nature medicine

Article

https://doi.org/10.1038/s41591-026-04239-3

Effects of daily multivitamin–multimineral and cocoa extract supplementation on epigenetic aging clocks in the COSMOS randomized clinical trial

Received: 11 March 2025

Accepted: 22 January 2026

Published online: 09 March 2026

Check for updates

A list of authors and their affiliations appears at the end of the paper

Large-scale randomized trials have found that multivitamin–multimineral (MVM) supplements and cocoa flavanols may benefit several age-related chronic conditions among older adults, but it remains unclear whether these two supplements directly slow the biological aging process. This prespecified ancillary study evaluated the 2-year effect of a daily MVM (Centrum Silver) and cocoa extract (500 mg cocoa flavanols per day, including 80 mg (−)-epicatechin) on five DNA methylation measures of biological aging (PCHannum, PCHorvath, PCPhenoAge, PCGrimAge and DunedinPACE) among 958 participants (482 women and 476 men) in the COcoa Supplement and Multivitamin Outcomes Study (COSMOS). Compared with placebo, daily MVM supplementation modestly reduced the rate of increase of second-generation epigenetic clocks, with a between-group difference in yearly change of −0.113 years (95% confidence interval (CI) −0.205 to −0.020; $P = 0.017$) for PCGrimAge and −0.214 years (−0.410 to −0.019; $P = 0.032$) for PCPhenoAge. MVM had a stronger effect on PCGrimAge among those with accelerated biological aging at baseline (−0.236 [−0.380 to −0.091]) compared with those with normal or decelerated biological aging (−0.013 [−0.130 to 0.104]; $P = 0.018$ for interaction). Cocoa extract did not have an effect on the five epigenetic clocks tested. Although the statistically significant but small effects of daily MVM supplementation on slowing biological aging are encouraging, additional studies are needed to determine the clinical relevance of daily MVM supplementation on epigenetic clocks and whether such effects can help explain the beneficial effects of MVM supplementation on aging-related chronic conditions.

It has been estimated that half of the global disease burden among adults is attributable to aging-related disease[1]. Biological aging reflects a gradual and progressive functional decline in system integrity occurring with advancing chronological age and exhibits considerable variation in populations. The geroscience hypothesis proposes that interventions targeted to slow or reverse molecular changes that occur with aging could promisingly delay, prevent or alleviate age-related diseases and conditions and extend healthy lifespan[2,3]. DNA methylation (DNAm) patterns usually undergo notable changes over the course of aging, potentially contributing to aberrant specific gene repression or activation[1]. Epigenetic clocks are derived from DNAm patterns, and their residuals regressed to chronological age (age deviation, AgeDev) have emerged as powerful tools to accurately estimate the biological aging process, predict age-related outcomes and potentially capture aging-related changes in age-reversal trials[4,5]. Thus far, relatively little is known about the responsiveness of epigenetic measures in clinical

e-mail: hsesso@bwh.harvard.edu

https://doi.org/10.1038/s41591-026-04239-3

trial settings, especially for the alignment between change in epigenetic biomarkers and the long-term benefits measured by clinical endpoints.

Inadequate micronutrient intakes and related deficiencies are a major challenge to global public health, and recent global estimates suggested that more than 5 billion people do not consume enough essential micronutrients from food, especially iodine, vitamin E and calcium[6]. MVM formulations are the most common dietary supplement in the USA and other high-income countries[7], typically providing a combination of essential vitamins and minerals to meet daily recommended nutrient intake and mitigate the risk of nutritional deficiencies or insufficiencies[8]. We previously found in large-scale randomized clinical trials (RCTs) that daily MVM use reduced cancer and cataracts among older male physicians in the Physicians' Health Study II (PHS II)[9,10] and slowed age-related memory loss and cognitive decline among older males and females in the COcoa Supplement and Multivitamin Outcomes Study (COSMOS)[11–14]. COSMOS also reported a 3% nonsignificant risk reduction in the primary outcome of cancer and a significant 38% lower risk of lung cancer[15]. Short-term RCTs have also suggested that some individual MVM components (for example, B vitamins, vitamin D) may reduce epigenetic aging acceleration[16–20], but findings were inconsistent and limited by small sample size[21]. Whether a daily MVM affects biological aging based on DNAm has yet to be elucidated.

Cocoa flavanols have been shown to reduce the risk of many age-related chronic diseases and cardiometabolic outcomes[22]. In COSMOS, we found that cocoa flavanols reduced the secondary endpoint of cardiovascular death by 27% in the overall population[23] and restored hippocampal-dependent memory among those with lower diet quality or lower habitual flavanol consumption[24]. Small trials also suggested that flavanols could modify DNAm by inhibiting DNA methyltransferases through binding in the active site of the enzyme[16,25], but their effects on epigenetic biological aging remain unclear.

In this prespecified ancillary study, we investigated the effects of 2-year cocoa extract and MVM supplementation on DNAm measures of epigenetic aging in 958 healthy older adults who remained free of major morbidity from COSMOS. Based upon the effects of an MVM and cocoa flavanols on several aging-related outcomes and prior evidence of individual MVM components on epigenetic aging, our primary hypothesis is that the COSMOS interventions could delay biological aging measured by DNAm.

Results

COSMOS recruited a total of 21,442 women aged ≥65 years and men aged ≥60 years, and the COSMOS Blood subcohort included 6,867 participants who provided baseline biospecimens (Methods). A total of 1,829 participants provided their blood samples at baseline, year 1 and year 2 of follow-up. In this COSMOS ancillary study, we included a total of 2,993 samples from 998 participants who were randomly selected for the DNAm assessment, with oversampling of people of color and balanced by biological sex. Because we were specifically interested in the impact of the COSMOS intervention on healthy aging in older adults remaining free of major morbidity, we included participants free of major chronic disease at baseline who also did not experience incident cardiovascular disease (CVD; including myocardial infarction, stroke, coronary revascularization) or invasive cancer between the date of baseline blood drawn and 2 years after the blood collection period. Baseline characteristics suggested that participants included in the current study were representative of the overall COSMOS Blood subcohort and the overall COSMOS trial[26].

A total of 958 participants and 2,815 samples were included in our analyses (958 samples were from baseline, 934 from year 1 and 923 from year 2) (Fig. 1). Table 1 presents the baseline characteristics according to randomized assignments. The mean age was 70.2 ± 5.6 years, 482 (50.3%) participants were female and 854 (89.1%) were white. Four hundred seventy-six participants were randomized to the active

