages use exact unsuppressed denominators and may not be reproduced by dividing displayed rounded counts. The 2 outcome-specific samples overlap substantially but are not identical. Descriptive values use observed data; variable-specific missingness is reported in Supplementary Table S2. SF-36 indicates 36-Item Short Form Health Survey.

Characteristic	Physical Functioning sample	Pain-burden sample
Age-51 Physical Functioning at floor, score = 0	20 (0.6%)	20 (0.6%)
Age-51 pain burden, median (IQR)	10 (10–32)	10 (10–32)
Age-51 pain burden at floor, score = 0	940 (24.7%)	940 (24.7%)
Age-51 pain burden at ceiling, score = 100	40 (1.2%)	40 (1.2%)

Note. Values are mean (SD), median (IQR), observed range, or n (%), as indicated. Participant counts are rounded to the nearest 10; cells containing 1 to 9 participants are suppressed. Percentages use exact unsuppressed denominators and may not be reproduced by dividing displayed rounded counts. The 2 outcome-specific samples overlap substantially but are not identical. Descriptive values use observed data; variable-specific missingness is reported in Supplementary Table S2. SF-36 indicates 36-Item Short Form Health Survey.

Primary Associations

In the fully adjusted primary linear analysis, a 10-percentage-point higher exposure corresponded to 0.426 Physical Functioning points (95% confidence interval, − 0.096 to 0.948; robust SE, 0.266; P=.109; Holm-adjusted P=.219) and − 0.099 pain-burden points (− 0.658 to 0.461; robust SE, 0.285; P=.729; Holm-adjusted P=.729), with rounded n = 3820 for each outcome (Table 2 and Fig. 2). The estimates were small relative to the 0–100 scales, but their intervals do not establish equivalence or rule out all practically important associations.

Table 2
Fully Adjusted Linear Associations Between Long-Bout Accumulation and Age-51 Outcomes

Outcome	Adjusted mean difference (95% CI)	Robust SE	P value	Holm-adjusted P value
SF-36 Physical Functioning	0.426 (− 0.096 to 0.948)	0.266	.109	.219
Pain burden (reverse-scored Bodily Pain)	−0.099 (− 0.658 to 0.461)	0.285	.729	.729

Note. Adjusted differences are reported per 10-percentage-point higher long-bout accumulation proportion and are conditional on the corresponding baseline outcome, age, sex, follow-up time, waking wear time, valid-day count, total sitting/lying time, steps, and measured sociodemographic and health covariates. Models used 50 imputations, sequential selection weights truncated at the 1st and 99th percentiles, fixed-weight HC3 robust standard errors, and Rubin pooling. Confidence intervals are nominal; Holm correction was applied only to the 2 primary fully adjusted outcome tests. Pain burden denotes reverse-scored SF-36 Bodily Pain.

Across the three adjustment specifications, Physical Functioning estimates were − 0.288 points in the core model (95% confidence interval, − 0.878 to 0.302), − 0.069 after education and employment were added (− 0.625 to 0.488), and 0.426 after the remaining health and movement variables were added (− 0.096 to 0.948). Corresponding pain-burden estimates were 0.115 (− 0.453 to 0.684), 0.033 (− 0.527 to

0.593), and − 0.099 (− 0.658 to 0.461). Reduced models excluding total sitting, steps, wear, and valid-day count yielded 0.138 (− 0.338 to 0.614) and − 0.100 (− 0.600 to 0.401), respectively. The direction change therefore reflects dependence on the conditioning set rather than a progressive approach to a single causal effect (Supplementary Table S4).

Figure 2. Fully Adjusted Linear Estimates and Nominal 95% Confidence Intervals

Stage 1 identifies archive device selection within the target population. Stage 2 branches directly from all device-selected participants to the 2 outcome-response samples and combines age-51 nonlinkage with absence of the relevant outcome. Linkage is displayed only as a descriptive annotation. Counts are rounded to the nearest 10; cells containing 1 to 9 participants are suppressed. Exact exclusion counts should not be inferred by subtracting rounded totals.

Analyses Accounting for Outcome Bounds and Nonlinearity

Fractional-logit average marginal score differences were 0.407 Physical Functioning points (95% confidence interval, − 0.073 to 0.887; P=.096) and − 0.079 pain-burden points (− 0.623 to 0.465; P=.777). Average marginal risk differences were − 0.281 percentage points for any Physical Functioning limitation (score < 100; −1.617 to 1.056; P=.681) and 0.645 percentage points for any pain burden (score > 0; −0.524 to 1.814; P=.279). Four-knot restricted cubic splines gave overall and nonlinear P values of .340 and .785 for Physical Functioning and .587 and .405 for pain burden, respectively. Full estimates, curves, and exposure support are shown in Supplementary Table S7 and Figure S2.

