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.

Increasing and maintaining physical activity in later life: a target trial emulation of daily living problems or death across three ageing cohorts

Zewei Chen
Second Xiangya Hospital of Central South University
Fengying Yu
Second Xiangya Hospital of Central South University
Yiyuan Huang
Second Xiangya Hospital of Central South University
Fei Luo
Second Xiangya Hospital of Central South University
Ting Luo
Second Xiangya Hospital of Central South University
Zhenfei Fang
fangzhenfei@csu.edu.cn
Second Xiangya Hospital of Central South University

Research Article

Keywords: physical activity, healthy ageing, activities of daily living, mortality, target trial emulation, cohort study, health promotion, population impact

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11396575/v1

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

Additional Declarations: No competing interests reported.

Increasing and maintaining physical activity in later life: a target trial emulation of daily living problems or death across three ageing cohorts

Zewei Chen[1†], Fengying Yu[1,3†], Yiyuan Huang[1], Fei Luo[1,2], Ting Luo[1,3*], Zhenfei Fang[1*]

1 Department of Cardiovascular Medicine, The Second Xiangya Hospital, Central South University, Changsha, Hunan 410011, China.

2 Ataturk University, Yakutiye, Erzurum, Türkiye.

3 Clinical Nursing Teaching and Research Section, The Second Xiangya Hospital of Central South University, Changsha, Hunan, China.

† Zewei Chen and Fengying Yu contributed equally to this work.

* Corresponding authors: Ting Luo (luoting1986@csu.edu.cn) and Zhenfei Fang (fangzhenfei@csu.edu.cn).

Correspondence address: Department of Cardiovascular Medicine, The Second Xiangya Hospital, Central South University, Changsha, Hunan 410011, China.

Abstract

Background: Maintaining independence and survival are central goals of healthy ageing. For older adults who have been less active, the association of increasing and maintaining physical activity with subsequent daily living problems or death has implications for health promotion and programme planning.

Methods: We emulated sequential, visit-defined target trials using longitudinal cohorts from the USA, Europe and South Korea. Participants were aged 65 years or older, previously less active and independent in five basic activities of daily living (ADLs). We compared increased activity frequency maintained at the first follow-up with persistently lower frequency. Overlap weighting and inverse-probability adjustment for adherence and outcome observation estimated cumulative risks over two subsequent visits, generally about four years. Post hoc analyses examined baseline prognosis and hypothetical programme coverage.

Results: The analysis included 19204 participants contributing 43754 eligible trial entries. Risk ratios for an ADL problem or death were 0.776 (95% confidence interval 0.728–0.827) in the US cohort, 0.764 (0.672–

0.868) in the European cohort and 0.725 (0.643–0.816) in the South Korean cohort (all P<0.001). Absolute risk differences were −7.08, −8.15 and −4.89 percentage points, respectively. Mortality contributed substantially to these differences; evidence for ADL problems alone was less consistent. Larger absolute differences at higher baseline risk remained supported after multiplicity adjustment in the US cohort. Risk priority did not clearly outperform age priority in any cohort. With 50% programme reach and 50% strategy achievement among those reached, modelled reductions relative to the natural activity course were approximately 9–11 events per 1000 eligible entries.

Conclusions: Increasing and maintaining activity frequency was associated with lower risk of daily living problems or death across three ageing cohorts, with the clearest common signal concerning survival. These findings support evaluating sustained activity support within healthy ageing programmes, with attention to programme reach and successful strategy achievement.

Keywords: physical activity; healthy ageing; activities of daily living; mortality; target trial emulation; cohort study; health promotion; population impact

Background

Maintaining independence in everyday activities is a priority for older people [1, 2] and a central goal of healthy ageing. Physical activity is a modifiable behaviour recommended to support health and physical function in later life, including among people with chronic conditions [3]. For public health practice, an important population comprises older adults who remain independent in basic activities of daily living (ADLs) but have previously been less active. Understanding the outcomes associated with increasing activity in this group can inform the evaluation of health promotion programmes that seek to support independent living and survival.

Trials incorporating physical activity, including LIFE, SPRINTT and REACT, have demonstrated benefits for mobility or lower limb function in older adults [4-6]. Whether these gains extend to basic self-care is less certain: a LIFE secondary analysis did not demonstrate a reduction in basic ADL disability [7]. A longitudinal meta-analysis associated higher activity with lower incidence and progression of basic ADL disability [8], and recent cohort studies linked increasing activity with more favourable functional or survival outcomes [9, 10]. A target trial emulation in Australian women associated sustained adherence to activity guidelines, or first

attainment at age 55, with better physical health-related quality of life in later life [11]. These findings motivate assessment of a specific behavioural strategy in older adults: increasing activity frequency after a period of lower activity, followed by maintenance, with daily living problems and death assessed together and separately.

The public health relevance of this strategy depends on absolute risk differences as well as relative associations. Baseline prognosis influences the absolute difference corresponding to a given relative association [12]. Programme planning also requires attention to how many eligible people can be reached and how many of those reached achieve the intended activity strategy. Evidence across different ageing populations can clarify the consistency of associations while identifying differences in measurement and context. We therefore examined increased and maintained activity frequency in relation to basic ADL problems or death across three ageing cohorts and assessed hypothetical programme-coverage scenarios.

Methods

Study design

We emulated sequential trials using the US Health and Retirement Study (HRS), the Survey of Health, Ageing and Retirement in Europe (SHARE), and the Korean Longitudinal Study of Aging (KLoSA) [13-15]. Target trial specification and reporting followed the target trial framework, TARGET and STROBE [16-18]. Baseline waves were HRS 8–14 (2006–2018), SHARE 5–7 (2013–2017) and KLoSA 2–8 (2008–2020), with outcomes through 2022, 2021–2022 and 2024, respectively. Cohort-specific analyses accommodated differences in measurements and interview spacing. This secondary observational analysis did not assign interventions and was not prospectively registered.

Participants and strategies

At each baseline, participants were aged ≥65 years, reported lower activity at the preceding visit, and had no problem in five basic activities of daily living (ADLs) at either visit. Eligibility also required intact indoor mobility. HRS/SHARE required community-dwelling, non-proxy interviews without an earlier memory-related diagnosis; equivalent flags were unavailable in KLoSA. SHARE required earlier SHARELIFE participation; country–wave cells with <10% coverage of follow-up activity or ADL/vital status were

78 excluded. Participants could contribute repeated eligible entries. Two future survey slots had to exist, but
79 completing them was not required (Figure 1; Table S1).

80 Strategies were higher activity at baseline and first follow-up after lower prior frequency, or lower activity at
81 both visits. Higher frequency meant moderate activity more than once weekly in HRS/SHARE and general
82 exercise at least twice weekly in KLoSA. Time zero was the baseline recording visit; exact activity-change
83 dates were unknown. Adherence was updated only among first-follow-up survivors without an outcome. Non-
84 adherent or unknown activity then ended second-interval contribution; all first-interval outcomes remained
85 included. Second-follow-up activity did not define earlier adherence [19].

86 **Outcomes and covariates**

87 The primary outcome was a first recorded problem in dressing, bathing, eating, bed transfer or toileting, or all-
88 cause death across two subsequent visits. ADL problems meant difficulty in HRS/SHARE and assistance in
89 KLoSA, where bed transfer and room mobility were combined. Components were analysed as first recorded
90 events; independent all-cause mortality retained participants after an earlier ADL problem. A recorded death
91 was observable without an interview; living participants with unavailable ADL information were not assumed
92 event-free. Exact ADL-onset dates were unavailable. Follow-up was visit-defined, generally about four years.