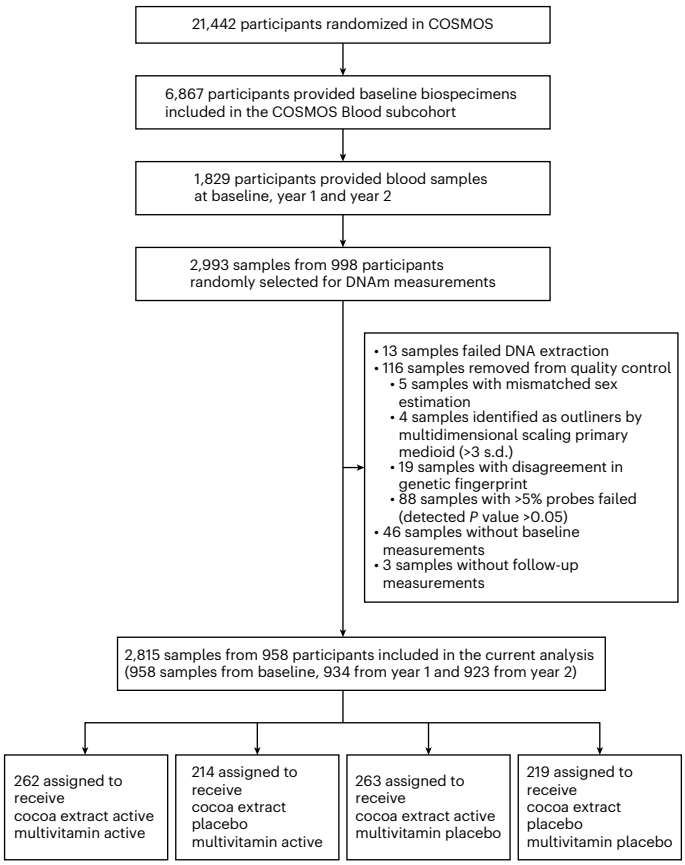


Fig. 1 | Flowchart of study enrollment. A total of 958 participants had available DNAm data at baseline and at least one follow-up assessment. For MVM supplementation, 482 were randomized to the placebo group and 476 to the active group. For cocoa extract, 433 were randomized to the placebo group and 525 to the active group.

MVM group, and 525 were randomized to the active cocoa extract group. There were no significant differences in baseline characteristics between randomized groups. At 2 years follow-up, the compliance rate remained high for both assignments, with 91.9% of participants adhering to randomization for MVM and 95.1% adhering to randomization for cocoa extract (defined as taking >75% of study pills and without using nonstudy pills).

In this COSMOS ancillary study, we prespecified epigenetic aging measures based on blood DNAm as primary outcome measures and we focused on five common DNAm measures of biological aging, including first-generation clocks predicting chronological age (Horvath[27] and Hannum[28]), second-generation clocks developed from analyses of mortality risk (PhenoAge[29] and GrimAge[30]) and third-generation clocks of epigenetic metrics developed to predict system integrity and Pace of Aging (DunedinPACE)[31]. All of these epigenetic clocks have been strongly associated with major morbidity and mortality[30–32]. We focused on epigenetic measures designed to maximize technical reliability; thus, we used the principal component (PC) versions with the method developed in ref. 33. Because DunedinPACE was initially created using CpGs with good replication performance to achieve high technical reliability, it was directly calculated using the method developed in ref. 31. The first two generations of clocks are estimates of biological age in 'years', and AgeDev was estimated by residualizing epigenetic clocks with respect to chronological age. DunedinPACE is expressed in pace-of-aging units, which are the rate of biological change with aging relative to a reference norm of 1 year of biological decline per chronological year. Measures are described in detail in the Methods. The correlations between epigenetic clocks and chronological age

https://doi.org/10.1038/s41591-026-04239-3

Table 1 | Baseline participant data stratified by the intervention arm in the COSMOS trial

	Overall (*n*=958)	Multivitamin		Cocoa extract	
		Placebo (*n*=482)	Active (*n*=476)	Placebo (*n*=433)	Active (*n*=525)
Demographics					
Age at randomization, years, mean±s.d.	70.2±5.6	70.3±5.7	70.0±5.6	69.7±5.6	70.3±5.7
Female (%)	482 (50.3)	250 (51.9)	232 (48.7)	210 (48.5)	272 (51.8)
Race (%)					
White	854 (89.1)	433 (89.8)	421 (88.4)	394 (91.0)	460 (87.6)
Black or African American	34 (3.5)	15 (3.1)	19 (4.0)	17 (3.9)	17 (3.2)
Hispanic	36 (3.8)	18 (3.7)	18 (3.8)	11 (2.5)	25 (4.8)
Native Hawaiian or Asian	21 (2.2)	11 (2.3)	10 (2.1)	4 (0.9)	17 (3.2)
Other or unknown	13 (1.4)	5 (1.0)	8 (1.7)	7 (1.6)	6 (1.1)
Lifestyle and health status					
Smoking (%)					
Never	556 (58.6)	275 (57.3)	281 (60.0)	252 (58.7)	304 (58.6)
Past	359 (37.9)	191 (39.8)	168 (35.9)	162 (37.8)	197 (38.0)
Current	33 (3.5)	14 (2.9)	19 (4.1)	15 (3.5)	18 (3.5)
Alcohol use (%)					
Rarely or never	239 (26.4)	132 (29.3)	107 (23.5)	109 (26.7)	130 (26.2)
Monthly	63 (7.0)	32 (7.1)	31 (6.8)	28 (6.8)	35 (7.0)
Weekly	340 (37.5)	164 (36.4)	176 (38.6)	140 (34.2)	200 (40.2)
Daily	264 (29.1)	122 (27.1)	142 (31.1)	132 (32.3)	132 (26.6)
Alternative health eating index, mean±s.d.	42.7±10.7	42.5±10.7	42.9±10.8	43.1±11.0	42.4±10.5
Body mass index, kg m^{-2}, mean±s.d.	27.2±5.2	27.3±5.2	27.2±5.1	27.3±5.4	27.1±5.0
Hypertension (%)	512 (53.6)	260 (54.1)	252 (53.2)	230 (53.5)	282 (53.7)
Diabetes (%)	124 (12.9)	61 (12.7)	63 (13.2)	65 (15.0)	59 (11.2)
Cholesterol-lowering medication (%)	398 (41.6)	199 (41.4)	199 (41.8)	177 (41.0)	221 (42.1)
Recruited from Women's Health Initiative (%)	103 (10.8)	53 (11.0)	50 (10.5)	41 (9.5)	62 (11.8)
Baseline epigenetic clocks, mean±s.d.					
PCHorvath	62.33±5.48	62.41±5.52	62.25±5.45	62.41±5.45	62.27±5.51
PCHannum	69.30±5.48	69.38±5.61	69.21±5.35	69.33±5.62	69.27±5.37
PCPhenoAge	65.50±6.60	65.55±6.72	65.45±6.48	65.52±6.58	65.48±6.62
PCGrimAge	76.62±4.84	76.59±5.00	76.66±4.67	76.40±4.78	76.81±4.88
DunedinPACE	0.987±0.105	0.987±0.104	0.987±0.106	0.988±0.106	0.987±0.104

at baseline are presented in Supplementary Fig. 1a. As expected, the first two generations of epigenetic clocks were moderately to highly correlated with chronological age (Spearman correlation coefficient, $r = 0.54–0.81$), and DunedinPACE was only weakly correlated with chronological age ($r = 0.13$). The unadjusted mean values for epigenetic clocks and their AgeDev residuals at each visit by randomized assignments are presented in Extended Data Table 1.

