# Physical activity across midlife and health-related quality of life in Australian women: A target trial emulation using a longitudinal cohort

Binh Nguyen    , Philip Clare, Gregore I. Mielke, Wendy J. Brown, Ding Ding

## Abstract

**Background**

There is little long-term causal evidence on the effect of physical activity on health-related quality of life. This study aimed to examine the associations between longitudinal patterns of physical activity over 15 years and health-related quality of life in both the physical and mental health domains, in a cohort of middle-aged Australian women.

**Methods and findings**

We used data collected at 3-year intervals (1998 to 2019) from 11,336 participants in the Australian Longitudinal Study on Women's Health (ALSWH) (1946 to 1951 birth cohort). Primary outcomes were the physical (PCS) and mental health component summary (MCS) scores (range from 0 to 100; higher scores indicate higher perceived physical/mental health) from the SF-36 in 2019 (when women aged 68 to 73 years). Using target trial emulation to imitate a randomized controlled trial (RCT), we tested 2 interventions: (1) meeting the World Health Organization (WHO) physical activity guidelines consistently throughout the 15-year "exposure period" (2001 to 2016; when women aged 50–55 to 65–70 years; physical activity assessed every 3 years); and (2) not meeting the guidelines at the beginning of the exposure period but starting to first meet the guidelines at age 55, 60, or 65; against the control of not meeting the guidelines throughout the exposure period. Analysis controlled for confounding using marginal structural models which were adjusted for sociodemographic and health variables and conditions. Consistent adherence to guidelines during the exposure period (PCS: 46.93 [99.5% confidence interval [CI]: 46.32, 47.54]) and first starting to meet the guidelines at age 55 (PCS: 46.96 [99.5% CI: 45.53, 48.40]) were associated with three-point higher PCS (mean score difference: 3.0 [99.5% CI: 1.8, 4.1] and 3.0 [99.5% CI:1.2, 4.8]) than consistent non-adherence (PCS: 43.90 [99.5% CI: 42.79, 45.01]). We found a similar pattern for most SF-36 subscales but no significant effects of the interventions on MCS. The main limitations of the study were that it may not account for all underlying health conditions and/or other unmeasured or insufficiently measured confounders, the use of self-reported physical activity and that findings may not be generalizable to all mid-age women.

**Conclusions**

Results from the emulated RCT suggest women should be active throughout mid-age, ideally increasing activity levels to meet the guidelines by age 55, to gain the most benefits for physical health in later life.

## Author summary

**Why was this study done?**

○ The evidence for an association between physical activity and health-related quality of life has been based primarily on cross-sectional studies and randomized controlled trials (RCTs) with short-term outcomes.
　○ To date, few longitudinal studies have examined the associations between long-term patterns of physical activity and health-related quality of life, measured physical activity at more than one time point, and examined the causal effects of physical activity.

**What did the researchers do and find?**

○ We leveraged data on >11,000 women enrolled in the Australian Longitudinal Study on Women's Health (ALSWH) and applied targeted maximum likelihood estimation (TMLE) to emulate various target trials of physical adherence regimes and explore how physical activity in midlife impacts health outcomes in later life.
　○ We found that consistent adherence to physical activity guidelines over 15 years was associated with better health-related quality of life in the physical health domain, especially physical functioning.

　○ Meeting physical activity guidelines earlier in midlife and especially by age 55 resulted in better physical health-related quality of life outcomes in later life but the same impact on mental health domain outcomes was less evident.

**What do these findings mean?**

○ This study provides insights into the long-term effects of different patterns of physical activity during mid-age on later health-related quality of life in Australian women.

○ Combined with existing evidence, it contributes to the growing evidence concerning the benefits of maintaining or adopting an active lifestyle in mid-age.

○ The study presents an important message for public health which encourages women to be active throughout mid-age, even if they start meeting physical activity guidelines in their mid-50s, to gain the most benefits for physical health-related quality of life latterly.

○ Limitations of the study included reliance on self-reported physical activity, an inability to account for all contributing factors and findings may not be generalizable to all mid-age Australian women or to women outside of Australia.

**Data Availability:** The data underlying this study are owned by the Australian Government Department of Health (ALSWH). The process for data access is documented on the Australian Longitudinal Study on Women's Health website [http://www.alswh.org.au] which includes all the survey questionnaires, data books of frequency tables for all surveys, meta-data, conditions of data access and request form.

**Abbreviations:** ALSWH, Australian Longitudinal Study on Women's Health; ARIA+, Accessibility-Remoteness Index of Australia Plus; BMI, body mass index; CES-D, Center for Epidemiological Studies-Depression; CI, confidence interval; DAG, directed acyclic graph; IRSD, Index of Relative Socio-economic Disadvantage; MCS, mental health component summary; PCS, physical health component summary; RCT, randomized controlled tria; SD, standard deviation; TMLE, targeted maximum likelihood estimation; WHO, World Health Organization

## Introduction

The health benefits of physical activity are well established [1–4]. The World Health Organization (WHO) recommends 150 to 300 min of moderate-intensity aerobic physical activity or 75 to 150 min of vigorous activity per week (or an equivalent combination of both) for health benefits [2]. Although maintaining an active lifestyle is optimal for health, levels of physical activity may change across the lifespan. For women, major transitions in life, such as marriage, childbirth, children leaving home, retirement, and bereavement, can be associated with changes in physical activity [5]. To date, most studies on physical activity and health have relied on physical activity measured at one point in time, not accounting for changes in activity patterns over time. Relying on a single static measure not only introduces measurement biases but also fails to answer an important question "what would happen if one starts to be physically active in later life?" The health outcomes of longitudinal patterns of meeting physical activity guidelines in women have seldom been explored [6–9].