93 Adjustment included trial wave, sex, education, cohort-specific race/ethnicity or country, prior activity, age
94 modelled with a three-degree-of-freedom natural spline, and preceding-visit and baseline health measures.
95 These comprised mobility, hearing, vision, self-rated health, cognition, depressive symptoms, IADLs, marital
96 status, smoking, alcohol use, hypertension, diabetes, heart disease, stroke, cancer, lung disease, arthritis,
97 healthcare and medication use, body mass index (BMI), wealth and income. First-follow-up health informed
98 later adherence and observation models. Education used years in HRS/SHARE and three ordinal levels in
99 KLoSA; KLoSA depressive scores were standardised within wave. Wealth/income used inverse hyperbolic
100 sine transformations after division by 10000, visit counts used log(1+count), and BMI was bounded to 12–60
101 kg/m². Numeric missingness received median replacement plus indicators; categorical missingness formed a
102 separate level. Unavailable or constant terms were omitted; no later cognition was carried backwards into
103 KLoSA waves 8–9. Missingness is reported in Table S3.

Statistical analysis

Four logistic models estimated baseline strategy, first-visit observation, adherence and second-visit observation probabilities. Overlap weights were 1−e for the increased-activity group and e for the lower-activity group, where e was the fitted baseline strategy probability; survey sampling weights were not used [20, 21]. First-interval weights divided overlap weights by observation probability. Second-interval weights additionally divided by adherence and second-observation probabilities among event-free, adherent survivors, with updated health predictors [22]. Probabilities were bounded at 0.025–0.975; combined weights were not separately trimmed. Weighted interval risks yielded cumulative risks, risk ratios (RRs) and risk differences (RDs). Component risks used composite survival so that ADL-first and death-first risks summed to composite risk. Mortality analyses refitted the models and retained survivors after ADL events. Number needed to benefit was the reciprocal absolute risk reduction, with finite benefit intervals reported only when the reduction interval excluded zero. Balance was assessed using standardised differences [23]. Person-clustered joint influence functions incorporated weighting-model uncertainty and repeated entries. Risk and RR intervals used logit and log scales; RD intervals used the normal scale. Two-sided P values and 95% confidence intervals were reported; 300 whole-person full-refit bootstraps assessed primary interval stability. Preprocessing and primary single imputations remained fixed during resampling.

Sensitivity analyses retained the earliest entry, outcomes before 2020, complete core covariates, no recent health deterioration, or previous-only/current-only health adjustment. Core fields were age, education, self-rated health, BMI, depressive symptoms, cognition, IADLs, previous self-rated health and ADLs, plus walking where available. Recent deterioration meant new heart disease, stroke or cancer, self-rated health worsening by ≥2 categories, or baseline hospital admission. E-values assessed sensitivity to unmeasured confounding [24]. Multiple-imputation sensitivity used 30 chained-equation datasets, 20 iterations and five predictive mean matching donors. Person–wave core records were imputed once and reused across entries, using current/earlier covariates and observed subsequent ADL/vital-status categories; later covariates were excluded. Cohort–wave fields with <30 observed donors remained missing; exposures, eligibility and outcomes were unchanged [25, 26]. All weighting models were refitted; Rubin pooling and Barnard–Rubin degrees of freedom combined estimates [27]. Post hoc age, sex and cardiovascular-history interactions on RR and RD scales used Holm correction across 18 tests.

Post hoc prognosis analyses used observation-weighted pooled logistic models with separate interval intercepts, baseline age, sex, self-rated health, IADL limitation, BMI and seven chronic-disease indicators. Age used a three-degree-of-freedom spline and a ridge penalty of 1 stabilised coefficients; activity assignment and future clinical predictors were excluded. Five person-level folds produced out-of-fold natural-course risk predictions. Cohort-specific overlap-weighted medians defined higher/lower-risk groups; calibration compared predictions with observation-corrected risks in fifths. Relative and absolute contrasts were standardised over common age-by-risk strata [12, 28]. Five hundred whole-person bootstraps refitted weighting and prediction models and reassigned risk groups. Twelve risk-group and allocation tests formed one Holm family.

Post hoc population-impact analyses compared hypothetical strategy achievement with the natural activity course [29, 30]. Natural-course risks retained later activity without adherence censoring, using observation weights and the propensity tilting function $e(1-e)$. Natural-course and assigned-strategy risks were standardised over common age/cardiovascular-history strata. At coverage q, policy risk mixed assigned-strategy and natural-course risks in proportions q and $1-q$. Uniform coverage used the same q across strata; priority allocation covered the selected age or disease category first at equal overall coverage. Potential impact fractions divided absolute risk reductions by natural-course risk. One thousand whole-person bootstrap refits recomputed models, target shares and allocation estimates; 18 allocation contrasts used Holm correction. Risk-priority comparisons at 50% coverage instead used common age-by-risk strata and the 500 refits described above. Programme reach multiplied by strategy achievement among those reached defined achieved coverage; each took values of 25%, 50% or 75%. These scenarios assumed achieved strategy delivery without spillover and estimated hypothetical impact, rather than delivered-programme effectiveness or costs. Sensitivity analyses varied probability bounds or retained the earliest entry with 300 bootstrap refits.

Supplementary baseline-classification analyses allowed later activity to vary. Mortality analyses used public death dates and a 48-month horizon, excluding incompatible chronology, undated deaths or unusable follow-up (Table S25). HRS exit-interview death month was preferred; public NDI NMONTH represented a quarter. SHARE dates had month precision and KLoSA dates had day/month precision. Weighted Cox models were stratified by baseline wave and combined overlap, initial-observation and monthly censoring weights. Censoring survival used a ridge-penalised Cox model (penalty 1), bounded below at 0.025. Three hundred

whole-person full-refit bootstraps estimated HR uncertainty and contrasted months 0–24 with 24–48; a sensitivity model omitted censoring weights. These HRs address baseline classification rather than the sustained strategy.

Use of artificial intelligence

During manuscript preparation, the authors used ChatGPT (OpenAI) for literature identification, manuscript drafting and English-language editing, and assistance with analysis and plotting code. Statistical analyses and figures were produced using R and Python. The authors reviewed and verified all AI-assisted text and take full responsibility for the manuscript.

Results

Study population

There were 19,204 participants and 43,754 eligible entries: 9,448/21,003 in HRS, 4,262/5,437 in SHARE and 5,494/17,314 in KLoSA (Figure 1). Before weighting, activity increasers generally had more favourable health and function (Table 1). Maximum absolute baseline standardised differences after overlap weighting were 0.002–0.003. Median intervals to second follow-up were 4.08, 4.58 and 4.00 years. Among observed first-follow-up survivors without an event, higher frequency was retained by 50.2%, 66.3% and 52.5%, respectively. Full selection, missingness, maintenance and weighting results are in Additional file 1.

Higher baseline activity accounted for 32.0%, 48.3% and 18.6% of entries in HRS, SHARE and KLoSA. After first-visit observation correction, maximum absolute standardised differences were 0.012, 0.092 and 0.011, respectively.