To test the effects of MVM and cocoa extract on biological aging, we conducted intention-to-treat (ITT) analysis to compare the change in epigenetic clocks from baseline between participants randomized to intervention versus placebo group using mixed linear models adjusting for chronological age, sex, baseline biological aging for that clock, recruitment source and the other randomized arm assignment. As expected, the first two generations of clocks significantly increased from baseline to 2-year follow-up in both MVM and placebo groups, and MVM significantly slowed the increase in PCPhenoAge (between-group difference, −0.443 [95% CI −0.848 to −0.037]) and PCGrimAge (−0.209 [95% CI −0.397 to −0.021]) at 2 years (Fig. 2 and Supplementary Table 1). For the first generation of clocks, the yearly change in PCHorvath (95% CI −0.019 [−0.127 to 0.089]; $P = 0.731$) and PCHannum (95% CI −0.058 [−0.195 to 0.079]; $P = 0.405$) did not differ between MVM and placebo groups (Table 2). For the second-generation clocks, a daily MVM significantly slowed epigenetic aging, with a between-group difference in yearly change of −0.214 years (95% CI −0.410 to −0.019; $P = 0.032$) for PCPhenoAge and −0.113 years (95% CI −0.205 to −0.020; $P = 0.017$) for PCGrimAge. The standardized mean difference (Cohen's *d*) was −0.038 (95% CI −0.074 to −0.003) for PCPhenoAge and −0.033 (95% CI −0.061 to −0.006) for PCGrimAge. As for DunedinPACE, it significantly increased among participants randomized to placebo but remained generally stable in the active MVM group. However, the yearly change in DunedinPACE did not significantly differ between groups (−0.003 [95% CI −0.007 to 0.001]; $P = 0.191$). Of note, although nonsignificant, the magnitude of between-group differences in DunedinPACE became larger from year 1 to year 2 (−0.004 [95% CI −0.012 to 0.003] at year 1 and −0.006 [95% CI −0.014 to 0.002] at year 2) (Fig. 2). The correlation between longitudinal changes in epigenetic clocks is presented in Supplementary Fig. 1b, and 2-year changes in PCPhenoAge and PCGrimAge were highly correlated ($r = 0.80$).

Compared with placebo, cocoa extract supplements did not significantly change any of these epigenetic clocks from baseline to either

https://doi.org/10.1038/s41591-026-04239-3

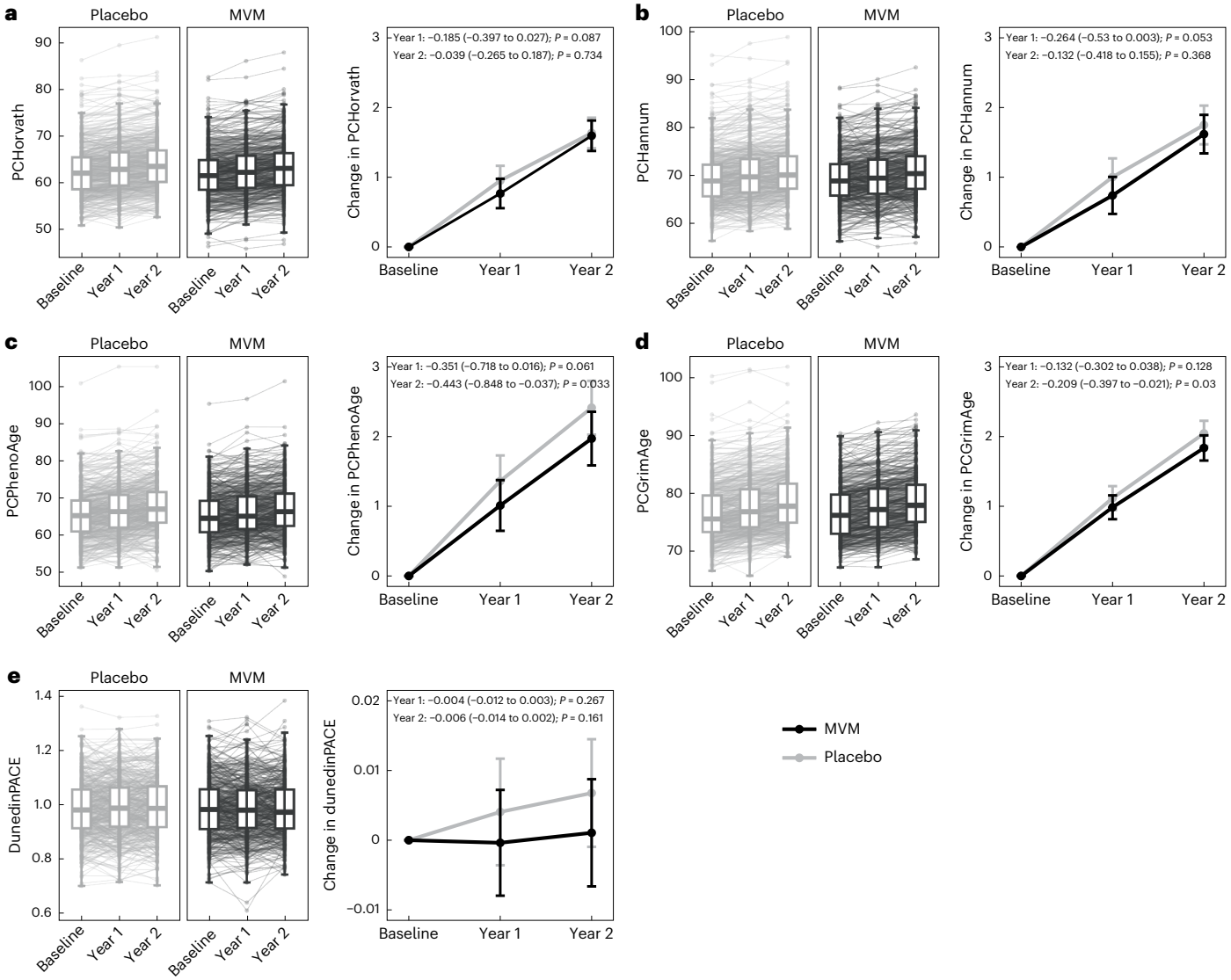


Fig. 2 | Changes from baseline to 1- and 2-year follow-ups in epigenetic clocks according to MVM supplementation and placebo groups in COSMOS. Longitudinal changes in epigenetic clocks by MVM assignments ($n = 482$ for placebo group; $n = 476$ for active group). **a**, PCHorvath. **b**, PCHannum. **c**, PCPhenoAge. **d**, PCGrimAge. Values are denominated in 'years' for the first two generations of epigenetic clocks, and the expected change under the null hypothesis is 1 year and 2 years at 1-year and 2-year follow-up, respectively. **e**, DunedinPACE. Values are denominated in pace-of-aging units, which are interpretable as the percentage difference in the rate of aging relative to the reference norm of 1 year of biological decline per chronological year; the expected change under the null hypothesis is zero. For each panel, the left-hand box plot shows the observed values of the epigenetic measures at baseline, year 1 and year 2. Individual participant data are plotted as dots and connected with lines. The box shows the median and interquartile range, and the whiskers show 1.5× the interquartile range. The right-hand line plot shows mean values of change from baseline and 95% CI values for each epigenetic clock estimated using mixed models for repeated measures, and the between-group differences comparing active intervention versus placebo at year 1 and year 2 are presented. Models included the treatment group, the time point of follow-up measurements (year 1 and year 2) and the treatment by follow-up time point interaction as fixed effects, and models were adjusted for the baseline biological aging from that clock, the follow-up time point by baseline age clock interaction, chronological age at baseline, sex, recruitment source, the other randomized arm and an individual-random intercept. Statistical significance with two-tailed *P* values was determined by the treatment by follow-up time point interaction without adjustment for multiple comparisons. Detailed data are shown in Supplementary Table 1.

year 1 or year 2 ($P > 0.05$) (Fig. 3). The between-group difference in yearly change was also not significant for PCHorvath, PCHannum, PCGrimAge or DunedinPACE, but cocoa extract significantly increased yearly change in PCPhenoAge (0.202 [95% CI 0.006 to 0.399]; $P = 0.044$) (Table 2).