Physical activity is important for healthy aging, defined by the WHO as "the process of developing and maintaining the functional ability that enables wellbeing in older age" [10]. An important measure of healthy aging is health-related quality of life, a multidimensional concept that captures perceived functioning and wellbeing in the physical, mental, and social domains of health [11]. Health-related quality of life is a strong predictor of mortality risk [12] and is widely used as a population health indicator [13]. Evidence suggests a link between physical activity and health-related quality of life in adults, but this evidence is primarily based on cross-sectional studies, which have inherent biases, and randomized controlled trials (RCTs), which tend to focus on short-term outcomes [14,15]. Few studies have examined the associations between long-term patterns of physical activity and health-related quality of life [16]. As it is impractical to conduct a physical activity RCT over many years, we use causal inference methods to emulate a long-term RCT from observational data [17,18]. This analysis approach can produce valid causal inferences under some assumptions [19]. Previous studies have not considered complex forms of bias that can arise in longitudinal data or examined causal effects.

The aim of this study was to examine the associations between longitudinal patterns of physical activity over 15 years and health-related quality of life, including physical and mental health domains and individual subscales, in a cohort of middle-aged Australian women. We hypothesized that adherence to physical activity guidelines in mid-age would lead to improved quality of life, in both physical and mental health domains. Specifically, we used target trial emulation [20] to test the following hypotheses:

1. Consistently meeting physical activity guidelines in all surveys of the exposure period (2001 to 2016, when women aged from 50–55 to 65–70) ("consistent adherence" thereafter) will result in better health-related quality of life in 2019 (when women aged 68 to 73) than not meeting guidelines in any survey (control: "consistent non-adherence").

2. Starting to meet guidelines during the exposure period will result in better health-related quality of life than the control condition of consistent non-adherence, with earlier initiation associated with better outcomes.

## Methods

### Target trial emulation

A target trial emulation approach enables the use of observational data to emulate a "target" trial (i.e., the RCT that would have been conducted, were it possible to do so), with exposures framed as hypothetical interventions (e.g., the intervention group following a particular pattern of physical activity over the treatment/exposure period) [20]. We emulated a target trial in which the "intervention" was conceptualized as participants being assigned to meet the guidelines following different patterns across the course of the study with 100% adherence, compared with a control defined as meeting guidelines in none of the surveys (Table 1).

**Table 1. Comparison of target trial and emulated trial using ALSWH data.**
https://doi.org/10.1371/journal.pmed.1004384.t001

**Study population**

We used data from the 9 surveys (1996 to 2019) of the 1946 to 1951 birth cohort of the Australian Longitudinal Study on Women's Health (ALSWH), a population-based prospective cohort study. More information can be found on the ALSWH website (http://www.alswh.org.au), and details of the sample, study design, and recruitment are provided elsewhere [21,22]. The analytical sample includes 11,336 women born in 1946 to 1951 who were 45 to 50 years of age when they first completed the baseline survey in 1996. Women were randomly selected from the Medicare Australia national insurance database and followed up approximately every 3 years (1998, 2001, 2004, 2007, 2010, 2013, 2016, and 2019) with mailed surveys offered initially and web-based surveys more recently. The response rate for initial recruitment was estimated to be 53% to 56% [21]. The ALSWH study has ongoing ethical approval from the Human Research Ethics Committees of the Universities of Newcastle and Queensland (approval numbers H-076-0795 and 2004000224). All participants provided signed consent before participation. Participants whose physical functioning was in the lowest 5th percentile at the study "baseline" in 1998 (immediately prior to the exposure period) were excluded from this study as their physical functioning may have been too impaired to participate in physical activity. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist is included in S1 STROBE Checklist, and study flow chart is presented in Fig 1.
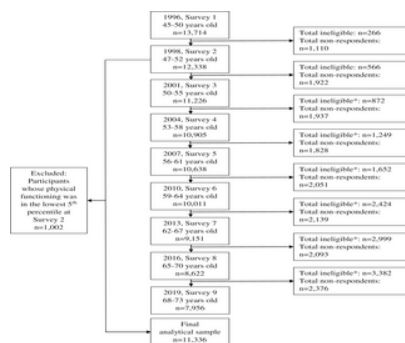


**Fig 1. Study flow chart.**
*Ineligible comprises: deceased, frailty and withdrawn.
https://doi.org/10.1371/journal.pmed.1004384.g001

**Outcome variables**

The primary outcomes are the physical health component (PCS) and mental health component (MCS) summary scores, derived from the 36-item Medical Outcomes Study short-form survey (SF-36) [23], a validated questionnaire covering various aspects of health-related quality of life. All except one of the 36 items were used to define 8 subscales: general health, physical functioning, physical role (role limitations from physical problems), bodily pain, vitality, social functioning, emotional role (role limitations from emotional problems), and mental health. These subscales were standardized and combined to form the PCS and MCS [24]. Scores range between 0 and 100 and higher scores indicate higher perceived physical/mental health. For this analysis, scales were standardized against data for Australian women of a similar age (age 45 to 54 years) [25]. We used the individual subscale scores as secondary outcomes. All outcomes were from survey 9 (2019) of the study.