Primary outcome

At second follow-up, composite RRs were HRS 0.776 (95% CI 0.728–0.827; P<0.001), SHARE 0.764 (0.672–0.868; P<0.001) and KLoSA 0.725 (0.643–0.816; P<0.001). RDs were −7.08, −8.15 and −4.89 percentage points, respectively (Table 2). Differences were larger at second than first follow-up (Figure 2). Visit-specific estimates and complete P values are in Additional file 1, Supplementary tables.

First-follow-up composite risks were 14.2% versus 15.9% in HRS (RR 0.89; P=0.002), 15.7% versus 17.5% in SHARE (RR 0.90; P=0.146), and 6.7% versus 8.4% in KLoSA (RR 0.80; P=0.002).

Secondary outcomes

ADL-first RRs were 0.88 in HRS (P=0.003), 0.80 in SHARE (P=0.013) and 0.87 in KLoSA (95% CI 0.73–1.04; P=0.126). Independent all-cause mortality RRs were 0.61, 0.66 and 0.59 (all P<0.001; Table 2). The death-first component contributed 68.7%, 49.4% and 78.9% of the composite RD arithmetically (Figure 2).

Subgroup analyses

Older participants and those with heart disease/stroke had larger absolute point estimates, but none of the 18 interactions by age, sex or cardiovascular history met the Holm threshold (Additional file 1).

Risk stratification

Higher-versus-lower-risk absolute benefit estimates were 9.6 versus 2.8 events per 100 entries in HRS, 11.6 versus 3.8 in SHARE and 6.7 versus 2.3 in KLoSA (Figure 3). Between-group differences remained supported after Holm correction in HRS (P=0.001), but not SHARE (P=0.759) or KLoSA (P=0.133). No relative-scale contrast met the adjusted threshold. Profiles, calibration and all contrasts are in Additional file 1.

Conditional baseline standardised differences were below 0.10 in both risk groups. Calibration of out-of-fold predictions is shown in Figure S5 (Additional file 1, Supplementary figures).

Population impact

At 50% uniform achieved coverage, potential impact fractions relative to natural activity were 7.7%, 5.9% and 11.0%. At equal coverage, risk priority yielded 14.2 additional events avoided per 1,000 entries in HRS versus uniform allocation (Holm P<0.001); corresponding contrasts were uncertain in SHARE and KLoSA. Risk priority did not clearly outperform age priority in any cohort (Additional file 1, Supplementary tables). With 50% programme reach and 50% achievement, modelled reductions were 11.3, 8.9 and 9.2 events per 1,000 entries (Figure 3).

Sensitivity analyses

Composite RRs remained below 1 in every same-strategy sensitivity analysis. Multiple-imputation estimates were HRS 0.78 (95% CI 0.73–0.83), SHARE 0.76 (0.67–0.86) and KLoSA 0.72 (0.64–0.81; all P<0.001). Bootstrap intervals supported the primary direction; E-values were 1.90–2.10 (Additional file 1). Baseline-classification composite RRs were 0.90, 0.89 and 0.90. Separate mortality HRs were 0.808 (0.741–0.882; P<0.001), 0.855 (0.710–1.029; P=0.097) and 0.883 (0.779–1.001; P=0.052). Full additional results are in Additional file 1.

The direction was retained after excluding recent health deterioration and restricting to the earliest eligible entry or pre-2020 outcomes. In the supplementary mortality analysis, the KLoSA contrast weakened between the first and second two-year periods (interaction P=0.036).

Discussion

Among previously less active adults aged 65 years or older who were independent in basic ADLs, increased and maintained activity frequency was associated with lower risk of a first recorded ADL problem or death across cohorts from the USA, Europe and South Korea. Risk ratios ranged from 0.725 to 0.776, corresponding to 4.9–8.2 fewer events per 100 eligible trial entries over two subsequent visits. Mortality contributed substantially to the differences, whereas evidence for ADL problems alone was less consistent, particularly in KLoSA. The primary direction persisted in sensitivity analyses. For health promotion, these findings identify a relevant behavioural strategy and quantify the associated absolute differences; the programme scenarios further show how modelled impact depends on reach and successful strategy achievement.

Randomised trial evidence is strongest for mobility and physical performance. LIFE reduced major mobility disability, defined by inability to walk 400 m [4], while SPRINTT demonstrated benefit in physically frail, sarcopenic participants with Short Physical Performance Battery scores of 3–7 [5]. REACT combined community exercise with behavioural maintenance and helped preserve lower limb function [6]. Mobility and self-care represent related but distinct outcomes: a LIFE secondary analysis did not demonstrate reductions in basic or instrumental ADL disability [7]. Lower limb performance predicts subsequent disability before basic ADL problems emerge [31]. An earlier Australian emulation examined adherence to activity guidelines from midlife and later health-related quality of life [11]. Our study adds a later-life comparison among previously

less active adults, incorporating assessed maintenance and reporting relative and absolute risks of daily living problems or death. These outcomes connect physical activity research to the healthy ageing goals of independent living and survival.

The survival findings are consistent with EPIC-Norfolk and a recent meta-analysis linking increasing or consistently higher activity with lower mortality [32, 33]. Life-course analyses based on recalled activity histories also associated increases during adulthood with lower mortality [34], while accelerometry-based evidence supports the broader activity–survival association [35]. Related HRS and Taiwanese trajectory studies reported more favourable functional or survival outcomes with increasing activity [9, 10]. Those analyses summarised longer activity histories; our comparison concerned an increase between adjacent visits followed by assessed maintenance. Reporting the outcome components clarified the contribution of survival to the overall difference. KLoSA assessed need for assistance, a more advanced functional threshold than reported difficulty, and its ADL estimate remained uncertain. Component allocation depended on the first recorded event; death between interviews could follow an unobserved ADL problem. Public health evaluation should therefore report functional and survival outcomes separately alongside their composite.

Baseline prognosis helps translate relative associations into absolute differences relevant to programme planning. Similar relative associations can correspond to larger absolute differences among people with higher underlying risk, as described by PATH guidance and its extension to observational databases [12, 28]. Higher-risk point estimates were larger in all three cohorts, but between-group differences remained supported after multiplicity adjustment in HRS and uncertain in SHARE and KLoSA. Age, sex and cardiovascular-history interactions were not supported after adjustment. At equal achieved coverage, risk priority was more favourable than uniform allocation in HRS, but did not clearly outperform age priority in any cohort. These results support further evaluation of simple and more detailed allocation approaches using achieved coverage, absolute outcomes and delivery costs, rather than assuming that a more complex risk score will improve programme performance.

The implementation question is how to support activity changes that can be achieved and sustained. Activity goals can be tailored to ability and health status, consistent with current guidance [3]. PACE-Lift combined primary-care consultations, activity monitoring and feedback and increased objectively measured activity

[36]. Some objectively measured gains in the PACE trials persisted for three to four years [37]. A broader primary-care meta-analysis found less consistent device-measured effects [38], while the ExWell cohort illustrates the practical challenges of participation and retention in community exercise [39]. Our scenarios distinguish programme reach from strategy achievement among those reached: with each at 50%, modelled reductions relative to the natural activity course were approximately 9–11 events per 1000 eligible entries. This calculation makes implementation assumptions explicit and provides scenarios for planning prospective evaluations. A pragmatic community or primary-care trial could test sustained activity support, measuring repeated activity, functional and survival outcomes, reach, adherence, staff time and costs.