In prespecified effect modification analyses, we were particularly interested in whether those with greater baseline biological versus chronological age (accelerated biological aging) would yield greater benefits from MVM use (Fig. 4 and Supplementary Table 2). For the first two generations of clocks, positive values of epigenetic age deviation residuals mean that biological age is older than chronological age, and vice versa. DunedinPACE corresponds to the number of biological years per chronological year, with values above 1 indicating a faster pace of aging. Among those with accelerated biological aging of PCGrimAge at baseline (PCGrimAge deviation >0), daily MVM use significantly reduced the rate of increases in PCGrimAge (between-group difference for MVM versus placebo, −0.236 [95% CI −0.380 to −0.091]) from baseline to year 2, while among those with lower baseline PCGrimAge deviation, there was no significant effect from daily MVM use (−0.013 [95% CI −0.130 to 0.104]; $P = 0.018$

https://doi.org/10.1038/s41591-026-04239-3

Table 2 | Yearly changes in epigenetic clocks in the intervention and placebo groups

	Yearly change from baseline (95% CI)		Difference (active versus placebo, 95% CI)	Cohen's *d* (95% CI)	*P* value
	Placebo	Active			
MVM					
PCHorvath	0.760 (0.684 to 0.836)	0.741 (0.647 to 0.834)	−0.019 (−0.127 to 0.089)	−0.003 (−0.024 to 0.017)	0.731
PCHannum	0.802 (0.705 to 0.898)	0.744 (0.625 to 0.862)	−0.058 (−0.195 to 0.079)	−0.011 (−0.039 to 0.016)	0.405
PCPhenoAge	1.116 (0.978 to 1.254)	0.902 (0.732 to 1.071)	−0.214 (−0.410 to −0.019)	−0.038 (−0.074 to −0.003)	0.032
PCGrimAge	1.017 (0.952 to 1.082)	0.904 (0.824 to 0.984)	−0.113 (−0.205 to −0.020)	−0.033 (−0.061 to −0.006)	0.017
DunedinPACE	0.003 (0 to 0.006)	0 (−0.003 to 0.004)	−0.003 (−0.007 to 0.001)	−0.026 (−0.065 to 0.013)	0.191
Cocoa extract					
PCHorvath	0.698 (0.618 to 0.778)	0.794 (0.699 to 0.889)	0.096 (−0.012 to 0.204)	0.018 (−0.003 to 0.038)	0.082
PCHannum	0.732 (0.630 to 0.833)	0.808 (0.687 to 0.929)	0.076 (−0.062 to 0.214)	0.015 (−0.013 to 0.042)	0.280
PCPhenoAge	0.899 (0.754 to 1.044)	1.101 (0.929 to 1.274)	0.202 (0.006 to 0.399)	0.036 (0.001 to 0.071)	0.044
PCGrimAge	0.920 (0.851 to 0.989)	0.995 (0.913 to 1.076)	0.074 (−0.019 to 0.167)	0.022 (−0.006 to 0.05)	0.117
DunedinPACE	0.001 (−0.002 to 0.004)	0.002 (−0.001 to 0.006)	0.001 (−0.003 to 0.005)	0.012 (−0.027 to 0.051)	0.548

Estimates were from mixed linear models using the change in epigenetic clocks as the outcome. Models were adjusted for chronological age at baseline, sex, baseline biological aging for that clock, recruitment source, the other randomization arm, and a random intercept of each individual. Cohen's *d* was estimated by scaling the repeated measures of age-difference values (that is, epigenetic clock age−chronological age) using the standard deviation of age-difference values at baseline for the first two generations of epigenetic clocks; DunedinPACE was directly scaled using the standard deviation at baseline. CI, confidence interval.

for interaction). The patterns were quite similar for other clocks, although the interaction was not statistically significant. MVM can significantly slow PCPhenoAge among participants with accelerated biological aging of PCPhenoAge (−0.394 [95% CI −0.675 to −0.112]) but not among those with lower baseline PCPhenoAge deviation (−0.075 [95% CI −0.340 to 0.191]; *P* for interaction = 0.106). Similarly, MVM significantly reduced DunedinPACE among those with a faster pace of aging at baseline (DunedinPACE >1: −0.007 [95% CI −0.013 to 0]) but not among those with normal or slower pace of aging (0 [95% CI −0.005 to 0.005]; *P* = 0.122 for interaction). In a post-hoc analysis among a subset of 166 COSMOS participants with available nutritional biomarkers[34], there was a pattern that lower levels of nutritional biomarkers were associated with accelerated biological aging at baseline, and larger improvements in nutritional biomarkers were inversely associated with epigenetic aging during follow-up (Extended Data Table 2). Further, a daily MVM may significantly increase folate (*P* = 0.03 for PCGrimAge) and lutein (*P* = 0.02 for DunedinPACE) levels among those with accelerated biological aging at baseline, along with a nonsignificant tendency for increases in other nutritional biomarkers, including zeaxanthin, vitamin B12 and vitamin D (Extended Data Table 3).

The effects of MVM were generally consistent across other prespecified subgroups by key baseline characteristics, including age, sex, race, body mass index (BMI), alternative healthy eating index (AHEI), smoking status, hypertension, diabetes and use of cholesterol-lowering medications (Extended Data Fig. 1). As for cocoa extract, the effects on PCHorvath, PCHannum and PCPhenoAge varied by baseline dietary quality measured by AHEI, with epigenetic aging significantly increased among participants with higher diet quality; other baseline characteristics did not significantly modify the effects of cocoa extract on any of these five epigenetic measures.

There was no evidence for interactions between MVM and cocoa extract supplementation for all epigenetic clocks (all *P* > 0.05 for interaction) (Extended Data Fig. 1). Compared with both placebo groups, the MVM active and cocoa extract placebo group could significantly slow biological aging measured by PCGrimAge (−0.137 [95% CI −0.275 to 0]; *P* = 0.050) and DunedinPACE (−0.006 [95% CI −0.012 to 0]; *P* = 0.043) (Extended Data Table 4). By contrast, the MVM placebo and cocoa extract active group did not significantly impact any epigenetic clocks. This suggested that the increased PCPhenoAge in the overall cocoa extract group may be partially explained by regression to the mean effects, because PhenoAge AgeDev measures were relatively lower at baseline in the cocoa extract group, although not significant.