**Exposure measurement**

Physical activity was measured using a validated and reliable modified version of the Active Australia survey questions [26]. Reported weekly minutes of activity (walking, moderate leisure activity, and vigorous leisure activity) were summed, with vigorous activity weighted double [27]. We dichotomized total physical activity based on reports of at least 150 min of weighted activity, as recommended by the WHO [28]. The exposure window was defined as a 15-year period, from 2001 to 2016, corresponding to study surveys 3 to 8. We did not use surveys 1 and 2 because the physical activity questions were different.

Emulated interventions.

We emulated the following 2 interventions against the control of consistent non-adherence to physical activity guidelines:

1. Consistent adherence to guidelines.

2. Starting to meet the guidelines on reaching the ages of 55, 60, or 65. Age interventions were distinct from each other and not cumulative.

More details on the methods used to define and evaluate the intervention regimes are included in S1 Text.

### Confounders

To illustrate the expected causal relationships among confounders, exposures, and outcomes over time, we developed a directed acyclic graph (DAG) (S1A and S1B Fig). A DAG is a conceptual presentation of the causal relationships between variables and can be used to inform research design and analysis [29]. In the case of the current study, the DAG helped us select confounders that could be common causes of both exposures and outcomes but not an intermediate outcome on the causal pathway (as the confounders were collected prior to the exposure). Baseline time-constant confounders, including highest level of education completed and country of birth, were from surveys 1 (1996) and 2 (1998). Time-varying confounders were from surveys 2 to 7 (1998 to 2016), including age (continuous), employment status (employed; not employed), area-level socioeconomic status, measured using the Index of Relative Socio-economic Disadvantage (IRSD) [30] (tertiles), geographical remoteness based on the postcode-level Accessibility-Remoteness Index of Australia Plus (ARIA+) [31] (major city; regional; remote), living with children (yes; no), marital status (married/de facto; separated/divorced/never married), diagnosis/treatment history coronary heart disease, stroke, arthritis, any cancer, anxiety, and depression (all yes; no), Center for Epidemiological Studies-Depression (CES-D) scale [32] (continuous), perceived stress scale [33] (continuous), SF-36 subscale scores, body mass index (BMI) (underweight; healthy; overweight; obese), lifetime risky alcohol consumption (>10 alcoholic drinks/week) based on the 2020 National Health Medical Research Council guidelines [34] (yes; no), heavy episodic alcohol consumption (more than 4 drinks on an occasion at least once a month) (yes; no), smoking status (never smoker; ex-smoker; current smoker), vegetable intake (<1; 1; 2; 3; 4; 5; ≥6 vegetables/day), and fruit intake (0; <1; 1; 2; 3; ≥4 pieces of fruit/day). S2 Text presents detail of the confounders.

### Statistical analysis

We preregistered the analysis (https://doi.org/10.17605/OSF.IO/24T6Z). We used models to emulate interventions, following the counterfactual framework, where each participant is considered to have a number of potential outcomes (corresponding to different exposures), only one of which is observed (because each person can only have 1 actual exposure history). Analytically, counterfactual methods use all available information to estimate the expected outcome had all participants followed a particular exposure pattern. More details about the counterfactual framework and its implementation are included in S1 Text.

Because the exposure group (e.g., adherence to physical activity guidelines) in the ALSWH is not randomly allocated, emulating the target trial requires controlling for confounding [35]. Because some of the confounding variables may themselves be affected by past exposure (e.g., $BMI_{t-1}$ is a confounder of the exposure time t, but the exposure may also likely affect subsequent observations of BMI; as indicated in S1A Fig by the arrows connecting $Exposure_t$ to $Time-varying\ confounders_t$), specialized analytical methods are required to produce unbiased causal effect estimates [19]. As such, we conducted all analyses using targeted maximum likelihood estimation (TMLE) [36], a consistent, doubly robust method for estimating causal effects, providing valid causal inference under the assumptions of conditional exchangeability [37], positivity [38], consistency [39], and no interference [40]. We conducted all analysis in R 4.2.1 via the "ltmle" package with models estimated via ensemble machine learning [41]. All analyses were weighted based on the probability of selection into the study [21]. Results are reported as expected mean scores, as well as mean difference versus control, and their 99.5% confidence intervals (CIs). This conservative alpha level (0.005) was selected based on the recommendations by Benjamin and colleagues [42]. Standard errors were estimated using the higher of the TMLE or influence-curve variance. Further details of the estimation methods and target estimands are included in S1 Text. Analysis code is available online (https://www.philipclare.com/code/alswh).

### Sensitivity analyses.

Additionally, we conducted 5 sets of sensitivity analyses:

1. We tested the hypothetical interventions of "starting the exposure period meeting the guidelines but stopping meeting them upon reaching 55, 60, and 65 years."

2. Analyses using a lower cut-point of 75 min instead of 150, to address the generally curvilinear relationship between physical activity and health-related outcomes, in which being active below the recommended levels could convey significant health benefits [2,3].

3. Analyses using a higher cut-point of 300 min to reflect the upper cut-point of the physical activity guidelines.

4. Analyses excluding a small number of variables that were wholly missing in a given survey because they were not included us that survey wave (S3 Text).

5. E-value analysis (S4 Text) to assess the potential impact of unmeasured confounding [43].

6. Post hoc sensitivity analyses excluding mental health confounders (depression, anxiety, and stress) in order to test for possible overadjustment.