Sensitivity Analyses

Across 11 comparable sensitivity analyses using the long-bout accumulation proportion, excluding the conventional influence screen and models with a different time-use estimand, Physical Functioning coefficients ranged from 0.138 to 0.504, with 95% CI bounds ranging from − 0.339 to 1.187. Pain-burden coefficients ranged from − 0.158 to 0.079, with 95% CI bounds ranging from − 0.868 to 0.577. Type-specific categorical imputation yielded 0.423 (− 0.099 to 0.944) and − 0.102 (− 0.661 to 0.458). Interview-month timing gave 0.425 (− 0.097 to 0.947) and − 0.104 (− 0.664 to 0.455). Device-quality-control estimates were 0.461 (− 0.059 to 0.982) and − 0.095 (− 0.653 to 0.464) (Supplementary Tables S4 and S9).

For each additional 30 minutes/day of long-bout sitting while total sitting/lying was held fixed, estimates were 0.191 Physical Functioning points (95% confidence interval, − 0.081 to 0.462; P=.168) and − 0.054 pain-burden points (− 0.347 to 0.238; P=.715). In a distinct time-partition model that held shorter-bout sitting constant and allowed total sitting to change, estimates were 0.137 (− 0.094 to 0.368; P=.244) and − 0.286 (− 0.520 to − 0.052; nominal P=.017). Because the latter analysis was specified after the primary analysis, estimated a different contrast, and was not multiplicity-adjusted, it was treated as exploratory rather than confirmatory (Supplementary Table S4).

Seventy participants (rounded) had 79 archived days with more than 24 hours of waking wear. The original and recalculated exposures correlated at r=.99924; the median absolute difference was 0.037 percentage points, fewer than 10 participants lost eligibility, and none gained eligibility. Approximately 5150 paired participant averages differed numerically, largely because day-level values were compared with rounded archive averages (Supplementary Table S9).

Weighting and Model Diagnostics

All selection models converged. Mean minimum Stage-1 probability was 0.302; mean minimum Stage-2 probabilities were 0.310 for Physical Functioning and 0.308 for pain burden. A mean 1.99% of final weights was truncated, effective sample sizes were 3580.2 and 3578.4, and mean maximum weights were 2.127 and 2.144. Overall maximum absolute standardized mean difference declined from 0.139 before sequential weighting to 0.017 afterward. Within the Stage-2 risk set, the largest maximum absolute standardized mean difference declined from 0.084 to 0.004 for each outcome when the outcome-response weight was applied (Supplementary Tables S5–S6 and Figure S5).

Imputation, Influence, and Weight-Uncertainty Diagnostics

Monte Carlo standard errors were approximately 1.01% of the pooled standard error for Physical Functioning and 0.92% for pain burden; relative efficiencies exceeded 0.9998 (Supplementary Table S3). Figure S3 shows imputation-chain traces. Across all 50 imputations, the conventional Cook's distance/DFBETA union flagged a mean of approximately 320 Physical Functioning and 290 pain-burden observations, and the screen excluded approximately 8% of each sample. Excluding the top 1% of absolute exposure DFBETA values yielded 0.204 (− 0.208 to 0.616) and 0.079 (− 0.418 to 0.577), respectively. Propagating propensity-coefficient uncertainty changed the intervals only slightly, to − 0.0964 to 0.9493 and − 0.6591 to 0.4616 (Supplementary Tables S5 and S8).

Discussion

In the target population, the fully adjusted primary linear models showed no clear conditional association of the age-46 long-bout proportion with age-51 Physical Functioning or pain burden. Fractional-logit and boundary-probability analyses were similar to the primary estimates, and restricted cubic spline analyses provided no clear evidence of either an overall association or nonlinearity across the 5th–95th percentile exposure range. Across comparable sensitivity analyses using the long-bout accumulation proportion, Physical Functioning estimates ranged from 0.138 to 0.504 points and pain-burden estimates ranged from − 0.158 to 0.079 points, with appreciable uncertainty. An exploratory time-partition model produced a nominal inverse pain association, but it estimated a different behavioral contrast and was not multiplicity-adjusted.