In addition to primary analyses using the original clock values as the outcome, we reanalyzed the effects of MVM using AgeDev measures for the first two generations of clocks and results were not materially changed (Extended Data Table 5). The results were also similar when evaluating the effects in per-protocol analyses (Extended Data Table 6). We also evaluated the effect of MVM on different components of PCGrimAge (Extended Data Table 7), and MVM could significantly change PC-based DNAm measures of telomere length (0.022 [95% CI 0.003 to 0.042]; *P* = 0.027), beta-2 microglobulin (−0.028 [95% CI −0.045 to −0.010], *P* = 0.002), cystatin C (−0.027 [95% CI −0.044, −0.011]; *P* = 0.001) and growth differentiation factor 15 (GDF-15) (−0.029 [95% CI −0.046, −0.012]; *P* = 0.001).

To determine the biological relevance of the observed improvements in epigenetic clocks in response to a daily MVM, we also performed post-hoc exploratory analyses that tested the mediation effects through epigenetic clocks for the MVM intervention on cognition (COSMOS-Mind)[11,35] and inflammaging biomarkers[36] among participants with available DNAm data (Extended Data Tables 8 and 9). Among 117 COSMOS participants, PCPhenoAge had a significant inverse association with global cognition and executive function at baseline (*P* < 0.05). MVM led to a nonstatistically significant increase in episodic memory (standardized effect size, 0.087 ± 0.16; *P* = 0.59), although the effect size was comparable with the overall COSMOS-Mind study[11]. Mediation analyses showed that changes in PCPhenoAge and PCGrimAge may mediate 4.1–5.2% of MVM-related improvements in episodic memory, but these results did not reach statistical significance. For the COSMOS Inflammaging study, PCPhenoAge, PCGrimAge and DunedinPACE were each associated with baseline and longitudinal changes in high-sensitivity C-reactive protein, interleukin 6 and interleukin 10 (*P* < 0.05). Changes in PCGrimAge and PCPhenoAge explained 25.1–70.0% of MVM-related decreases in interleukin 6 and interleukin 10 at 2 years, but these mediation effects were not statistically significant.

Discussion

COSMOS represents a large-scale RCT that examined the effect of a daily MVM and cocoa extract supplementation on biological aging using longitudinal measurements of epigenetic clocks. We found that 2 years of daily MVM supplementation had a small but statistically

https://doi.org/10.1038/s41591-026-04239-3

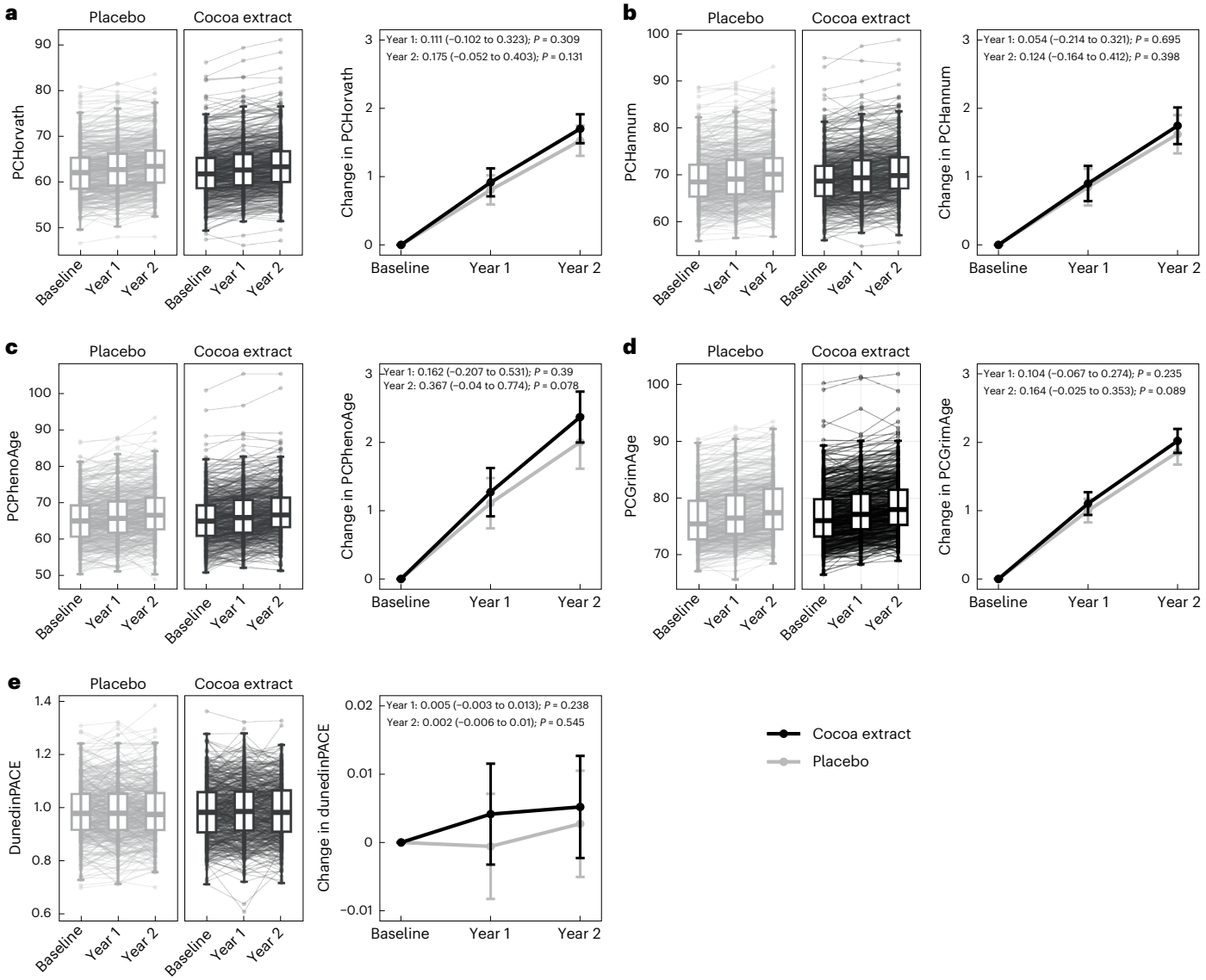


Fig. 3 | Changes from baseline to 1- and 2-year follow-ups in epigenetic clocks according to cocoa extract supplementation and placebo groups in COSMOS. Longitudinal changes in epigenetic clocks by cocoa extract assignments ($n = 433$ for placebo group; $n = 525$ for active group). **a**, PCHorvath. **b**, PCHannum. **c**, PCPhenoAge. **d**, PCGrimAge. Values are denominated in 'years' for the first two generations of epigenetic clocks, and the expected change under the null hypothesis is 1 year and 2 years at 1-year and 2-year follow-up, respectively. **e**, DunedinPACE. Values are denominated in pace-of-aging units, which are interpretable as percentage difference in the rate of aging relative to the reference norm of 1 year of biological decline per chronological year; the expected change under the null hypothesis is zero. For each panel, the left-hand box plot shows the observed values of the epigenetic measures at baseline, year 1 and year 2. Individual participant data are plotted as dots and connected with lines. The box shows the median and interquartile range, and the whiskers show 1.5× the interquartile range. The right-hand line plot shows the mean values of changes from baseline and 95% CIs for each epigenetic clock estimated using mixed models for repeated measures, and between-group differences comparing active intervention versus placebo and associated 95% CIs at year 1 and year 2 are presented in the figures. Models included the treatment group, the time point of follow-up measurements (year 1 and year 2), and the treatment by follow-up time point interaction as fixed effects, and models were adjusted for the baseline biological aging from that clock, the follow-up time point by baseline age clock interaction, chronological age at baseline, sex, recruitment source, the other randomized arm and an individual-random intercept. Statistical significance with two-tailed P values was determined by the treatment by follow-up time point interaction without adjustment for multiple comparisons. Detailed data are shown in Supplementary Table 1.

significant protective effect on delaying biological aging as represented by PCPhenoAge and PCGrimAge epigenetic clocks in older adults, but not for three other epigenetic clocks. Further, participants with higher biological aging at baseline may have a more pronounced benefit from daily MVM use, suggesting greater geroprotection among those at an accelerated aging process. However, we did not find evidence of delaying biological aging by cocoa extract as measured by these five epigenetic clocks. Further studies are needed to determine whether modest, statistically significant reductions in the rate of increase in epigenetic clocks after long-term MVM use ultimately translate into subsequent, meaningful clinical benefits.