### Missing data.

The data contained missing values from 2 scenarios: (1) intermittent missing data within surveys due to non-response to individual questions; and (2) missing data due to non-response to any entire survey (including loss to follow-up defined as those who did not complete a survey wave (or subsequent survey waves) of the study). The proportion of missing physical activity data ranged from 16% (survey 3) to 34% (survey 8) (Table A in S3 Text). The median (interquartile range) number of waves completed was 5 (3, 7). We handled intermittent missing data using multiple imputation [19] using chained equations, with M = 40 imputations [19], and data imputed using random forests [19]. Missing waves was handled using inverse probability of censoring weights [16]. Further details of the missing data and the procedures used to address them are included in S3 Text. A post hoc sensitivity analysis was subsequently conducted in which both intermittent missing data and missing waves were imputed, and the main analyses were conducted again using that version of the data (S7 Text).

## Results

### Sample characteristics

Characteristics of the weighted sample (n = 11,336) are presented in Table 2. Characteristics of the unweighted sample are presented in Table A in S5 Text. At baseline (1996), participants were aged 45 to 50 years (mean: 49.5 years; standard deviation [SD] 1.5). Three quarters of participants were born in Australia and almost two thirds had up to high school education only. Most participants were employed, married, or in a de facto relationship, and more than half lived with children (60.5%). Participants were

relatively equally distributed across area-level socioeconomic status. More than half lived in regional areas and nearly a third in major cities. One in 8 were lifetime risky drinkers, almost one third were heavy episodic drinkers, and almost 1 in 6 were current smokers. Forty five percent was overweight or obese.

| Variable | Categories | Mean (SD)/ n (%) (N = 11,336) |
|---|---|---|
| Age | | 49.5 (1.8) |
| Highest level of education | High school or less | 6,878 (60.0%) |
| | Trade/apprentice/certificate/ diploma | 2,280 (20.0%) |
| | University | 1,968 (20.0%) |
| Country of birth | Australia | 7,953 (70.0%) |
| | Other | 3,155 (30.0%) |
| Employment status | Not employed | 2,130 (20.0%) |
| | Employed | 8,976 (79.2%) |
| Marital status | Married/de facto | 9,094 (80.0%) |
| | Separated/divorced/never married | 1,813 (20.0%) |
| | Widowed | 257 (0.0%) |
| Live with children under 18 years | No | 6,693 (70.0%) |
| | Yes | 3,476 (30.0%) |
| Live with children aged 18+ years | No | 5,928 (60.0%) |
| | Yes | 4,299 (40.0%) |
| Area-level SES (IRSD) | Score for lowest tertile | 939 (72) |
| | Score for middle tertile | 985 (24) |
| | Score for highest tertile | 1,071 (82) |
| Remoteness (ARIA+) | Major city | 7,277 (70.0%) |
| | Regional | 3,680 (30.0%) |
| | Remote | 235 (0.0%) |
| Lifetime risky drinking[a] | No | 8,814 (90.0%) |
| | Yes | 1,395 (10.0%) |
| Heavy episodic drinking[b] | No | 7,007 (70.0%) |
| | Yes | 3,463 (30.0%) |
| Smoking status | Never smoker | 5,961 (60.0%) |
| | Ex smoker | 2,830 (30.0%) |
| | Current smoker | 1,689 (20.0%) |
| CESD-10 Depression Score | | 6.2 (7.1) |
| Mean stress score | | 0.6 (0.7) |
| BMI category | Underweight | 144 (0.0%) |
| | Healthy | 4,877 (50.0%) |
| | Overweight | 3,032 (30.0%) |
| | Obese | 1,783 (20.0%) |
| Ever diagnosed/treated for heart disease | No | 11,013 (100.0%) |
| | Yes | 216 (0.0%) |
| Ever diagnosed/treated for stroke | No | 11,148 (100.0%) |
| | Yes | 82 (0.0%) |
| Ever diagnosed/treated for cancer | No | 10,766 (100.0%) |
| | Yes | 464 (0.0%) |
| Ever diagnosed/treated for depression | No | 9,377 (80.0%) |
| | Yes | 1,853 (20.0%) |
| Ever diagnosed/treated for anxiety | No | 9,704 (90.0%) |
| | Yes | 1,526 (10.0%) |
| SF-36 scores | Physical component score | 51.14 (9.50) |
| | Mental component score | 47.51 (15.22) |
| | Physical functioning | 89.44 (14.27) |
| | Role physical | 83.98 (39.39) |
| | Bodily pain | 73.22 (27.60) |
| | General health | 74.59 (23.38) |
| | Vitality | 60.03 (24.95) |
| | Social functioning | 84.27 (27.07) |
| | Role emotional | 79.61 (43.49) |
| | Mental health | 73.77 (21.42) |

Note: percentages calculated as percent of total sample; percentages do not add to 100% due to missing data within variables.
[a]Lifetime risky alcohol consumption defined as >10 alcoholic drinks/week based on the 2020 National Health Medical Research Council guidelines [34].
[b]Heavy episodic alcohol consumption defined as >4 alcoholic drinks on an occasion at least once a month [34].
ARIA+, Accessibility-Remoteness Index of Australia Plus; BMI, body mass index; CESD-10, 10-item Centre for Epidemiological Studies Depression Scale; IRSD, Index of Relative Socio-economic Disadvantage; SD, standard deviation; SF-36: 36-item Medical Outcomes Study short-form survey.

https://doi.org/10.1371/journal.pmed.1004384.t002

**Table 2. Descriptive statistics of the weighted analytical sample of Australian women at baseline.**
https://doi.org/10.1371/journal.pmed.1004384.t002