Our findings differ from the cross-sectional association between bouts of at least 60 minutes and poorer objective function in older European adults[4]. They also extend a cross-sectional BCS70 analysis of upright and stepping event patterns by examining subsequent self-reported outcomes[5]. Direct

comparison with EPIC-Norfolk is limited because that study examined broader movement behaviors and later walking or chair-stand performance in older participants.6

Prospective evidence is strongest in vulnerable older groups. Sedentary time predicted functional loss in adults at risk of knee osteoarthritis[33] and disability in mobility-limited older adults[34]. In EPIC-Norfolk, associations between total sedentary time and physical capability attenuated after accounting for moderate-to-vigorous activity[35]. The Maastricht Study found only marginal independent associations between activPAL sedentary measures and objective function among adults aged 40 to 75 years[36]. In another older cohort, sitting-related associations attenuated after activity adjustment and were most apparent at low activity levels[37].

The outcome distributions required direct attention. Nearly half of participants scored 100 for Physical Functioning, compressing variation at the upper bound, while one quarter had pain burden of 0. Fractional-logit marginal score differences closely resembled the linear estimates, and models for any Physical Functioning limitation or any pain burden did not show clear probability differences. These checks reduce concern that the primary conclusion is solely an artifact of treating bounded scores as Gaussian, but no model restores information lost at a scale boundary. The spline curves likewise provided no clear evidence of a systematic pattern across the 5th–95th percentile exposure range, while estimates remained imprecise for small or localized departures.

The ratio coefficient has a specific behavioral interpretation. With total sitting/lying held fixed, a higher long-bout proportion reallocates sedentary time from bouts shorter than 30 minutes to bouts at least 30 minutes; it does not add sedentary time. The absolute-duration model that also conditioned on total sitting gave similar weak estimates. By contrast, the time-partition model held shorter-bout time fixed and allowed total sitting to rise: pain burden decreased by 0.286 points per additional 30 minutes of long-bout time (nominal P=.017). This exploratory result is not evidence that prolonged sitting is beneficial; it represents a different estimand and illustrates that conclusions about direction can depend on how the time-use composition is parameterized.

Trial evidence also remains limited. A Cochrane review of 7 trials found no clear improvement in Short Physical Performance Battery scores or 6-minute walking distance, with low certainty[38]. A later review identified only 5 unique studies and noted difficulty separating sedentary interruptions from added physical activity[39]. This cautions against causal interpretation.

The pain finding is compatible with uncertain prior evidence. CARDIA reported directionally unexpected associations among sedentary time, activity, and knee symptoms[7]. A broader systematic review found insufficient prospective evidence relating sedentary behavior to musculoskeletal pain[8]. A prospective-cohort review found no association with new low-back pain (pooled risk ratio, 1.01; 95% confidence interval, 0.98 to 1.04)[40]. Most studies used self-reported exposure. Null longitudinal findings also came from a co-twin study[41] and a cohort of 76,438 Danish adults[42]. Conversely, objectively measured prolonged sitting can provoke immediate back discomfort, but this has not established an effect on later

clinical episodes[43]. Our general pain-burden result does not exclude site-specific, occupational, immediate, or clinically defined pain associations.

The primary point estimates were small on 0–100 scales, but published important-change thresholds vary by population, anchor, direction, and whether the contrast concerns within-person change or between-participant differences[44–46]. We therefore based interpretation on the observed estimates and confidence intervals rather than equivalence testing; equivalence would require independently justified margins specified before a confirmatory analysis.

The imputation and weighting diagnostics strengthen reproducibility but do not eliminate assumptions. Imputation requires a missing-at-random mechanism conditional on included variables. Selection weighting requires conditional exchangeability, positivity, and adequate model specification; favorable measured balance cannot exclude unmeasured selection. The main HC3 variance treated fitted weights as fixed. Parametric coefficient simulation suggested little additional variance from estimating the two logistic coefficient vectors, but it was not a participant-level bootstrap and does not capture all sources of analytic uncertainty.

Strengths include prospective linkage in a national birth cohort, posture-sensitive thigh-worn monitoring, baseline-outcome adjustment, 50 imputations, outcome-specific sequential weighting, explicit stage-specific balance assessment, robust covariance estimation, a clearly defined primary analysis, and extensive sensitivity analyses. Reporting effect ranges and 95% CI bounds across alternative specifications avoids judging robustness solely by whether intervals cross zero and makes the dependence on model form, covariate roles, and time-use estimand explicit.

Several limitations remain. First, a single 7-day monitoring period may not represent habitual accumulation during the subsequent 5 years. A repeated-measures study found moderate 6-month reproducibility for activPAL posture measures and potential attenuation from within-person variability[47]. Second, the exposure was reconstructed from archive-level mean duration bins. It cannot distinguish the number, exact duration, timing, or day-to-day consistency of sitting/lying events. Similar totals can arise from different bout distributions[48], and estimated associations can vary by bout definition[49]. The 30-minute threshold is therefore operational rather than a universal biological boundary. Event-level data would permit threshold-free accumulation curves, transition metrics, time-of-day patterns, and reliability analyses.