Among generally healthy older adults, daily MVM use slowed biological aging as measured by PCGrimAge and PCPhenoAge, the second generation of epigenetic clocks developed to reflect the variations in biological aging among same-aged individuals by training on phenotypic biomarkers of mortality (PhenoAge)[29] or time to all-cause mortality (GrimAge)[30]. Prospective studies have reported associations of these two age clocks with a range of healthy lifespan metrics[30,37],

https://doi.org/10.1038/s41591-026-04239-3

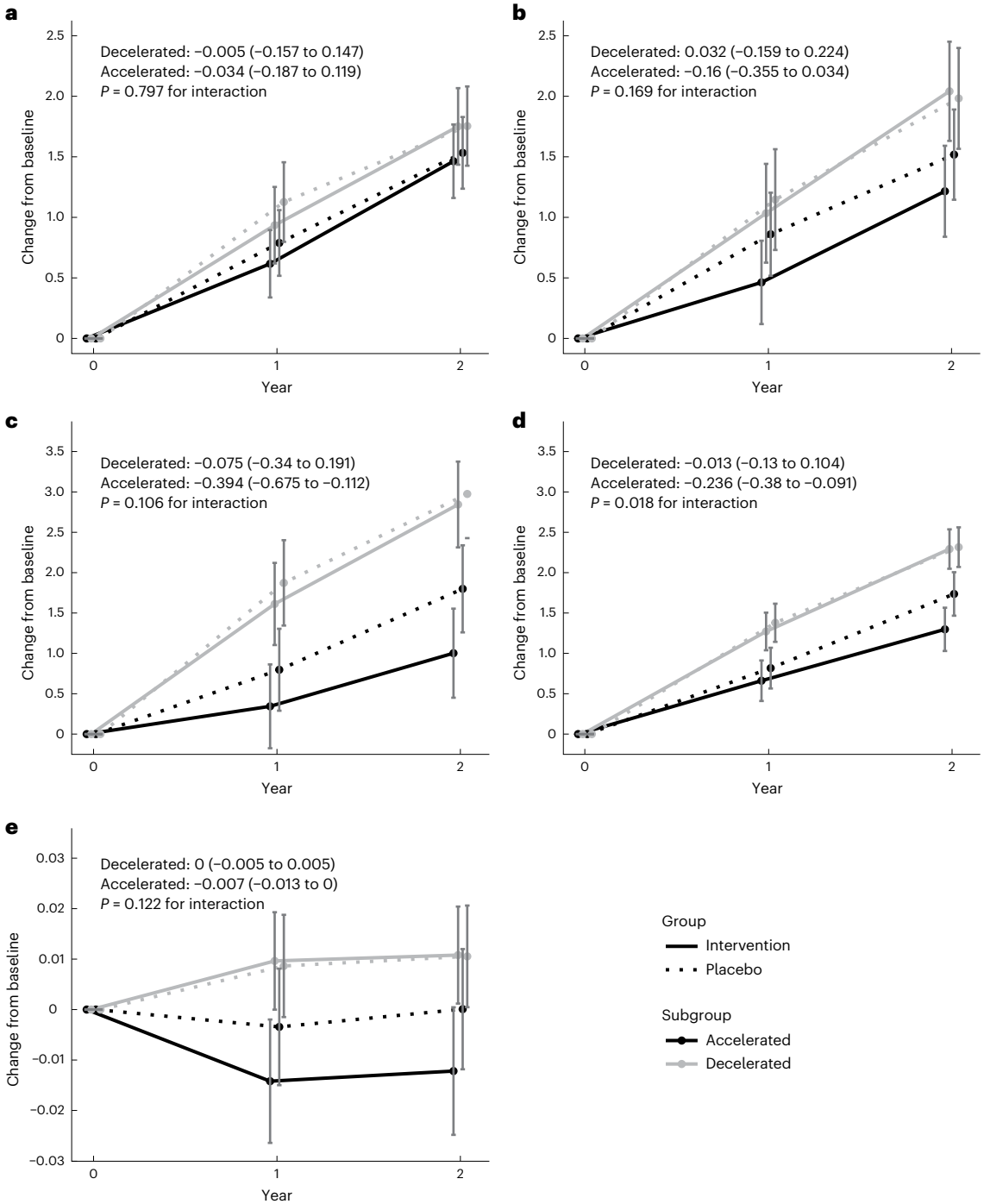


Fig. 4 | Changes in epigenetic clocks according to MVM supplementation and placebo groups by baseline values of the clocks. Mean values of changes from baseline and 95% CIs for each epigenetic clock by accelerated biological aging status (black line denotes accelerated aging and gray line denotes decelerated aging) and intervention group (solid line denotes MVM intervention group and dotted line denotes placebo group) estimated using mixed models for repeated measures. Between-group differences in yearly changes comparing active intervention versus placebo and associated 95% CIs stratified by baseline epigenetic aging acceleration are also presented. Estimates were determined from linear mixed-effects models, using changes in epigenetic clocks as the outcome. The models were adjusted for chronological age at baseline, sex, baseline biological aging for that clock, recruitment source, the other randomized arm and a random intercept for each individual. Statistical significance for the consistency between subgroups, as determined by two-tailed *P* values, was assessed through the three-way interactions of intervention, time and the stratified variables, without adjustment for multiple comparisons. **a**, PCHorvath. **b**, PCHannum. **c**, PCPhenoAge. **d**, PCGrimAge. Accelerated biological aging was defined as participants with corresponding age deviation residuals >0 at baseline for the first two generations of epigenetic clocks. **e**, DunedinPACE. Accelerated biological aging was defined as participants with DunedinPACE >1 at baseline. Detailed data are shown in Supplementary Table 2.

and they generally stand out for age-related conditions compared with other epigenetic clocks—especially with the PC version adopted in our study[33].