Strengths include independent analyses in three ageing cohorts, explicit strategies, retention of early outcomes before adherence updating, and adjustment for measured adherence and observation processes, consistent with target trial principles [16, 19]. This study has several limitations. Activity change was already observed at baseline, and eligibility required survival and ADL independence up to that visit; the contrast therefore concerns eligible survivors, with an unknown change date. Self-reported frequency did not establish weekly volume or continuous adherence, and cohort-specific measurements, eligibility criteria and interview intervals limited comparability. Residual confounding by underlying illness, functional reserve or motivation remains possible; evidence that declining function predicts lower subsequent activity reinforces this concern [40]. Same-visit health covariates may partly reflect earlier activity, complicating temporal ordering. Weighting requires adequate overlap and correctly specified models for activity, adherence and observation. The main analysis did not propagate imputation uncertainty, and some covariates were structurally unavailable despite supportive multiple-imputation sensitivity analyses. Covariate balance and E-values do not verify absence of confounding or other biases [24]. Overlap weighting targets participants with comparable strategy probabilities, and survey sampling weights were not used; estimates are not national population effects [21]. Risk stratification and implementation scenarios were post hoc and require external and prospective evaluation before programme allocation or delivery. Finally, supplementary mortality HRs compared baseline activity with subsequent activity unrestricted, addressing a different question from the sustained strategy. These comparisons also cannot isolate the incremental effect of maintenance.

Conclusions

In previously less active older adults who were independent in basic ADLs, increased and maintained activity frequency was associated with lower risk of daily living problems or death across three ageing cohorts. The clearest common signal concerned survival. These findings support evaluating sustained activity support within healthy ageing programmes, with reach and strategy achievement assessed alongside functional and survival outcomes.

Abbreviations

ADL, activities of daily living; BMI, body mass index; CI, confidence interval; HRS, Health and Retirement Study; HR, hazard ratio; IADL, instrumental activities of daily living; KLoSA, Korean Longitudinal Study of Aging; NNB, number needed to benefit; RD, risk difference; RR, risk ratio; SD, standard deviation; SHARE, Survey of Health, Ageing and Retirement in Europe; STROBE, Strengthening the Reporting of Observational Studies in Epidemiology; TARGET, Transparent Reporting of Observational Studies Emulating a Target Trial.

Declarations

Ethics approval and consent to participate

This study used de-identified data from the Health and Retirement Study (HRS), the Survey of Health, Ageing and Retirement in Europe (SHARE), and the Korean Longitudinal Study of Aging (KLoSA). The original cohort studies obtained ethical approval and informed consent from participants. The Ethics Committee of The Second Xiangya Hospital of Central South University waived ethical review for the present secondary analysis.

Consent for publication

Not applicable.

Availability of data and materials

The HRS data are available from https://hrsdata.isr.umich.edu/. SHARE release 9.0.0 data are available from https://share-eric.eu/data/data-access. KLoSA data are available from

https://survey.keis.or.kr/eng/klosa/klosa01.jsp. Access is subject to each provider's applicable registration requirements and conditions of use. Participant-level data are not redistributed by the authors. Analysis code is available from the corresponding authors upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the CLEVER TAVR study under Grant CSCF2021B01; the project 'Research on Key Technologies for the Assessment of Valvular Heart Disease in Older Adults and the Construction of a Comprehensive Information Big Data Platform' under Grant 2020YFC2008101; and the Hunan Provincial Natural Science Foundation of China under Grant 2024JJ5501 (to Yiyuan Huang). The funders had no role in the study design; collection, analysis or interpretation of data; writing of the manuscript; or the decision to submit the results for publication.

Authors' contributions

ZC contributed to conceptualisation, methodology, formal analysis and writing the original draft. FY contributed to methodology, data curation and writing the original draft. YH contributed to funding acquisition and manuscript review and editing. FL contributed to validation and manuscript review and editing. TL contributed to conceptualisation, supervision and manuscript review and editing. ZF contributed to conceptualisation, supervision, project administration and manuscript review and editing. All authors read and approved the final manuscript.

Acknowledgements

We thank the participants and research teams of HRS at the University of Michigan, SHARE-ERIC, and KLoSA at the Korea Employment Information Service for making the cohort data available.

Supplementary information

Additional file 1 (.docx): Tables S1–S25 and Figures S1–S10, with accompanying notes. Includes the target trial protocol, cohort selection and measurement summaries, follow-up counts, risk estimates, subgroup and

sensitivity results, and programme-impact scenarios. STROBE and TARGET reporting checklists are supplied separately.

References

1. Graham L, Brundle C, Harrison N, Andre D, Clegg A, Forster A, et al. What are the priorities for research of older people living in their own home, including those living with frailty? A systematic review and content analysis of studies reporting older people's priorities and unmet needs. Age Ageing. 2024;53:afad232. https://doi.org/10.1093/ageing/afad232.

2. Beard JR, Officer A, de Carvalho IA, Sadana R, Pot AM, Michel JP, et al. The World report on ageing and health: a policy framework for healthy ageing. Lancet. 2016;387:2145-2154. https://doi.org/10.1016/S0140-6736(15)00516-4.

3. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med. 2020;54:1451-1462. https://doi.org/10.1136/bjsports-2020-102955.

4. Pahor M, Guralnik JM, Ambrosius WT, Blair S, Bonds DE, Church TS, et al. Effect of structured physical activity on prevention of major mobility disability in older adults: the LIFE study randomized clinical trial. JAMA. 2014;311:2387-96. https://doi.org/10.1001/jama.2014.5616.

5. Bernabei R, Landi F, Calvani R, Cesari M, Del Signore S, Anker SD, et al. Multicomponent intervention to prevent mobility disability in frail older adults: randomised controlled trial (SPRINTT project). BMJ. 2022;377:e068788. https://doi.org/10.1136/bmj-2021-068788.

6. Stathi A, Greaves CJ, Thompson JL, Withall J, Ladlow P, Taylor G, et al. Effect of a physical activity and behaviour maintenance programme on functional mobility decline in older adults: the REACT (Retirement in Action) randomised controlled trial. Lancet Public Health. 2022;7:e316-e326. https://doi.org/10.1016/S2468-2667(22)00004-4.

7. Manini TM, Beavers DP, Pahor M, Guralnik JM, Spring B, Church TS, et al. Effect of Physical Activity on Self-Reported Disability in Older Adults: Results from the LIFE Study. J Am Geriatr Soc. 2017;65:980-988. https://doi.org/10.1111/jgs.14742.

8. Tak E, Kuiper R, Chorus A, Hopman-Rock M. Prevention of onset and progression of basic ADL disability by physical activity in community dwelling older adults: a meta-analysis. Ageing Res Rev. 2013;12:329-38. https://doi.org/10.1016/j.arr.2012.10.001.

9. Vasilopoulos T, Drozda D, Vincent HK. Physical activity positively impacts disability outcomes during transition from midlife to early older age irrespective of body mass index. Arch Gerontol Geriatr. 2024;120:105339. https://doi.org/10.1016/j.archger.2024.105339.

10. Yu YY, Lai HY, Loh CH, Chen LK, Hsiao FY. Multi-trajectories of physical activity and health in aging: A 20-year nationwide population-based longitudinal study. Arch Gerontol Geriatr. 2025;135:105878. https://doi.org/10.1016/j.archger.2025.105878.

11. Nguyen B, Clare P, Mielke GI, Brown WJ, Ding D. Physical activity across midlife and health-related quality of life in Australian women: A target trial emulation using a longitudinal cohort. PLoS Med. 2024;21:e1004384. https://doi.org/10.1371/journal.pmed.1004384.