**Main analyses**

Consistent adherence to physical activity guidelines (PCS: 46.93 [99.5% CI: 46.32, 47.54]) was associated with three-point higher PCS (mean difference: 3.0 [99.5% CI: 1.8, 4.1]) than the control (consistent non-adherence, PCS: 43.90 [42.79, 45.01]) (Fig 2 and Table A in S6 Text). While starting to meet guidelines upon reaching 55 years was associated with a similarly higher PCS (46.96 [45.53, 48.40]; mean difference versus control: 3.0 [99.5% CI: 1.2, 4.8]), starting to meet guidelines upon reaching 60 (mean difference versus control: 1.1 [99.5% CI: −0.6, 2.8]) and 65 (mean difference versus control: 0.2 [99.5% CI: −0.9, 1.3]) were not associated with statistically different PCS compared with the control. In terms of subscales (Tables B and J and Fig A in S6 Text), analyses show 5- to 9-point differences between consistent adherence to guidelines and consistent non-adherence in all 4 physical subscales, with the largest effect observed for "physical functioning" (mean difference: 9.3 [99.5% CI: 6.2, 12.3]). For all 4 physical health subscales, there was a similar magnitude of difference for starting to meet guidelines on reaching age 55 compared with the control, but no evidence for meaningful difference for starting to meet the guidelines on reaching age 60 and 65.



**Fig 2. Primary analysis of the effect of age at which started to meet physical activity guidelines on SF-36 component scores.**

Abbreviations: SF-36, 36-item Medical Outcomes Study short-form survey. This figure shows the effect of meeting physical activity upon reaching a particular age ("starters") but not prior to that, with a range of age thresholds considered (55, 60, and 65 years) on the physical health component and mental health component scores at survey 9. The points represent the estimates and the bars the 99.5% confidence intervals. Models were adjusted for: highest level of education, country of birth, age, employment status, living with children, marital status, SEIFA IRSD, geographical remoteness (ARIA+), history of coronary heart disease diagnosis/treatment, history of stroke diagnosis/treatment, history of arthritis diagnosis/treatment, history of any cancer diagnosis/treatment, history of anxiety diagnosis/treatment, and history of depression diagnosis/treatment, CES-D scale, stress, SF-36 subscale scores, BMI, lifetime risky alcohol consumption based on the 2020 National Health Medical Research Council guidelines, heavy episodic alcohol consumption, smoking status, vegetable intake, and fruit intake. ARIA+, Accessibility-Remoteness Index of Australia Plus; BMI, body mass index; CES-D, Center for Epidemiological Studies-Depression; SEIFA IRSD, Socio-economic Index for Areas Index of Relative Socio-economic Disadvantage.
https://doi.org/10.1371/journal.pmed.1004384.g002

We found no statistically significant differences in MCS between any hypothetical intervention and the control. In terms of mental subscales, consistent adherence to guidelines resulted in statistically higher scores for "vitality" (5-point difference). We also found a significant difference in "vitality" (around 5-point differences) linked to meeting guidelines on reaching 55 years. However, we did not find statistically significant differences between starting to meet guidelines on reaching 60 and 65, compared with the control, for any of the mental subscales.

**Sensitivity analyses**

Tables A and B and Figs B and C in S6 Text outlined the potential interventions of stopping meeting guidelines at ages 55, 60, and 65. Starting the exposure period meeting the guidelines, but then stopping meeting them on reaching age 65 was associated with nearly a three-point higher PCS than consistent non-adherence. The higher PCS was similar in magnitude to that associated with consistent adherence. Stopping meeting guidelines on reaching 55 or 60 years was not associated with statistically different PCS from the control. There were no statistically significant differences in MCS between any hypothetical intervention and the control.

In sensitivity analyses using a lower threshold of 75 min per week of physical activity, patterns for the primary (Table C and Figs D and E in S6 Text) and secondary analyses (Table D and Figs F and G in S6 Text) were similar, but differences were less marked. In sensitivity analyses using a higher threshold of 300 min per week of physical activity, patterns for the primary (Table E and Figs H and I in S6 Text) and secondary analyses were similar (Table F and Figs J and K in S6 Text). Findings from sensitivity analyses excluding confounders in surveys where they were completely missing (Tables G and H and Figs L–O in S6 Text) were similar to those from the main analyses.

The results of the E-Value analysis suggest that a moderate amount of unmeasured confounding would be required to alter the interpretation of the results (Tables I and J in S6 Text), with unmeasured confounding needing to double (or halve) the risk of the exposure in order to alter the significant findings reported.

Post hoc sensitivity analyses excluding mental health confounders returned very similar results to the primary analysis, suggesting that results were not due to overadjustment (Figs P–S in S6 Text). The unplanned post hoc sensitivity analyses where both intermittent missing data and missing waves were imputed showed findings similar to those from the main analyses (Table A in S7 Text).

## Discussion

Overall, the findings from this study highlight the importance of being physically active for as long as possible to gain the most benefits for physical health. Our findings show that consistent adherence to physical activity guidelines over 15 years was associated with better health-related quality of life in the physical health domain (three-point higher PCS on a scale of 1 to 100), compared with not adhering to guidelines in all surveys. Even partial adherence to guidelines during the 15 years, namely starting to meet the guidelines on reaching age 55 and maintaining meeting the guidelines from the start of the exposure period until reaching 65, was associated with better physical health-related quality of life at around age 70. Consistent or partial adherence to physical activity guidelines seemed to have weaker effects on the mental health domain, with some significant differences observed in subscales only, but not with the overall MCS.