Based on associations reported in previous studies[30,37], the treatment effect of a daily MVM on slowing age acceleration in COSMOS (a yearly decrease of 0.113 years for PCGrimAge and 0.214 years for

PCPhenoAge) corresponds to an approximately 3–7% lower risk of cancer over 10 years, which was also similar to the significant effect previously reported in PHS II[9] and even the nonsignificant 3% reduction over 3.6 years reported in COSMOS[15]. Moreover, the effects on epigenetic aging were also comparable to MVM-related improvements in cognition observed in COSMOS (Cohen's d = 0.033–0.038 for epigenetic clocks versus 0.06–0.07 for global cognition and episodic memory)[11–14]. Our randomized trial findings in COSMOS for the effect of a daily MVM on epigenetic aging may have important implications on mortality, because a meta-analysis of MVM trials by the US Preventive Services Task Force reported odds ratios (95% CI) of mortality, cancer and lung cancer of 0.94 (0.87 to 1.01), 0.93 (0.87 to 0.99) and 0.75 (0.58 to 0.95), respectively[38]. Recent observational studies examining the association between self-reported MVM use and mortality remain highly susceptible to confounding and highlight the importance of RCTs[39].

Ultimately, it is critical to determine the clinical relevance of our findings; that is, do improvements in epigenetic clocks for those taking an MVM improve healthy aging outcomes such as cognition, biomarkers or other outcomes[11–15]. However, our study focused on COSMOS participants who remained free of major health outcomes. Among COSMOS participants who overlapped with other ancillary studies, we found promising but largely statistically nonsignificant signals that changes in PCPhenoAge and PCGrimAge in response to the MVM intervention may mediate a portion of MVM-related changes in both cognitive function and inflammaging biomarkers, another hallmark of aging[40]. Determining whether changes in epigenetic clocks from daily MVM use ultimately translate into subsequent clinical endpoints of aging-related chronic diseases requires larger sample sizes and longer follow-up. With continued follow-up and expanded methylation measurements, the COSMOS trial can definitively answer these important questions in future studies.

Our exploratory analyses of GrimAge components also suggested that MVM could significantly increase the DNAm-based biomarker of telomere length, another biomarker of aging, which was consistent with a prior cross-sectional analysis of 586 women aged 35–74 reporting significant associations between daily MVM use and longer telomere length[41]. Several DNAm-based key biomarkers in human aging were also significantly decreased, suggesting that MVM may have a broad and comprehensive impact on the aging process through modifying cellular senescence (telomere length), inflammation and immune system (beta-2 microglobulin), kidney function (cystatin C) and mitochondrial dysfunction (GDF-15). Considering that extending healthy lifespan by 1 year may save $38 trillion in the US population[42], our findings suggest that a daily MVM may represent a highly cost-efficient, accessible intervention to improve public health.

To date, only a few small-scale RCTs have tested limited combinations of vitamins and minerals on epigenetic aging with mixed findings[21], versus the broad-based MVM tested in COSMOS. Previous trials mainly focused on vitamin B (folic acid, vitamin B6 and vitamin B12)[16–18] and vitamin D[19]. One single-arm, 12-week RCT tested a supplement containing selected vitamins and nutraceuticals (including vitamin B3, vitamin C, vitamin D, omega 3 fish oils, resveratrol, olive fruit phenols and astaxanthin) among healthy older adults (31 men and 49 women, mean 71.9 ± 6.2 years)[20]. This trial found no significant effect on short-term changes in Horvath and Hannum clocks and age deviation measures, but there was a nonsignificant decline of 1.98 years in the Horvath clock among those with biological age acceleration over 2 years at baseline. Similarly, we found that a daily MVM was more beneficial for participants with accelerated biological aging at baseline, as observed for other dietary interventions[20,43]. A daily MVM exerted larger improvements in nutritional biomarkers among those with accelerated biological aging at baseline. Therefore, nutritional deficits might be the driving reason for biological aging acceleration at baseline, and the greater benefits from a daily MVM according to the clocks among age-accelerated participants at baseline may be driven by improvements in nutritional status. Because COSMOS participants were generally healthy and most already had adequate nutrient intake[15], the protective effect of MVM supplementation in poorly nourished populations may be more pronounced than that observed in COSMOS. Further work is required to elucidate how specific determinants influence epigenetic aging and those more likely to benefit from MVM use to facilitate individualized nutritional interventions aimed at promoting healthy lifespan. For example, the recently initiated CEDIRA trial (NCT06666660) will test a daily MVM (as Centrum) in 400 healthy middle-aged adults aged 40–59 years—specifically including those of greater biological than chronological age—on 12-month changes in epigenetic clocks.

In the current study, we found that the effect of MVM varied for these five updated and commonly used epigenetic clock measures, although the directions were quite consistent. Although these epigenetic clocks have each been identified as valid biomarkers of aging, they are composed of distinct methylation sites and designed to estimate different age-related measures, suggesting that they likely reflect unique aspects of aging processes[44–46]. Recent comparative studies also demonstrated that the first generation of clocks was sensitive to stochastic changes and less reliable in responding to aging interventions, whereas the clocks trained to predict mortality or pace of aging had the strongest responses across all interventions and showed consistent agreement with each other[47,48]. However, the potential clinical utility and relevance of changes in each epigenetic clock may also differ in trials testing different interventions. Recent data from the CALERIE trial reported a 2–3% (Cohen's d = 0.2–0.3) reduction in DunedinPACE after 2 years of caloric restriction, but not for GrimAge and PhenoAge[49]. In COSMOS, a daily MVM resulted in a comparatively more modest, but significant, yearly decrease of 0.113 years for PCGrimAge and 0.214 years for PCPhenoAge, corresponding to a decrease of 2.7–5.1 months after 2 years of intervention (Cohen's d = 0.03). For DunedinPACE, we reported a nonsignificant yearly reduction of 0.3%. The stronger effects found in CALERIE versus COSMOS are difficult to directly compare given differences in study populations (for example, CALERIE mean age, 38 years; COSMOS mean age, 70 years), intensity of the interventions tested and sample size. Moreover, DunedinPACE was developed to predict multi-system integrity decline over two decades of follow-up from early adulthood to midlife (aged 26–45 years) and quantify aging as a dynamic construct reflecting change in risk accumulation[31], which may limit its generalizability to older COSMOS participants with dynamic changes due to daily MVM use over 2 years of intervention. It may be difficult for a daily MVM to counter decades of decline in system integrity in 2 years among older adults who have already accumulated extensive molecular DNA changes[45,50]. The DO-HEALTH trial, a similar prevention trial in 777 older Swiss adults (mean age, 75.5 years), reported comparable effects for 3 years of omega 3 fatty acid supplementation, slowing GrimAge and PhenoAge by a similar 2.9–3.8 months and DunedinPACE by 1% with no effect for the first generation of clocks[51]. Taken together, the wide variability in populations and interventions tested among trials to date makes it difficult to clearly delineate the potential clinical utility and relevance of changes in each epigenetic clock.

Although cocoa extract has shown promising signals of reducing cardiovascular death by 27% and major cardiovascular events by 16% in COSMOS[23], we did not find evidence of delaying biological aging for cocoa extract as measured by these five epigenetic clocks. Given that all participants selected for DNAm were free of major chronic conditions, the effect of cocoa extract could be modest because they lack basal cardiometabolic disorders or inflammation status[52]. Similarly, the only small-scale and nonrandomized trial among 13 nonobese and healthy male smokers also found that an 8-week supplementation of monomeric and oligomeric flavanols derived from grape seeds did not significantly change the Horvath clock[16]. Previous mechanism studies suggested that cocoa flavanols can modify the DNAm profiles of certain

genes involved in vascular inflammatory processes[53] and may inhibit the activity of DNA methyltransferases and then lead to decreased global DNA methylation[16,54], which is inversely associated with CVD[55]. Of note, although epigenetic clocks were associated with CVD, a majority of CVD-related CpG sites were not covered by clocks (including Horvath, Hannum and GrimAge)[55]. Thus, the cardioprotective effects of cocoa extract may not be fully captured by these epigenetic clocks, and this also supports what we emphasized earlier, that the clinical utility of epigenetic clocks may vary by intervention.