Third, activPAL improves postural classification relative to count-only monitors but does not identify behavioral context[16]. Occupational loading and reasons for lying were not measured. Fourth, device participation and follow-up were selective. Distribution procedures, wear location, and valid-day rules can influence recruitment, adherence, and data loss[50, 51]. The weighting target was restricted to the predefined age-46 target population rather than the full cohort or a nationally representative sample of UK adults in midlife. The official age-51 paper self-completion weight targets a broader survey population and was therefore only a sensitivity. Neither analysis-specific nor official weights remove selection on unmeasured variables or structural device eligibility.

Fifth, residual confounding and bidirectionality remain possible despite baseline-outcome adjustment. Several age-46 health and behavioral covariates were contemporaneous with monitoring, so reduced and extended models answer different conditional questions. Device monitoring and the interview were not always concurrent; however, accepted monitor months were usually identical to interview month and the all-interview-month sensitivity was nearly unchanged. Sixth, both outcomes were self-reported and bounded, pain burden did not localize pain, and the archive's partial-scale rules can score scales from limited item response.

Analyses accounting for outcome bounds and restricted cubic spline analyses addressed—but could not eliminate—concerns related to model form.

The inference is therefore limited to conditional associations for an archive-constrained 30-minute accumulation measure in the target population. The study did not estimate an intervention effect, establish equivalence or assess calibration, discrimination, or external predictive validation. Event-level data, repeated behavior measurement, and performance-based or site-specific outcomes remain important next steps.

The next step is to use event-level activPAL data to derive a limited set of distinct pattern measures. These could include threshold-free accumulation, transitions, break characteristics, day-to-day reliability, and time-of-day patterns. Future analyses should apply multiplicity control and prioritize replication. Repeated behavior measurement and performance-based outcomes would help separate behavioral change from functional decline.

Conclusions

In the BCS70 target population, and under the stated assumptions regarding selection and missing data, the primary linear models showed no clear evidence of an association between the long-bout accumulation proportion and SF-36 Physical Functioning or pain burden at age 51. Restricted cubic spline analyses similarly provided no clear evidence of a consistent nonlinear association across the 5th–95th percentile exposure range. Analyses accounting for outcome bounds yielded estimates similar to those from the primary analyses. A distinct exploratory partition model for pain underscored that time-use parameterization matters. These findings do not establish causality or equivalence and do not exclude associations involving other sedentary metrics, outcomes, or populations.

Abbreviations

BCS70
1970 British Cohort Study
CI
Confidence interval
HC3
Heteroskedasticity-consistent covariance estimator type 3

IQR
Interquartile range
SD
Standard deviation
SE
Standard error
SF-36
36-Item Short Form Health Survey
STROBE
Strengthening the Reporting of Observational Studies in Epidemiology

Declarations

Ethics approval and consent to participate

The study was conducted in accordance with the Declaration of Helsinki. The age-46 sweep of the 1970 British Cohort Study (BCS70) received ethical approval from the South East Coast–Brighton and Sussex Research Ethics Committee (15/LO/1446), and the age-51 sweep received ethical approval from the London–Brighton & Sussex Research Ethics Committee (19/LO/1757). Participants provided informed consent. The present study analyzed deidentified secondary data under UK Data Service licensing and governance conditions.

Consent for publication

Not applicable.

Availability of data and materials

The datasets analyzed in this study are available through the UK Data Service. The analysis used the BCS70 age-46 main sweep (SN 8547), age-46 accelerometry data (SN 8611), and age-51 main sweep (SN 9347). The authors are not permitted to redistribute these data. Eligible researchers may apply for access through the UK Data Service under the applicable licensing and governance conditions.

The analysis code is not publicly archived but is available from the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was funded by Macao Polytechnic University (Grant No. RP/FCSD-03/2022).

Authors' contributions

RZ contributed to conceptualization, methodology, formal analysis, investigation, data curation, visualization, writing the original draft, and reviewing and editing the manuscript. YW contributed to investigation, data curation, validation, and reviewing and editing the manuscript. DW contributed to methodology, formal analysis, validation, and reviewing the manuscript. HT contributed to conceptualization, methodology, and reviewing the manuscript. JN contributed to supervision and reviewing and editing the manuscript. QS contributed to conceptualization, supervision, project administration, and reviewing and editing the manuscript. All authors read and approved the final manuscript.

Acknowledgments

We thank the BCS70 cohort members, the Centre for Longitudinal Studies, and the UK Data Service. These organizations bear no responsibility for the analyses or interpretations reported here.