12. Kent DM, Paulus JK, van Klaveren D, D'Agostino R, Goodman S, Hayward R, et al. The Predictive Approaches to Treatment effect Heterogeneity (PATH) Statement. Ann Intern Med. 2020;172:35-45. https://doi.org/10.7326/M18-3667.

13. Sonnega A, Faul JD, Ofstedal MB, Langa KM, Phillips JW, Weir DR. Cohort Profile: the Health and Retirement Study (HRS). Int J Epidemiol. 2014;43:576-85. https://doi.org/10.1093/ije/dyu067.

14. Borsch-Supan A, Brandt M, Hunkler C, Kneip T, Korbmacher J, Malter F, et al. Data Resource Profile: the Survey of Health, Ageing and Retirement in Europe (SHARE). Int J Epidemiol. 2013;42:992-1001. https://doi.org/10.1093/ije/dyt088.

15. Korea Employment Information Service. Korean Longitudinal Study of Aging study introduction. https://survey.keis.or.kr/eng/klosa/klosa01.jsp. Accessed 16 Oct 2026.

16. Hernan MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol. 2016;183:758-64. https://doi.org/10.1093/aje/kwv254.

17. Cashin AG, Hansford HJ, Hernan MA, Swanson SA, Lee H, Jones MD, et al. Transparent Reporting of Observational Studies Emulating a Target Trial-The TARGET Statement. JAMA. 2025;334:1084-1093. https://doi.org/10.1001/jama.2025.13350.

18. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007;4:e296. https://doi.org/10.1371/journal.pmed.0040296.

19. Hernan MA, Sauer BC, Hernandez-Diaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. J Clin Epidemiol. 2016;79:70-75. https://doi.org/10.1016/j.jclinepi.2016.04.014.

20. Li F, Morgan KL, Zaslavsky AM. Balancing covariates via propensity score weighting. J Am Stat Assoc. 2018;113:390–400. https://doi.org/10.1080/01621459.2016.1260466.

21. Li F, Thomas LE, Li F. Addressing Extreme Propensity Scores via the Overlap Weights. Am J Epidemiol. 2019;188:250-257. https://doi.org/10.1093/aje/kwy201.

22. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. Am J Epidemiol. 2008;168:656-64. https://doi.org/10.1093/aje/kwn164.

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

24. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. Ann Intern Med. 2017;167:268-274. https://doi.org/10.7326/M16-2607.

25. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011;30:377-99. https://doi.org/10.1002/sim.4067.

26. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. J Stat Softw. 2011;45:1–67. https://doi.org/10.18637/jss.v045.i03.

27. Barnard J, Rubin DB. Miscellanea. Small-sample degrees of freedom with multiple imputation. Biometrika. 1999;86:948–55. https://doi.org/10.1093/biomet/86.4.948.

413 28. Rekkas A, van Klaveren D, Ryan PB, Steyerberg EW, Kent DM, Rijnbeek PR. A standardized framework
414 for risk-based assessment of treatment effect heterogeneity in observational healthcare databases. NPJ Digit
415 Med. 2023;6:58. https://doi.org/10.1038/s41746-023-00794-y.

416 29. Westreich D. From exposures to population interventions: pregnancy and response to HIV therapy. Am J
417 Epidemiol. 2014;179:797-806. https://doi.org/10.1093/aje/kwt328.

418 30. Rudolph JE, Cartus A, Bodnar LM, Schisterman EF, Naimi AI. The Role of the Natural Course in Causal
419 Analysis. Am J Epidemiol. 2022;191:341-348. https://doi.org/10.1093/aje/kwab248.

420 31. Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-extremity function in persons
421 over the age of 70 years as a predictor of subsequent disability. N Engl J Med. 1995;332:556-61.
422 https://doi.org/10.1056/NEJM199503023320902.

423 32. Mok A, Khaw KT, Luben R, Wareham N, Brage S. Physical activity trajectories and mortality: population
424 based cohort study. BMJ. 2019;365:l2323. https://doi.org/10.1136/bmj.l2323.

425 33. Yu R, Duncombe SL, Nemoto Y, Araujo RH, Chung HF, Mielke GI. Physical activity trajectories and
426 accumulation over adulthood and their associations with all-cause and cause-specific mortality: a systematic
427 review and meta-analysis. Br J Sports Med. 2025;59:1228-1241. https://doi.org/10.1136/bjsports-2024-
428 109122.

429 34. Saint-Maurice PF, Coughlan D, Kelly SP, Keadle SK, Cook MB, Carlson SA, et al. Association of
430 Leisure-Time Physical Activity Across the Adult Life Course With All-Cause and Cause-Specific Mortality.
431 JAMA Netw Open. 2019;2:e190355. https://doi.org/10.1001/jamanetworkopen.2019.0355.

432 35. Ekelund U, Tarp J, Steene-Johannessen J, Hansen BH, Jefferis B, Fagerland MW, et al. Dose-response
433 associations between accelerometry measured physical activity and sedentary time and all cause mortality:
434 systematic review and harmonised meta-analysis. BMJ. 2019;366:l4570. https://doi.org/10.1136/bmj.l4570.

435 36. Harris T, Kerry SM, Victor CR, Ekelund U, Woodcock A, Iliffe S, et al. A primary care nurse-delivered
436 walking intervention in older adults: PACE (pedometer accelerometer consultation evaluation)-Lift cluster
437 randomised controlled trial. PLoS Med. 2015;12:e1001783. https://doi.org/10.1371/journal.pmed.1001783.

438 37. Harris T, Kerry SM, Limb ES, Furness C, Wahlich C, Victor CR, et al. Physical activity levels in adults
439 and older adults 3-4 years after pedometer-based walking interventions: Long-term follow-up of participants
440 from two randomised controlled trials in UK primary care. PLoS Med. 2018;15:e1002526.
441 https://doi.org/10.1371/journal.pmed.1002526.

442 38. Kettle VE, Madigan CD, Coombe A, Graham H, Thomas JJC, Chalkley AE, et al. Effectiveness of
443 physical activity interventions delivered or prompted by health professionals in primary care settings:
444 systematic review and meta-analysis of randomised controlled trials. BMJ. 2022;376:e068465.
445 https://doi.org/10.1136/bmj-2021-068465.

446 39. O'Shea M, Doody P, McCaffrey N, O'Deashunaigh A, Skelly F, Kelly L, et al. Community-based,
447 supervised, group exercise for people living with chronic conditions and multimorbidity: a cohort study. Age
448 Ageing. 2026;55:afag174. https://doi.org/10.1093/ageing/afag174.

449 40. Greenwood-Hickman MA, Zhu W, Idu A, Harrington LB, McCurry SM, LaCroix AZ, et al. Associations
450 Between 10-Year Physical Performance and Activities of Daily Living Trajectories and Physical Behaviors in
451 Older Adults. Int J Environ Res Public Health. 2025;22:704. https://doi.org/10.3390/ijerph22050704.