The magnitude of benefits regarding the physical health domain, either from consistent adherence to or earlier adoption or longer-term maintenance of meeting physical activity guidelines, is small but meaningful. Physical functioning is critical for maintaining independent functional ability, which is a predictor of various health outcomes including disability, hospitalization, and mortality [44,45]. For example, a 10-point difference in physical functioning scores may be interpreted as the difference between those with and without mild osteoarthritis [46]. A recent meta-analysis has found that even a one-unit increases in SF-36 physical functioning and the PCS are associated with lower risks of mortality in the general population [15].

To date, few studies have examined longitudinal associations between physical activity at multiple time points and health-related quality of life [16,47]. Previous analyses of data from the 1946 to 1951 ALSWH cohort, using the 6-year period from surveys 3 to 5, found small improvements in both PCS and MCS, but meaningful improvements in physical functioning and vitality with increases in total physical activity from 2001 to 2007 [16]. Analyses of data from the Nurses' Health Study also showed associations between increasing physical activity over 10 years from 1986 to 1996 and improvements in health-related quality of life scores (eight-point improvement in physical function, four-point improvement in vitality, and a two-point improvement in the mental health subscale) [47]. The magnitude of the differences in health-related quality of life scores in our data were similar in the mental domain to those reported in these earlier papers [16,47] and similar in the physical domain to those reported from the Nurses' Health Study [47]. The differences in PCS and physical functioning scores were higher in our study than in the previous 6-year analysis of data from this ALSWH cohort [16,47]. However, it is important to note that our analyses were not directly comparable with these earlier studies because we did not compare women with different physical activity patterns, but rather, the expected differences in health-related quality of life if all women followed a specific intervention during the exposure period. Therefore, using a causal inference framework, our analyses assumed that women with different physical activity interventions and the control are the same ("exchangeable") with respect to observed confounders, except for their physical activity levels.

Our modeling of dynamic interventions defined by women changing physical activity in midlife is based on the consideration that some women may not have been physically active enough until their fifties. It is therefore important to understand whether adopting an active lifestyle in midlife still conveys benefits. Our findings suggest that to maintain good physical health-related quality of life at around age 70, one may be able to "make up" for not being active earlier by becoming active in the mid-50s. This finding supports

public health initiatives for messaging around "turning back the clock" in midlife through lifestyle changes such as physical activity. The nonsignificant findings about being physically active only at age 60 or 65 may not be interpreted as the window opportunity being closed, but rather, that there had not been sufficient accumulation of physical activity for the health benefits to be evident by around age 70.

The main strengths of the study are the large cohort and the repeated physical activity measures over a long exposure period. Another strength is that we accounted for a wide range of potential time-varying confounders, using robust causal inference methods. Provided structural assumptions hold (positivity, no interference, consistency, and no unmeasured confounding), these methods allow us to estimate causal effects rather than associations, although even if those assumptions cannot be relied upon, these methods provide at least as reliable estimates as methods that only examine associations. In this case, the assumptions of positivity, no interference and consistency at least appeared reasonable. Meeting physical activity guidelines is common enough that there were no positivity violations, and TMLE is robust to near-positivity violations. Because the study uses data from a population cohort, and is based on clear and unambiguous definitions of treatment/control, there is little chance of violations of either no interference or consistency. Unmeasured confounding is possible despite adjusting for health variables and conditions, and excluding data from women with poor physical functioning, we cannot guarantee that the study controls for all confounders for the relationship between physical activity and health-related quality of life, such as underlying health conditions and/or other unmeasured or insufficiently measured confounders. However, E-Value analysis suggests that the key differences observed would require fairly substantial unmeasured confounding to result in different conclusions (requiring a 50%+ change in exposure and outcome in order to affect the conclusions in most cases). Another limitation is the use of self-reported physical activity; however, this measure has good reliability and validity [26]. In this study, we could not emulate all possible interventions. Finally, although the ALSWH included nationally representative samples at baseline [21], healthier women have remained in the study, so the findings may not be generalizable to all mid-age Australian women [48].

In conclusion, this study contributes to our understanding of the prospective associations between physical activity and health-related quality of life in mid-age women. An important public health message is that being active for as many years as possible, even if women start to meet physical activity guidelines in their mid-50s, could have important health benefits in terms of physical health, especially in physical functioning. Our study contributes to growing evidence of the health benefits of maintaining or adopting an active lifestyle in mid-age. Such public health messages should be used to encourage middle-aged women to become and stay active.

## Supporting information

**S1 STROBE Checklist. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist.**
https://doi.org/10.1371/journal.pmed.1004384.s001
(DOCX)

**S1 Text. Model estimands.**
https://doi.org/10.1371/journal.pmed.1004384.s002
(DOCX)

**S2 Text. Confounding variables.**
https://doi.org/10.1371/journal.pmed.1004384.s003
(DOCX)

**S3 Text. Missing data.**
https://doi.org/10.1371/journal.pmed.1004384.s004
(DOCX)

**S4 Text. E-value analysis.**
https://doi.org/10.1371/journal.pmed.1004384.s005
(DOCX)

**S5 Text. Descriptive statistics of the unweighted analytical sample.**
https://doi.org/10.1371/journal.pmed.1004384.s006
(DOCX)

**S6 Text. Additional results.**
https://doi.org/10.1371/journal.pmed.1004384.s007
(DOCX)

**S7 Text. Post hoc sensitivity analysis findings.**
https://doi.org/10.1371/journal.pmed.1004384.s008
(DOCX)

**S1 Fig. Directed acyclic graph.**
https://doi.org/10.1371/journal.pmed.1004384.s009
(DOCX)

## Acknowledgments

## References

1. Lear SA, Hu W, Rangarajan S, Gasevic D, Leong D, Iqbal R, et al. The effect of physical activity on mortality and cardiovascular disease in 130 000 people from 17 high-income, middle-income, and low-income countries: the PURE study. Lancet. 2017;390(10113):2643–2654.
   View Article • Google Scholar

2. World Health Organization. WHO guidelines on physical activity and sedentary behaviour. Geneva: World Health Organization; 2020.