References

1. Tremblay MS, Aubert S, Barnes JD, et al. Sedentary Behavior Research Network (SBRN) - Terminology Consensus Project process and outcome. Int J Behav Nutr Phys Activity. 2017;14:75. 10.1186/s12966-017-0525-8.
2. World Health Organization. WHO Guidelines on Physical Activity and Sedentary Behaviour. World Health Organization; 2020. https://www.who.int/publications/i/item/9789240015128.
3. Diaz KM, Howard VJ, Hutto B, et al. Patterns of Sedentary Behavior and Mortality in U.S. Middle-Aged and Older Adults: A National Cohort Study. Ann Intern Med. 2017;167(7):465–75. 10.7326/M17-0212.
4. Wilson JJ, McMullan I, Blackburn NE, et al. Associations of sedentary behavior bouts with community-dwelling older adults' physical function. Scand J Med Sci Sports. 2021;31(1):153–62. 10.1111/sms.13827.
5. Culverhouse J, Hillsdon M, Pulsford R. Cross-sectional associations between temporal patterns and composition of upright and stepping events with physical function in midlife: Insights from the 1970 British Cohort Study. Scand J Med Sci Sports. 2024;34:e14645. 10.1111/sms.14645.
6. Yerrakalva D, Hajna S, Wijndaele K, et al. Bidirectional associations of accelerometer-assessed physical activity and sedentary time with physical function among older English adults: the EPIC-Norfolk cohort study. Eur J Ageing. 2022;19(4):1507–17. 10.1007/s10433-022-00733-y.
7. Whitaker KM, Pettee Gabriel K, Laddu D, et al. Bidirectional associations of accelerometer measured sedentary behavior and physical activity with knee pain, stiffness, and physical function: The CARDIA study. Prev Med Rep. 2021;22:101348. 10.1016/j.pmedr.2021.101348.
8. Dzakpasu FQS, Carver A, Brakenridge CJ, et al. Musculoskeletal pain and sedentary behaviour in occupational and non-occupational settings: a systematic review with meta-analysis. Int J Behav

Nutr Phys Activity. 2021;18(1):159. 10.1186/s12966-021-01191-y.

9. Sullivan A, Brown M, Hamer M, Ploubidis GB, Cohort Profile Update. The 1970 British Cohort Study (BCS70). Int J Epidemiol. 2023;52(3):e179–86. 10.1093/ije/dyac148.
10. Hamer M, Stamatakis E, Chastin S, et al. Feasibility of Measuring Sedentary Time Using Data From a Thigh-Worn Accelerometer. Am J Epidemiol. 2020;189(9):963–71. 10.1093/aje/kwaa047.
11. University of London I of E Centre for Longitudinal Studies. 1970 British Cohort Study: Age 46, Sweep 10, 2016–2018. Published online 2024. 10.5255/UKDA-SN-8547-2
12. University of London I of E Centre for Longitudinal Studies. 1970 British Cohort Study: Age 46, Sweep 10 Accelerometry Data, 2016–2018. Published online 2024. 10.5255/UKDA-SN-8611-1
13. University of London I of E Centre for Longitudinal Studies. 1970 British Cohort Study: Age 51, Sweep 11, 2021–2024. Published online 2025. 10.5255/UKDA-SN-9347-1
14. Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. Stat Methods Med Res. 2013;22(3):278–95. 10.1177/0962280210395740.
15. Cole SR, Hernán MA. Constructing Inverse Probability Weights for Marginal Structural Models. Am J Epidemiol. 2008;168(6):656–64. 10.1093/aje/kwn164.
16. Edwardson CL, Winkler EAH, Bodicoat DH, et al. Considerations when using the activPAL monitor in field-based research with adult populations. J Sport Health Sci. 2017;6(2):162–78. 10.1016/j.jshs.2016.02.002.
17. Ware JEJ, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992;30(6):473–83.
18. Ware JE Jr, Snow KK, Kosinski M, Gandek B. SF-36 Health Survey: Manual and Interpretation Guide. The Health Institute, New England Medical Center; 1993.
19. VanderWeele TJ. Principles of confounder selection. Eur J Epidemiol. 2019;34(3):211–9. 10.1007/s10654-019-00494-6.
20. Sterne JAC, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393. 10.1136/bmj.b2393.
21. Rubin DB. Multiple Imputation for Nonresponse in Surveys. Wiley; 1987. 10.1002/9780470316696.
22. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011;30(4):377–99. 10.1002/sim.4067.
23. Morris TP, White IR, Royston P. Tuning multiple imputation by predictive mean matching and local residual draws. BMC Med Res Methodol. 2014;14(1):75. 10.1186/1471-2288-14-75.
24. Abayomi K, Gelman A, Levy M. Diagnostics for Multivariate Imputations. J Royal Stat Soc Ser C: Appl Stat. 2008;57(3):273–91. 10.1111/j.1467-9876.2007.00613.x.
25. Royston P, Carlin JB, White IR. Multiple Imputation of Missing Values: New Features for Mim. Stata J. 2009;9(2):252–64. 10.1177/1536867X0900900205.
26. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational

studies. Stat Med. 2015;34(28):3661–79. 10.1002/sim.6607.