Table 1. Unweighted baseline characteristics of eligible trial entries

Characteristic	HRS			SHARE			KLoSA		
	Increase	Lower	P	Increase	Lower	P	Increase	Lower	P
Trial entries, n	6716	14287	—	2625	2812	—	3228	14086	—
Age, mean (SD), years	74.2 (6.7)	75.2 (7.0)	2.33e-21	73.7 (6.4)	75.3 (7.1)	1.37e-16	73.5 (6.2)	74.9 (6.8)	1.89e-27
Women	3836 (57.1)	8941 (62.6)	4.53e-11	1489 (56.7)	1710 (60.8)	0.003	1735 (53.7)	9005 (63.9)	5.04e-21
Education: below upper secondary	1245 (18.5)	3242 (22.7)		1305 (49.7)	1670 (59.4)		2209 (68.4)	11813 (83.9)	
Upper secondary / vocational	4009 (59.7)	8868 (62.1)	7.80e-26	840 (32.0)	752 (26.7)	3.21e-11	739 (22.9)	1817 (12.9)	2.84e-62
Tertiary education	1459 (21.7)	2170 (15.2)		480 (18.3)	390 (13.9)		280 (8.7)	456 (3.2)	
BMI, mean (SD), kg/m²	27.6 (5.0)	28.4 (5.7)	3.88e-17	27.0 (4.4)	27.4 (4.7)	0.001	23.4 (2.8)	23.0 (2.9)	3.13e-11
Married / partnered	4137 (61.6)	8198 (57.4)	2.68e-07	1868 (71.3)	1932 (69.0)	0.080	2303 (71.3)	9021 (64.0)	5.57e-13
Current smoking	585 (8.8)	1595 (11.2)	1.47e-06	219 (14.0)	256 (15.2)	0.341	330 (10.2)	1668 (11.8)	0.016
Alcohol consumption	3539 (52.7)	6264 (43.9)	4.63e-27	1426 (54.3)	1165 (41.4)	1.74e-20	888 (27.5)	3577 (25.4)	0.026
Fair / poor self-rated health	1281 (19.1)	4047 (28.4)	5.59e-43	907 (34.6)	1490 (53.0)	7.93e-41	2057 (63.7)	10394 (73.8)	1.52e-25
Walking difficulty	360 (5.4)	2103 (14.8)	2.53e-94	147 (5.6)	529 (18.8)	3.20e-46	—	—	—
Any IADL limitation	417 (6.2)	1316 (9.2)	5.97e-14	76 (2.9)	290 (10.3)	4.90e-27	244 (7.6)	1415 (10.0)	6.76e-06
Hypertension	4412 (65.7)	10081 (70.6)	3.47e-10	1697 (64.6)	1989 (70.8)	2.42e-06	1685 (52.2)	6997 (49.7)	0.019
Diabetes	1578 (23.5)	4108 (28.8)	7.55e-13	530 (20.2)	678 (24.1)	9.43e-04	753 (23.3)	2980 (21.2)	0.015
Heart disease	1809 (26.9)	4320 (30.2)	1.18e-05	672 (25.6)	845 (30.1)	4.43e-04	356 (11.0)	1568 (11.1)	0.876
Stroke	510 (7.6)	1298 (9.1)	0.001	193 (7.4)	284 (10.1)	0.001	221 (6.8)	796 (5.7)	0.023
Cancer	1202 (17.9)	2743 (19.2)	0.045	293 (11.2)	377 (13.4)	0.018	231 (7.2)	754 (5.4)	7.03e-04
Chronic lung disease	591 (8.8)	1724 (12.1)	8.66e-11	322 (12.3)	458 (16.3)	7.48e-05	125 (3.9)	629 (4.5)	0.180
Arthritis	4266 (63.5)	9752 (68.3)	3.01e-09	1407 (53.6)	1695 (60.3)	1.46e-06	977 (30.3)	5318 (37.8)	4.00e-13
Hospital admission	1735 (25.9)	4086 (28.7)	3.86e-05	379 (14.4)	455 (16.2)	0.078	380 (11.8)	1855 (13.2)	0.031

Values are n (%) among known entries or mean (SD). P values use participant-clustered Wald tests; education uses one overall test. Increase and Lower denote baseline activity groups after lower prior frequency. Repeated entries were permitted. Education is harmonised for display. Weighted balance and missingness are in Figure S2 and Table S3.

Table 2. Estimated risks and differences for the activity strategies

Outcome	Cohort	Strategy risks, % Increase / Lower	RR (95% CI)	RD, pp (95% CI)	Strategy NNB (95% CI)	P for RR
ADL problem or death	HRS	24.5 / 31.6	0.78 (0.73, 0.83)	−7.08 (−8.76, −5.39)	14.1 (11.4, 18.6)	4.61e-15
	SHARE	26.3 / 34.5	0.76 (0.67, 0.87)	−8.15 (−12.01, −4.29)	12.3 (8.3, 23.3)	3.49e-05
	KLoSA	12.9 / 17.8	0.72 (0.64, 0.82)	−4.89 (−6.53, −3.25)	20.4 (15.3, 30.7)	1.12e-07
ADL first recorded	HRS	15.7 / 17.9	0.88 (0.80, 0.96)	−2.22 (−3.67, −0.76)	45.1 (27.2, 131.2)	0.003
	SHARE	16.4 / 20.5	0.80 (0.67, 0.95)	−4.12 (−7.38, −0.87)	24.3 (13.6, 115.5)	0.013
	KLoSA	6.9 / 8.0	0.87 (0.73, 1.04)	−1.03 (−2.30, 0.24)	96.8*	0.126
All-cause mortality	HRS	9.8 / 16.0	0.61 (0.55, 0.68)	−6.18 (−7.43, −4.94)	16.2 (13.5, 20.2)	1.16e-18
	SHARE	10.8 / 16.4	0.66 (0.53, 0.83)	−5.57 (−8.53, −2.61)	18.0 (11.7, 38.3)	3.27e-04
	KLoSA	6.5 / 11.0	0.59 (0.50, 0.70)	−4.52 (−5.79, −3.26)	22.1 (17.3, 30.7)	2.01e-09

NNB is the reciprocal of the absolute risk reduction; * indicates an RD interval crossing zero, with no finite benefit interval. P values test RR=1. ADL denotes difficulty in HRS/SHARE and assistance in KLoSA. Independent mortality includes deaths after ADL problems. Death-first estimates are in Table S8.

Figure 1. Trial design and cohort selection

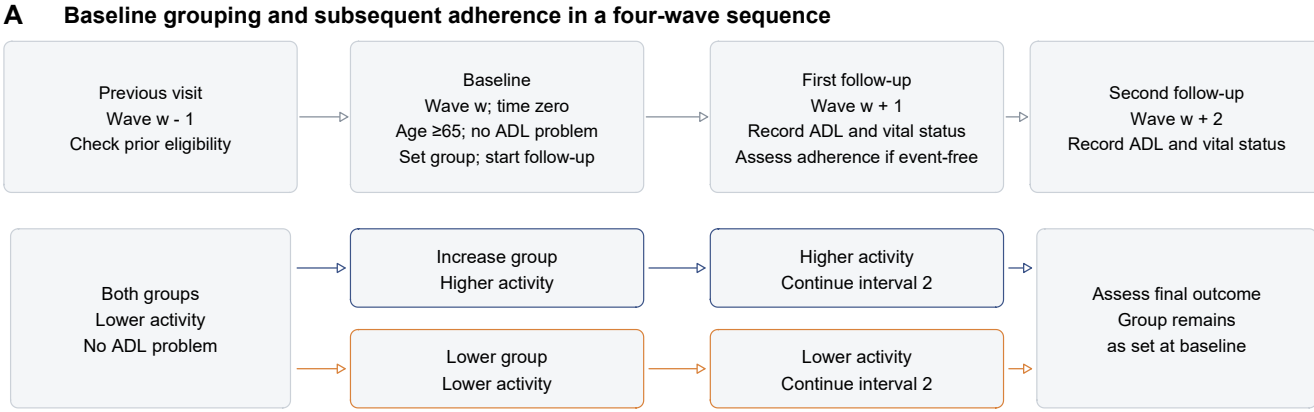
A, Four-visit trial sequence. Baseline activity defined the groups; first-follow-up outcomes were counted before adherence was assessed among event-free survivors. B–D, Cohort selection, allowing repeated eligible entries.