3. U.S. Department of Health and Human Services. Physical Activity Guidelines for Americans. 2nd ed. Washington, DC: U.S. Department of Health and Human Services; 2018.

4. Moore SC, Lee IM, Weiderpass E, Campbell PT, Sampson JN, Kitahara CM, et al. Association of leisure-time physical activity with risk of 26 types of cancer in 1.44 million adults. JAMA Intern Med. 2016;176(6):816–825. pmid:27183032
   View Article  •  PubMed/NCBI  •  Google Scholar

5. Brown W, Heesch KC, Miller YD. Life events and changing physical activity patterns in women at different life stages. Ann Behav Med. 2009;37(3):294–305. pmid:19506989
   View Article  •  PubMed/NCBI  •  Google Scholar

6. Mok A, Khaw KT, Luben R, Wareham N, Brage S. Physical activity trajectories and mortality: population based cohort study. BMJ. 2019;365:l2323. pmid:31243014
   View Article  •  PubMed/NCBI  •  Google Scholar

7. Saint-Maurice PF, Coughlan D, Kelly SP, Keadle SK, Cook MB, Carlson SA, et al. Association of leisure-time physical activity across the adult life course with all-cause and cause-specific mortality. JAMA Netw Open. 2019;2(3):e190355. pmid:30848809
   View Article  •  PubMed/NCBI  •  Google Scholar

8. Mielke GI. Relevance of life course epidemiology for research on physical activity and sedentary behavior. J Phys Act Health. 2022;19(4):225–226. pmid:35299155
   View Article  •  PubMed/NCBI  •  Google Scholar

9. Gomes GAO, Brown WJ, Codogno JS, Mielke GI. Twelve year trajectories of physical activity and health costs in mid-age Australian women. Int J Behav Nutr Phys Act. 2020;17(1):101. pmid:32778110
   View Article  •  PubMed/NCBI  •  Google Scholar

10. World Health Organization. Healthy ageing and functional ability [Internet]. Available from: https://www.who.int/news-room/questions-and-answers/item/healthy-ageing-and-functional-ability. Accessed [2023 Jan 19].

11. Kaplan RM, Hays RD. Health-related quality of life measurement in public health. Annu Rev Public Health. 2022;43(1):355–373. pmid:34882431
   View Article  •  PubMed/NCBI  •  Google Scholar

12. Phyo AZZ, Freak-Poli R, Craig H, Gasevic D, Stocks NP, Gonzalez-Chica DA, et al. Quality of life and mortality in the general population: a systematic review and meta-analysis. BMC Public Health. 2020;20(1):1596. pmid:33153441
   View Article  •  PubMed/NCBI  •  Google Scholar

13. Hennessy CH, Moriarty DG, Zack MM, Scherr PA, Brackbill R. Measuring health-related quality of life for public health surveillance. Public Health Rep. 1994;109(5):665–672. pmid:7938388
   View Article  •  PubMed/NCBI  •  Google Scholar

14. Bize R, Johnson JA, Plotnikoff RC. Physical activity level and health-related quality of life in the general adult population: a systematic review. Prev Med. 2007;45(6):401–415. pmid:17707498
   View Article  •  PubMed/NCBI  •  Google Scholar

15. Marquez DX, Aguiñaga S, Vásquez PM, Conroy DE, Erickson KI, Hillman C, et al. A systematic review of physical activity and quality of life and well-being. Transl Behav Med. 2020;10(5):1098–1109. pmid:33044541
   View Article  •  PubMed/NCBI  •  Google Scholar

16. Heesch KC, van Uffelen JG, van Gellecum YR, Brown WJ. Dose-response relationships between physical activity, walking and health-related quality of life in mid-age and older women. J Epidemiol Community Health. 2012;66(8):670–677. pmid:22544920
   View Article  •  PubMed/NCBI  •  Google Scholar

17. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology. 2000;11(5):550–560. pmid:10955408
   View Article  •  PubMed/NCBI  •  Google Scholar

18. van der Laan MJ, Gruber S. Targeted minimum loss based estimation of causal effects of multiple time point interventions. Int J Biostat. 2012;8(1). pmid:22611591
   View Article  •  PubMed/NCBI  •  Google Scholar

19. Clare PJ, Dobbins TA, Mattick RP. Causal models adjusting for time-varying confounding—a systematic review of the literature. Int J Epidemiol. 2019;48(1):254–265. pmid:30358847
   View Article  •  PubMed/NCBI  •  Google Scholar

20. Hernán MA, Wang W, Leaf DE. Target trial emulation: a framework for causal inference from observational data. JAMA. 2022;328(24):2446–2447. pmid:36508210
   View Article  •  PubMed/NCBI  •  Google Scholar

21. Brown WJ, Bryson L, Byles JE, Dobson AJ, Lee C, Mishra G, et al. Women's Health Australia: recruitment for a national longitudinal cohort study. Women Health. 1999;28(1):23–40.
   View Article  •  Google Scholar

22. Dobson AJ, Hockey R, Brown WJ, Byles JE, Loxton DJ, McLaughlin D, et al. Cohort profile update: Australian Longitudinal Study on Women's Health. Int J Epidemiol. 2015;44(5):1547,1547a–1547f. pmid:26130741
View Article   •   PubMed/NCBI   •   Google Scholar

23. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). Conceptual framework and item selection. Med Care. 1992;30(6):473–483.
View Article   •   Google Scholar

24. Ware JE, Kosinski M, Keller SK. SF-36 physical and mental health summary scales: A user's manual. Boston, MA: Health Assessment Lab, 1994.