27. MacKinnon JG, White H. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. J Econ. 1985;29(3):305–25. 10.1016/0304-4076(85)90158-7.
28. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat. 1979;6(2):65–70.
29. Papke LE, Wooldridge JM. Econometric methods for fractional response variables with an application to 401(k) plan participation rates. J Appl Econom. 1996;11(6):619–32. 10.1002/(SICI)1099-1255(199611)11:6%3C619::AID-JAE418%3E3.0.CO;2-1.
30. Durrleman S, Simon R. Flexible regression models with cubic splines. Stat Med. 1989;8(5):551–61. 10.1002/sim.4780080504.
31. Harrell FE Jr. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Springer International Publishing; 2015. 10.1007/978-3-319-19425-7.
32. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche 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. 10.1016/S0140-6736(07)61602-X.
33. Semanik PA, Lee J, Song J, et al. Accelerometer-Monitored Sedentary Behavior and Observed Physical Function Loss. Am J Public Health. 2015;105(3):560–6. 10.2105/AJPH.2014.302270.
34. Mankowski RT, Anton SD, Axtell R, et al. Device-measured physical activity as a predictor of disability among mobility-limited older adults. J Am Geriatr Soc. 2017;65(10):2251–6. 10.1111/jgs.15037.
35. KEEVIL VL, COOPER AJM, WIJNDAELE K, et al. Objective Sedentary Time, Moderate-to-Vigorous Physical Activity, and Physical Capability in a British Cohort. Med Sci Sports Exerc. 2016;48(3):421–9. 10.1249/MSS.0000000000000785.
36. van der Velde JHPM, Savelberg HHCM, van der Berg JD, et al. Sedentary Behavior Is Only Marginally Associated with Physical Function in Adults Aged 40–75 Years—the Maastricht Study. Front Physiol. 2017;8. 10.3389/fphys.2017.00242.
37. Walker RL, Greenwood-Hickman MA, Bellettiere J, et al. Associations between physical function and device-based measures of physical activity and sedentary behavior patterns in older adults: moving beyond moderate-to-vigorous intensity physical activity. BMC Geriatr. 2021;21:216. 10.1186/s12877-021-02163-4.
38. Chastin S, Gardiner P, Harvey J et al. Interventions for reducing sedentary behaviour in community-dwelling older adults. *Cochrane Database of Systematic Reviews*. 2021;2021. 10.1002/14651858.cd012784.pub2
39. Swartz A, Steinbrink G, Strath S, Mitra T, Morelli W. A Systematic Review of the Effects of Breaking up/Interrupting Sedentary Behavior on Health Outcomes Among Community-Dwelling Adults 60 + Years. J Aging Phys Act. 2024;33:1–22. 10.1123/japa.2023-0416.
40. Damato TM, Christofaro DGD, Pinheiro MB, et al. Does sedentary behaviour contribute to the development of a new episode of low back pain? A systematic review of prospective cohort studies. Eur J Pain. 2022;26(7):1412–23. 10.1002/ejp.1977.

