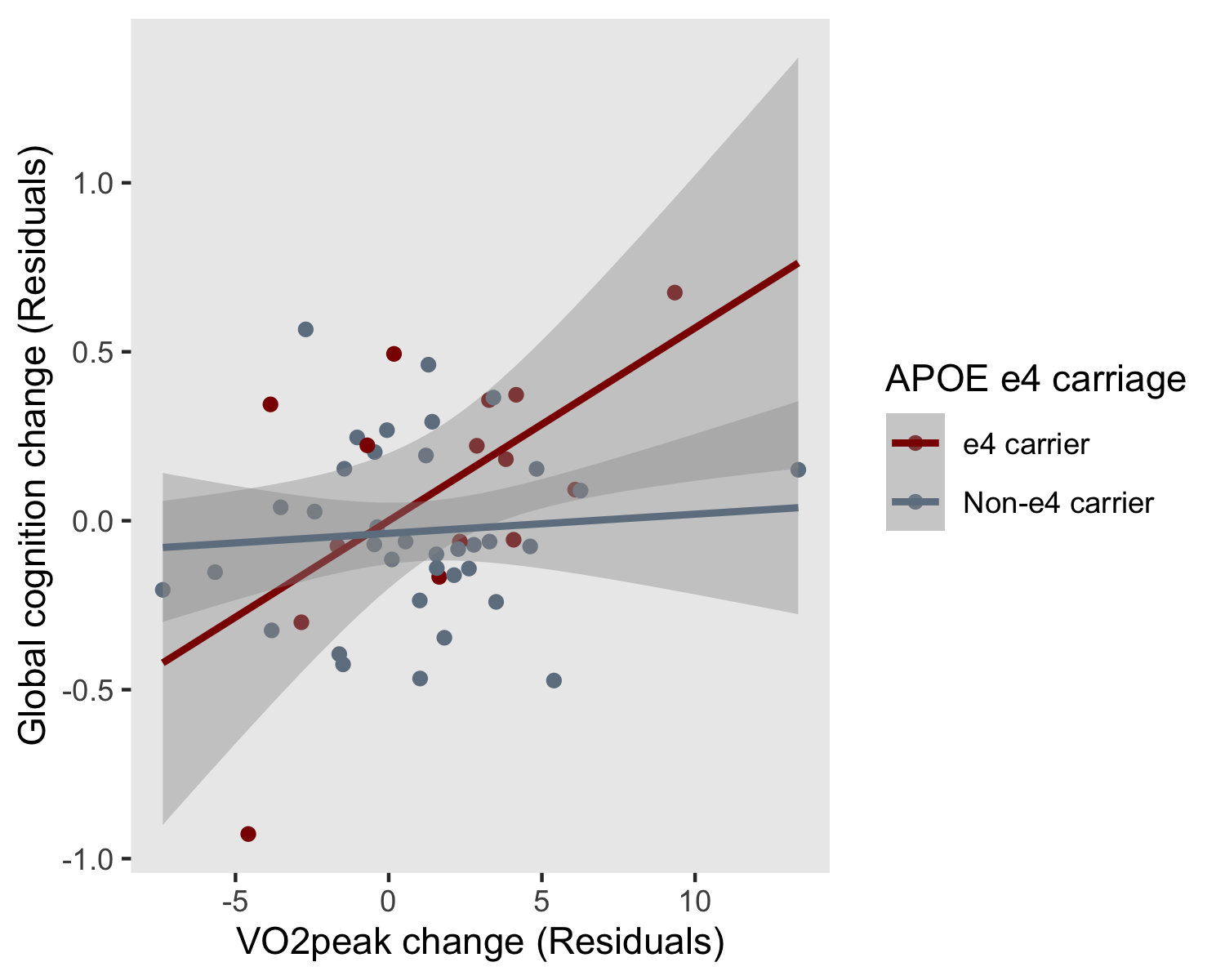


**eFigure 1:** Percentage of peak power for each month of the intervention for the high-intensity and moderate-intensity exercise groups. Each data point represents a participant’s mean peak power (from all attended exercise sessions) for that month. A VO2peak test following the third month was used for re-calculation of power. Corresponding months are different between groups (p < 0.001); whereby the high-intensity group had higher percentage peak power during all months of the intervention. Abbreviations: M1-6, Month 1 – 6; W, watts.