25. Australian Bureau of Statistics. 4399.0—National Health Survey: SF 36 Population Norms, Australia, 1995. Available from: https://www.abs.gov.au/ausstats/abs@.nsf/mf/4399.0. Accessed [2024 Jan 19].
View Article   •   Google Scholar

26. Brown WJ, Burton NW, Marshall AL, Miller YD. Reliability and validity of a modified self-administered version of the Active Australia physical activity survey in a sample of mid-age women. Aust N Z J Public Health. 2008;32(6):535–41. pmid:19076744
View Article   •   PubMed/NCBI   •   Google Scholar

27. Australian Institute of Health and Welfare. The Active Australia Survey: a guide and manual for implementation, analysis and reporting. Canberra: AIWH; 2003.

28. 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(24):1451–1462. pmid:33239350
View Article   •   PubMed/NCBI   •   Google Scholar

29. Pearl J. Causal diagrams for empirical research. Biometrika. 1995;82(4):669–688.
View Article   •   Google Scholar

30. Australian Bureau of Statistics. Technical Paper: Socio-Economic Indexes for Areas (SEIFA). 2016.
View Article   •   Google Scholar

31. Australian Bureau of Statistics. The Australian statistical geography standard (ASGS) remoteness structure. Australian Bureau of Statistics Canberra; 2018.

32. Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. Appl Psychol Measur. 1977;1(3):385–401.
View Article   •   Google Scholar

33. Bell S, Lee C. Development of the perceived stress questionnaire for young women. Psychol Health Med. 2002;7(2):189–201.
View Article   •   Google Scholar

34. Australian Government National Health and Medical Research Council. Australian guidelines to reduce health risks from drinking alcohol. 2020. Available from: https://www.nhmrc.gov.au/file/15923/download?token=t0Hrxdvq. Accessed [2024 Jan 19].
View Article   •   Google Scholar

35. Moreno-Betancur M, Lynch JW, Pilkington RM, Schuch HS, Gialamas A, Sawyer MG, et al. Emulating a target trial of intensive nurse home visiting in the policy-relevant population using linked administrative data. Int J Epidemiol. 2023 Feb 8;52(1):119–131. pmid:35588223
View Article   •   PubMed/NCBI   •   Google Scholar

36. van der Laan MJ, Rubin DB. Targeted maximum likelihood learning. Int J Biostat. 2006;2(1).
View Article   •   Google Scholar

37. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. Epidemiology. 1999;10(1):37–48. pmid:9888278
View Article   •   PubMed/NCBI   •   Google Scholar

38. Petersen ML, Porter KE, Gruber S, Wang Y, van der Laan MJ. Diagnosing and responding to violations in the positivity assumption. Stat Methods Med Res. 2012;21(1):31–54. pmid:21030422
View Article   •   PubMed/NCBI   •   Google Scholar

39. Cole SR, Frangakis CE. Commentary: the consistency statement in causal inference: a definition or an assumption? Epidemiology. 2009;20(1):3–5.
View Article   •   Google Scholar

40. Liu L, Hudgens MG, Becker-Dreps S. On inverse probability-weighted estimators in the presence of interference. Biometrika. 2016;103(4):829–42. pmid:29422692
View Article   •   PubMed/NCBI   •   Google Scholar

41. Lendle SD, Schwab J, Petersen ML, van der Laan MJ. ltmle: an R package implementing targeted minimum loss-based estimation for longitudinal data. J Stat Softw. 2017;81(1):1–21.
View Article   •   Google Scholar

42. Benjamin DJ, Berger JO, Johannesson M, Nosek BA, Wagenmakers EJ, Berk R, et al. Redefine statistical significance. Nat Hum Behav. 2018;2:6–10. pmid:30980045
View Article   •   PubMed/NCBI   •   Google Scholar

43. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-Value. Ann Intern Med. 2017;167(4):268–274. pmid:28693043
   View Article • PubMed/NCBI • Google Scholar

44. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol. 1994;49(2):M85–94. pmid:8126356
   View Article • PubMed/NCBI • Google Scholar

45. Penninx BW, Ferrucci L, Leveille SG, Rantanen T, Pahor M, Guralnik JM. Lower extremity performance in nondisabled older persons as a predictor of subsequent hospitalization. J Gerontol. 2000;55(11):M691–M697. pmid:11078100
   View Article • PubMed/NCBI • Google Scholar

46. Brook RH, Ware JE Jr, Rogers WH, Keeler EB, Davies AR, Donald CA, et al. Does free care improve adults' health? Results from a randomized controlled trial. N Engl J Med. 1983;309(23):1426–1434. pmid:6355851
   View Article • PubMed/NCBI • Google Scholar

47. Wolin KY, Glynn RJ, Colditz GA, Lee IM, Kawachi I. Long-term physical activity patterns and health-related quality of life in U.S. women. Am J Prev Med. 2007;32(6):490–9. pmid:17533064
   View Article • PubMed/NCBI • Google Scholar

48. Young AF, Powers JR, Bell SL. Attrition in longitudinal studies: who do you lose? Aust N Z J Public Health. 2006;30:353e61.
   View Article • Google Scholar