Figure 2. Composite risks and first recorded components

A, C, E, Weighted cumulative composite risks; lines connect survey estimates. B, D, F, ADL-first and death-first components of the second-follow-up absolute risk difference. Error bars indicate 95% confidence intervals. ADL denotes difficulty in HRS/SHARE and assistance in KLoSA.

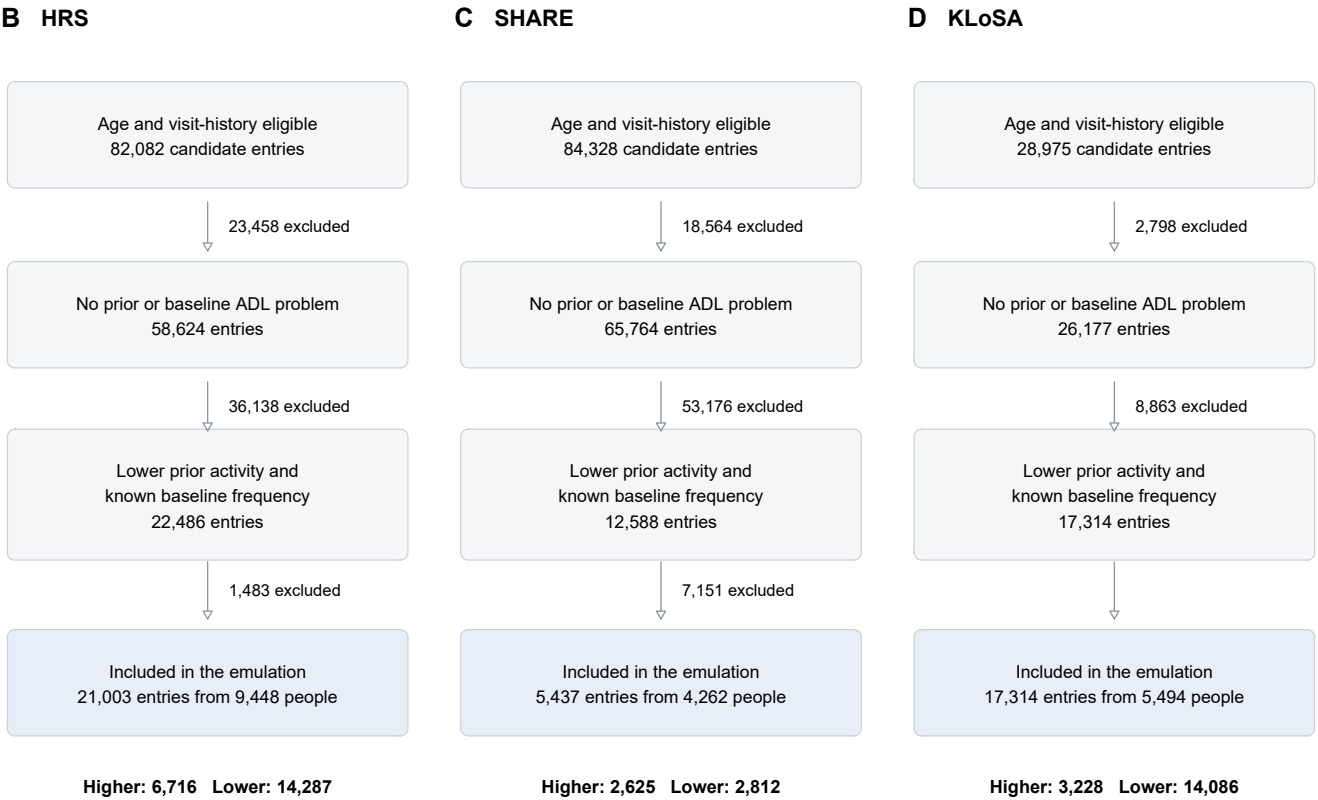
Figure 3. Baseline prognosis and programme coverage

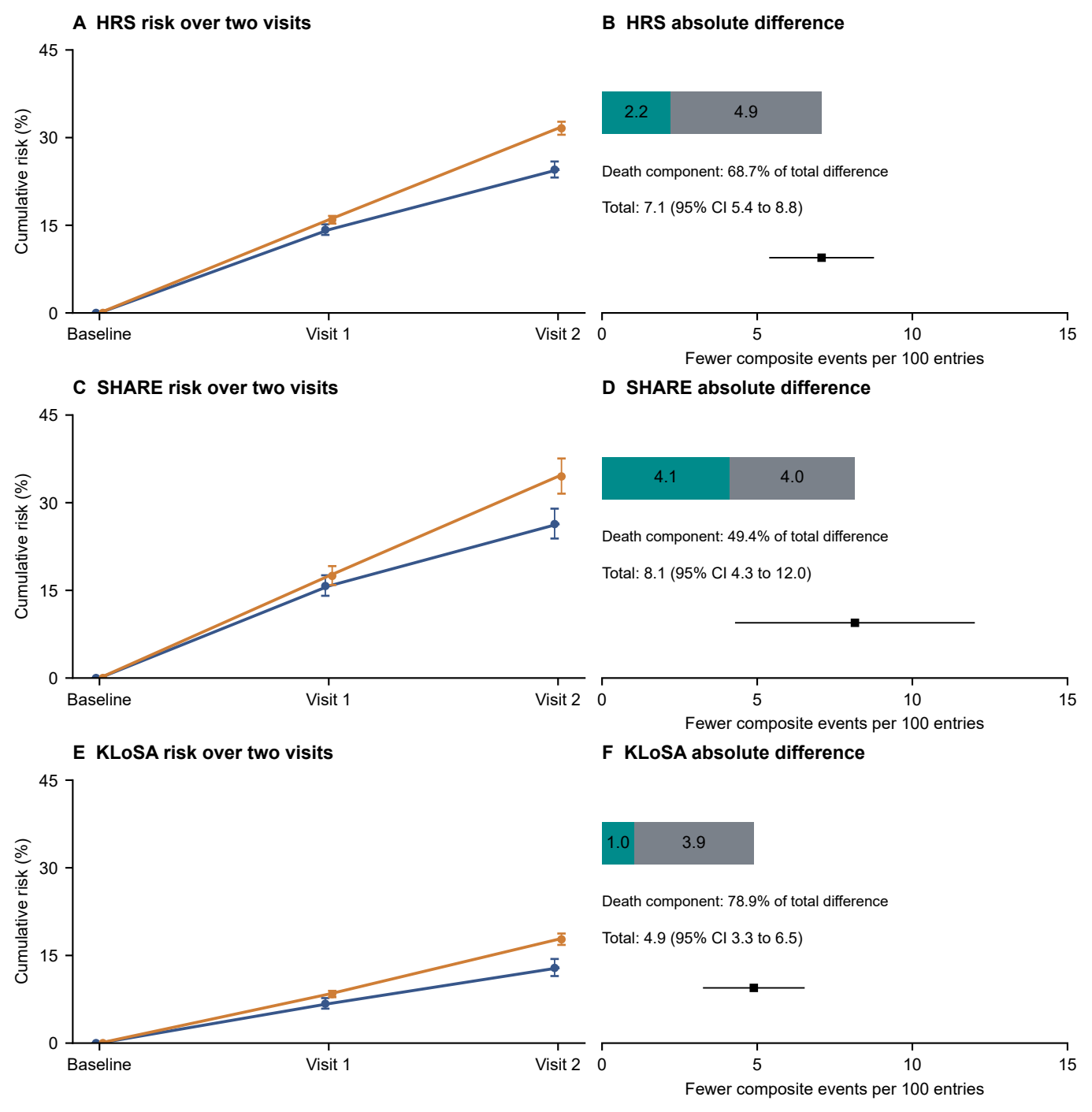
A–C, Absolute risk reductions by baseline predicted-risk group, with 95% bootstrap confidence intervals and Holm-adjusted between-group P values. D–F, Modelled events avoided per 1,000 entries relative to the natural activity course. Achieved coverage equals programme reach × strategy achievement.