41. Amorim AB, Levy GM, Pérez-Riquelme F, et al. Does sedentary behavior increase the risk of low back pain? A population-based co-twin study of Spanish twins. Spine J. 2017;17(7):933–42. 10.1016/j.spinee.2017.02.004.
42. Balling M, Holmberg T, Petersen C, Aadahl M, Meyrowitsch D, Tolstrup J. Total sitting time, leisure time physical activity and risk of hospitalization due to low back pain: The Danish Health Examination Survey cohort 2007–2008. Scand J Public Health. 2018;47:45–52. 10.1177/1403494818758843.
43. De Carvalho DE, de Luca K, Funabashi M, et al. Association of Exposures to Seated Postures With Immediate Increases in Back Pain: A Systematic Review of Studies With Objectively Measured Sitting Time. J Manipulative Physiol Ther. 2020;43(1):1–12. 10.1016/j.jmpt.2019.10.001.
44. Hays RD, Morales LS. The RAND-36 measure of health-related quality of life. Ann Med. 2001;33(5):350–7. 10.3109/07853890109002089.
45. Ward MM, Guthrie LC, Alba MI. Clinically important changes in Short Form 36 Health Survey scales for use in rheumatoid arthritis clinical trials: the impact of low responsiveness. Arthritis Care Res. 2014;66(12):1783–9. 10.1002/acr.22392.
46. Clement ND, Weir D, Deehan D. Meaningful values in the Short Form Health Survey-36 after total knee arthroplasty - an alternative to the EuroQol five-dimension index as a measure for health-related quality of life: minimal clinically important difference, minimal important change, patient-acceptable symptom state thresholds, and responsiveness. Bone Joint Res. 2022;11(7):477–83. 10.1302/2046-3758.117.BJR-2021-0493.R1.
47. Saint-Maurice PF, Sampson JN, Keadle SK, Willis EA, Troiano RP, Matthews CE. Reproducibility of Accelerometer and Posture-derived Measures of Physical Activity. Med Sci Sports Exerc. 2020;52(4):876–83. 10.1249/MSS.0000000000002206.
48. Chastin SFM, Granat MH. Methods for objective measure, quantification and analysis of sedentary behaviour and inactivity. Gait Posture. 2010;31(1):82–6. 10.1016/j.gaitpost.2009.09.002.
49. Kim Y, Welk GJ, Braun SI, Kang M. Extracting Objective Estimates of Sedentary Behavior from Accelerometer Data: Measurement Considerations for Surveillance and Research Applications. PLoS ONE. 2015;10(2):e0118078. 10.1371/journal.pone.0118078.
50. Roth MA, Mindell JS. Who provides accelerometry data? Correlates of adherence to wearing an accelerometry motion sensor: the 2008 Health Survey for England. J Phys Activity Health. 2013;10(1):70–8. 10.1123/jpah.10.1.70.
51. Pulsford RM, Brocklebank L, Fenton SAM, et al. The impact of selected methodological factors on data collection outcomes in observational studies of device-measured physical behaviour in adults: A systematic review. Int J Behav Nutr Phys Act. 2023;20:26. 10.1186/s12966-022-01388-9.

Figures

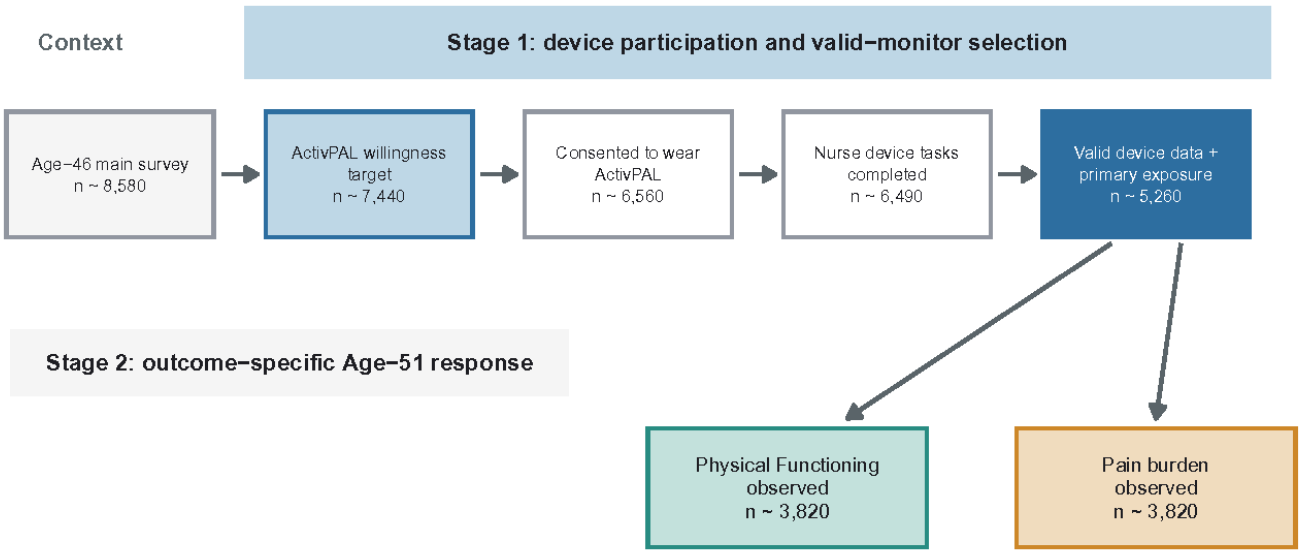


Figure 1

Outcome-Specific Response Risk Sets

Stage 1 identifies archive device selection within the target population. Stage 2 branches directly from all device-selected participants to the 2 outcome-response samples and combines age-51 nonlinkage with absence of the relevant outcome. Linkage is displayed only as a descriptive annotation. Counts are rounded to the nearest 10; cells containing 1 to 9 participants are suppressed. Exact exclusion counts should not be inferred by subtracting rounded totals.