******

**eFigure 2:** Linear relationship between change in cardiorespiratory fitness (residuals) and change in global cognition (residuals) from pre- to immediately post-intervention (6-months) in APOE ε4 carriers and non-carriers Abbreviations: APOE, Apolipoprotein E; VO2peak, peak aerobic capacity (fitness measurement).

**eTable 1:** Comparison of baseline characteristics between participants included and excluded from per-protocol analyses

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Included in per-protocol analyses**  (*n* = 86) | **Excluded from per-protocol analyses**  (*n* = 13) | **Test statistic** |
| **Age,** years | 69.3 ± 5.2 | 67.9 ± 5.0 | *t* = -0.87 |
| **Gender,** % Female (n) | 52.3 (45) | 69.2 (9) | *χ2*= 1.30 |
| ***APOE* ε4 allele carriers,** % (n) | 29.1 (25) | 7.7 (1) | *χ2*= 2.66 |
| ***BDNF* Val66Met carriers,** % (n) | 38.4 (33) | 38.5 (5) | *χ2*= 0.00 |
| **Years of education** | 13.9 ± 2.3 | 15.3 ± 2.1 | *t* = 2.17\* |
| **Global cognition,** MoCA score | 26.4 ± 2.4 | 26.6 ± 2.0 | *t =* 0.36 |
| **Baseline VO2peak** (ml/kg/min) | 23.1 ± 6.2 | 23.6 ± 6.7 | *t =* 0.24 |
| **Baseline peak power** (W) | 133.0 ± 43.2 | 138.0 ± 67.8 | *t =* 0.35 |
| **Alcohol,** Units per week | 5.8 ± 5.9 | 5.5 ± 5.4 | *t =* -0.16 |
| **Physical activity** (Met.min/wk-1) | 4100.0 ± 2786.3 | 4041.7 ± 3162.5 | *t =* -0.07 |
| **DASS Depression score** | 1.9 ± 2.4 | 1.5 ± 2.3 | *t =* -0.58 |
| **Daily kilojoule intakea** | 6852.5 ± 2711.5 | 5777.5 ± 1782.3 | *t =* -1.27 |
| **BMI** (kg/m2) | 25.7 ± 3.8 | 25.6 ± 2.8 | *t =* -0.05 |
| **Waist-hip ratio** | 0.88 ± 0.08 | 0.87 ± 0.05 | *t =* -0.19 |

*\*P* < 0.05, Test statistics determined by independent samples t-test for continuous variables and chi-square for categorical variables. Abbreviations: *APOE*, Apolipoprotein E; *BDNF* Val66Met, brain-derived neurotrophic factor Valine66Methionine single nucleotide polymorphism; BMI, body mass index; DASS, Depression, Anxiety and Stress Scales; Met.min/wk-1, metabolic minutes per week (subjective habitual physical activity measurement); MoCA, Montreal Cognitive Assessment; VO2peak, peak aerobic capacity (fitness measurement); W, wattage. aDaily kilojoule intake quantified from the Cancer Council of Victoria Food Frequency Questionnaire.

**eTable 2:** Beta coefficients and standard error of time\*group from linear mixed models, and projected sample size for each group to detect a significant effect from longpower function in R statistical package

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Cognitive composite variable** | **Baseline to 6 months; Intention-to-treat analysis** | | | | | | **All timepoints; Intention-to-treat analysis** | | | | | | |
| **Control versus High-intensity** | | | **Control versus Moderate-intensity** | | | **Control versus High-intensity** | | | **Control versus Moderate-intensity** | | | |
| *B (SE)* | *p-value* | *Projected sample size*  *(per group)* | *B (SE)* | *p-value* | *Projected sample size*  *(per group)* | *B (SE)* | *p-value* | *Projected sample size*  *(per group)* | *B (SE)* | *p-value* | *Projected sample size*  *(per group)* |
| **Global Cognitive Composite** | -0.076 (0.136) | 0.58 | 492 | 0.080 (0.134) | 0.55 | 692 | -0.048 (0.052) | 0.35 | 129 | -0.035 (0.050) | 0.48 | 479 |
| **Executive Function composite** | 0.010 (0.177) | 0.97 | 782 | 0.216 (0.174) | 0.22 | 625 | 0.042 (0.077) | 0.59 | >1000 | 0.055 (0.076) | 0.47 | >1000 |
| **Episodic Memory composite** | 0.035 (0.213) | 0.87 | >1000 | 0.046 (0.205) | 0.82 | >1000 | -0.079 (0.078) | 0.31 | 207 | -0.047 (0.076) | 0.53 | 531 |
| **Attention composite** | 0.017 (0.294) | 0.95 | >1000 | 0.492 (0.282) | 0.08 | 123 | -0.053 (0.104) | 0.61 | 420 | 0.001 (0.102) | 0.99 | >1000 |