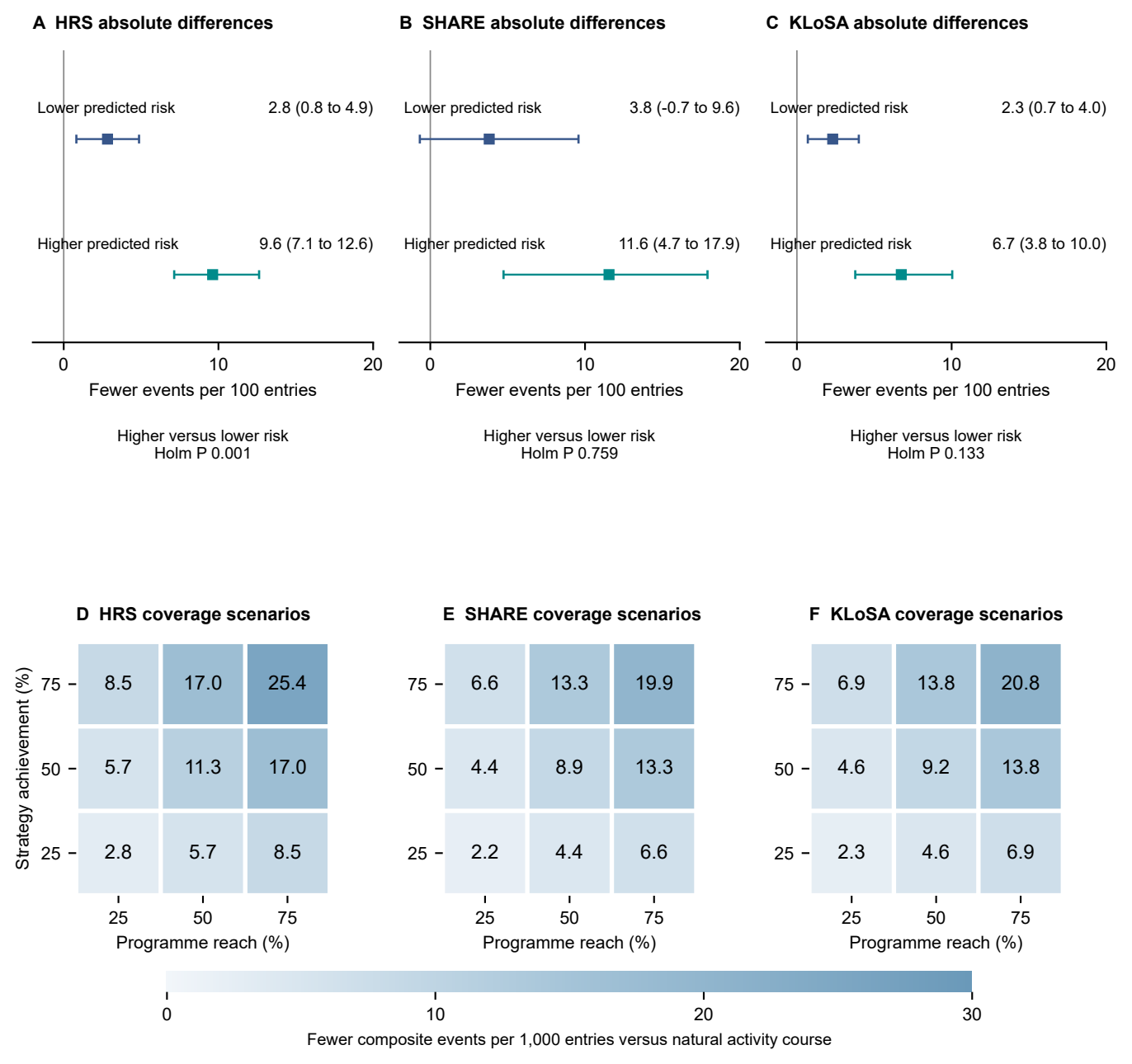


At wave w + 1: an ADL event or death is counted in the baseline group; no adherence check is required.

If event-free: changed or unknown activity stops interval 2 contribution; adherence weighting is applied.




A HRS risk over two visits
B HRS absolute difference
Cumulative risk (%)
45
30
15
0
Baseline
Visit 1
Visit 2
2.2
4.9
Death component: 68.7% of total difference
Total: 7.1 (95% CI 5.4 to 8.8)
0
5
10
15
Fewer composite events per 100 entries
C SHARE risk over two visits
D SHARE absolute difference
4.1
4.0
Death component: 49.4% of total difference
Total: 8.1 (95% CI 4.3 to 12.0)
E KLoSA risk over two visits
F KLoSA absolute difference
1.0
3.9
Death component: 78.9% of total difference
Total: 4.9 (95% CI 3.3 to 6.5)
Increase and maintain
Remain lower
ADL first
Death first


A HRS absolute differences
Lower predicted risk
2.8 (0.8 to 4.9)
Higher predicted risk
9.6 (7.1 to 12.6)
0
10
20
Fewer events per 100 entries
Higher versus lower risk
Holm P 0.001
B SHARE absolute differences
Lower predicted risk
3.8 (-0.7 to 9.6)
Higher predicted risk
11.6 (4.7 to 17.9)
0
10
20
Fewer events per 100 entries
Higher versus lower risk
Holm P 0.759
C KLoSA absolute differences
Lower predicted risk
2.3 (0.7 to 4.0)
Higher predicted risk
6.7 (3.8 to 10.0)
0
10
20
Fewer events per 100 entries
Higher versus lower risk
Holm P 0.133
D HRS coverage scenarios
Strategy achievement (%)
75
50
25
8.5
17.0
25.4
5.7
11.3
17.0
2.8
5.7
8.5
25
50
75
Programme reach (%)
E SHARE coverage scenarios
75
50
25
6.6
13.3
19.9
4.4
8.9
13.3
2.2
4.4
6.6
25
50
75
Programme reach (%)
F KLoSA coverage scenarios
75
50
25
6.9
13.8
20.8
4.6
9.2
13.8
2.3
4.6
6.9
25
50
75
Programme reach (%)
0
10
20
30
Fewer composite events per 1,000 entries versus natural activity course

Supplementary Files

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

- Additionalfile1Supplementarymaterial.docx