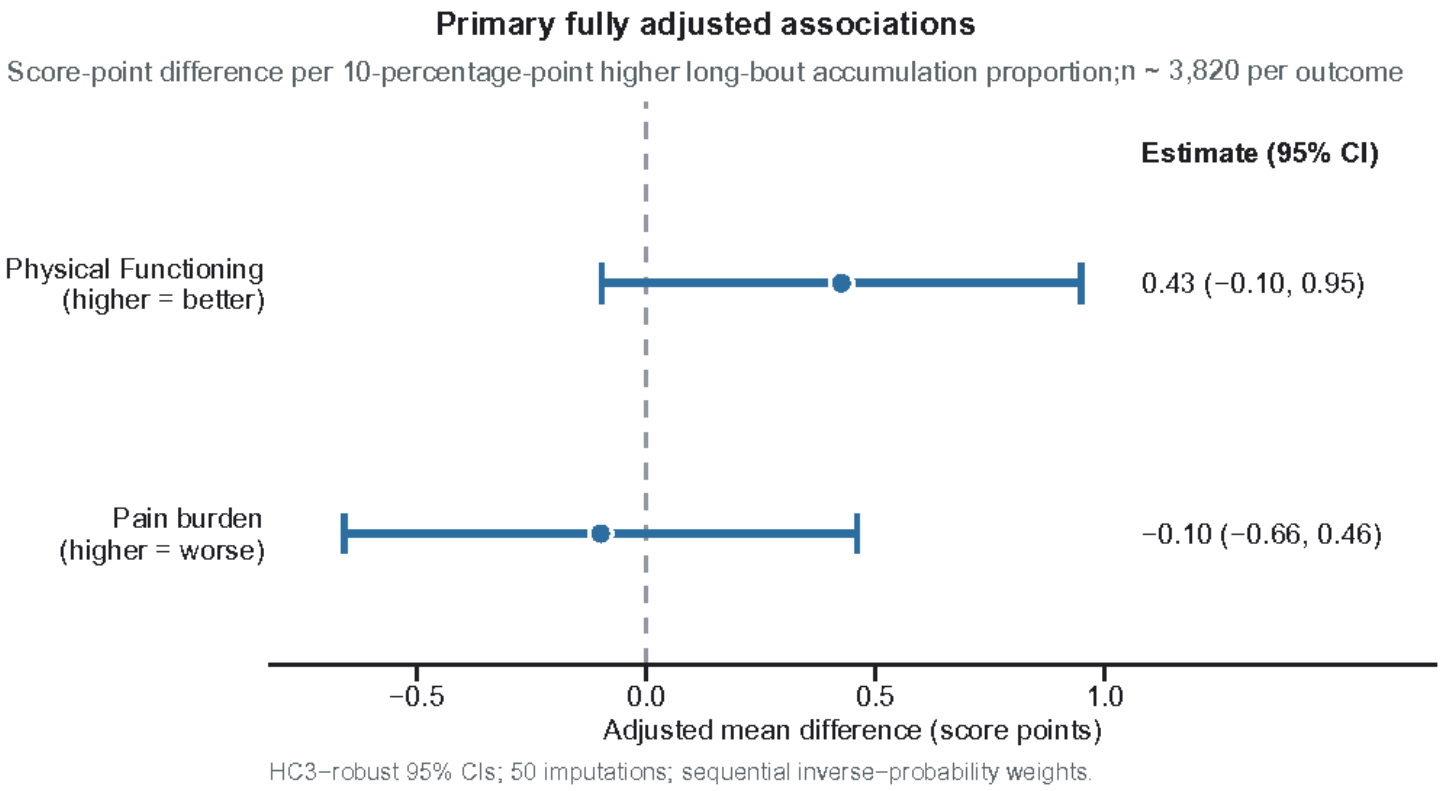


Figure 2

Fully Adjusted Linear Estimates and Nominal 95% Confidence Intervals

Stage 1 identifies archive device selection within the target population. Stage 2 branches directly from all device-selected participants to the 2 outcome-response samples and combines age-51 nonlinkage with absence of the relevant outcome. Linkage is displayed only as a descriptive annotation. Counts are rounded to the nearest 10; cells containing 1 to 9 participants are suppressed. Exact exclusion counts should not be inferred by subtracting rounded totals.

Supplementary Files

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

- Supplementary18Aug.docx