Abbreviations: B (SE), Unstandardised beta (standard error). A positive beta represents a larger positive slope for the moderate/high-intensity groups, compared with the control group.

**Supplementary methods**

**Exercise interventions**

The high-intensity exercise sessions commenced and finished with a 10-minute warm-up and cool-down during which participants cycled at a perceived exertion rating of 11 (30-40% aerobic capacity). The active portion of the intervention involved completion of 11 cycling-based intervals of 1 minute of high exertion (>80% aerobic capacity; 18.0 Borg Scale) interspersed with 2 minutes of active recovery (30-40% aerobic capacity; 12.0 Borg Scale).

The moderate-intensity exercise group participants were required to cycle continuously at a rating of perceived exertion of 13 (50-60% aerobic capacity) for 50 minutes. The high-intensity and moderate-intensity protocols were work-matched based on an 80 kg person with a maximal aerobic capacity of 27 ml.kg-1.min-1, to yield approximately 386 Met.min-1 for a moderate intensity session and 380 Met.min-1 for a high-intensity session.

**Cognitive assessments and composite scores**

The battery included the Montreal Cognitive Assessment (MoCA), Wechsler Adult Intelligence Scale-III Digit Span, California Verbal Learning Test (CVLT-II), Brief Visual Memory Test (BVMT), Trail Making Test forms A and B, and the NIH EXAMINER Verbal fluency task, Flanker, and Set-shifting. A computerised Cogstate battery (www.cogstate.com) was administered including Groton Maze learning and recall, and identification, detection, one-card learning, and one-back tasks.

We calculated cognitive composite scores using z-scores of individual’s performance across all timepoints. For scores where a lower score indicates better performance (i.e., speed), we inversed the score ([score]\*-1). The composite scores included the following tasks: 1) Global cognitive composite: Digit Span, Cogstate one-back, Cogstate identification task, CVLT (learning, short delay recall, long delay recall, and recognition d`), BVMT (learning and long delay recall), Cogstate Groton Maze recall, Trails B, Phonemic fluency, Flanker, and Set-shifting; 2) Attention: Digit Span (Forward only) and Cogstate identification task; 3) Episodic Memory: CVLT (short delay recall, long delay recall, and recognition d`), BVMT long delay recall, and Groton Maze recall; and 4) Executive function: Trails B, Phonemic fluency, Flanker, and Set-shifting.

**Graded exercise test**

The graded exercise test followed a step protocol using two-minute stages increasing in intensity until volitional fatigue. The initial power output and step progressions were standardised to baseline body mass to enable similar test durations for individuals: 1) under 70kg, 2) between 70 to 100kg and 3) over 100kg. During the graded exercise test, heart rate was continuously monitored and expired ventilation collected using a Parvo TrueOne metabolic cart (ParvoMedics, USA). VO2peak was determined as the highest 15-s mean VO2 value recorded in the final 2 minutes of the test.

Peak aerobic power was determined using the following equation: PLCS + Fst \* BMP; where PLCS is the power at the last stage completed, Fst is the fraction of the last uncompleted stage and BMP is the body-mass specific increase in work rate per stage.

**Statistical analysis: longpower**

Based on our finding that the exercise interventions did not directly impact cognition, we calculated the post-hoc power that would be required to detect significant differences between groups in each of the cognitive composite scores (using the lmmpower function [23]). Beta coefficients, standard errors, and estimated required samples