The responsiveness of epigenetic measures in trials may vary by the epigenetic measures themselves, intensity and duration of the interventions tested, study population characteristics and baseline biological aging status, consistent with a recent integrated study that compared the pre- and post-intervention differences for various epigenetic measures from 51 small-scale aging intervention trials[47]. Nevertheless, based on the consistency in effect sizes in epigenetic biomarkers and long-term benefits measured by clinical endpoints after MVM intervention, our study supports the use of epigenetics as surrogate outcomes for the anti-aging effects of an intervention. Our findings also emphasize the need to account for heterogeneity in treatment response based on initial factors to more precisely identify those most likely to benefit from the intervention.

We assessed the 2-year intervention effects of MVM and cocoa extract on multiple epigenetic aging clocks based on COSMOS, a large, randomized, placebo-controlled trial with high compliance and follow-up. Nevertheless, several potential limitations should be considered. First, this study included only older adults, and most were non-Hispanic white, limiting our generalizability to other populations. Second, we did not apply multiple comparison adjustments in this study because such corrections assume test independence, which is not applicable here because of the high intercorrelation across these epigenetic clocks[47]. Applying these adjustments could increase the risk of excluding true positives. Although false positives are likely to occur, we observed generally consistent directions for different DNA measures and the effect sizes also aligned with our previous findings showing the protective effects for age-related comorbidities and functional decline, suggesting the effect on slowing biological aging by MVM may not be explained by chance. Third, any random measurement error caused by technical limitations of the DNAm assays may bias our effect estimates toward the null[33]. Thus, the effect of a daily MVM on biological aging as measured by our included epigenetic clocks might have been underestimated. Fourth, although we included updated and commonly used epigenetic clocks representing multiple generations of biological aging, all biological changes occurring with age have yet to be fully captured by these DNAm measures[56]. Emerging epigenetic clocks have included more physiological biomarkers[57] or quantify aging heterogeneity across different physiological systems[58], but their utility remains to be fully evaluated in geroprevention trials. We will further explore the effects of a daily MVM on biological aging by applying more advanced epigenetic measures. Fifth, our study sample was not equally distributed by cocoa extract arm, which may be related to the exclusion of participants who had outcomes at baseline and during follow-up. Sixth, although we assessed a 2-year effect in the current study, it remains unclear whether a longer intervention duration would have a greater effect on slowing the aging process. Seventh, although consistent effect sizes of MVM on epigenetic clocks and selected clinical outcomes such as inflammaging and episodic memory were observed, power was limited for outcome-based mediation analyses in these initially healthy COSMOS participants. Further studies to expand DNAm analyses to COSMOS participants with stored longitudinal biospecimens who develop subsequent incident clinical outcomes can directly address the question of whether MVM-driven or cocoa extract-driven changes in epigenetic clocks over time explain clinical outcomes.

In conclusion, we provide evidence from a large-scale and long-term RCT that a daily MVM is a safe, readily accessible and low-cost intervention that may slow epigenetic aging measured by PhenoAge and GrimAge among older adults, and especially among those with accelerated biological aging at baseline. Cocoa extract did not impact these epigenetic aging measurements. Our effects for an MVM, as measured by epigenetic clocks, appear to align with benefits measured by clinical endpoints in both COSMOS and PHS II. COSMOS is an important prevention trial designed to evaluate whether the MVM intervention prevents, delays or alleviates aging, and additional trials are needed to confirm these findings and determine the role of MVM in extending healthy aging not only among older adults, but also across the lifespan.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-026-04239-3.

References

1. Wu, Z., Zhang, W., Qu, J. & Liu, G. H. Emerging epigenetic insights into aging mechanisms and interventions. *Trends Pharmacol. Sci.* **45**, 157–172 (2024).
2. Kennedy, B. K. et al. Geroscience: linking aging to chronic disease. *Cell* **159**, 709–713 (2014).
3. Campisi, J. et al. From discoveries in ageing research to therapeutics for healthy ageing. *Nature* **571**, 183–192 (2019).
4. Sandalova, E. & Maier, A. B. Targeting the epigenetically older individuals for geroprotective trials: the use of DNA methylation clocks. *Biogerontology* **25**, 423–431 (2024).
5. Moqri, M. et al. Biomarkers of aging for the identification and evaluation of longevity interventions. *Cell* **186**, 3758–3775 (2023).
6. Passarelli, S. et al. Global estimation of dietary micronutrient inadequacies: a modelling analysis. *Lancet Glob. Health.* **12**, e1590–e1599 (2024).
7. Rautiainen, S., Manson, J. E., Lichtenstein, A. H. & Sesso, H. D. Dietary supplements and disease prevention – a global overview. *Nat. Rev. Endocrinol.* **12**, 407–420 (2016).
8. Yetley, E. A. Multivitamin and multimineral dietary supplements: definitions, characterization, bioavailability, and drug interactions. *Am. J. Clin. Nutr.* **85**, 269s–276s (2007).
9. Gaziano, J. M. et al. Multivitamins in the prevention of cancer in men: the Physicians' Health Study II randomized controlled trial. *JAMA* **308**, 1871–1880 (2012).
10. Christen, W. G. et al. Effects of multivitamin supplement on cataract and age-related macular degeneration in a randomized trial of male physicians. *Ophthalmology* **121**, 525–534 (2014).
11. Baker, L. D. et al. Effects of cocoa extract and a multivitamin on cognitive function: a randomized clinical trial. *Alzheimers Dement.* **19**, 1308–1319 (2023).
12. Sachs, B. C. et al. Impact of multivitamin–mineral and cocoa extract on incidence of mild cognitive impairment and dementia: results from the COcoa Supplement and Multivitamin Outcomes Study for the Mind (COSMOS-Mind). *Alzheimers Dement.* **19**, 4863–4871 (2023).
13. Yeung, L. K. et al. Multivitamin supplementation improves memory in older adults: a randomized clinical trial. *Am. J. Clin. Nutr.* **118**, 273–282 (2023).
14. Vyas, C. M. et al. Effect of multivitamin–mineral supplementation versus placebo on cognitive function: results from the clinic subcohort of the COcoa Supplement and Multivitamin Outcomes Study (COSMOS) randomized clinical trial and meta-analysis of 3 cognitive studies within COSMOS. *Am. J. Clin. Nutr.* **119**, 692–701 (2024).

15. Sesso, H. D. et al. Multivitamins in the prevention of cancer and cardiovascular disease: the COcoa Supplement and Multivitamin Outcomes Study (COSMOS) randomized clinical trial. *Am. J. Clin. Nutr.* **115**, 1501–1510 (2022).
16. Sae-Lee, C. et al. Dietary intervention modifies DNA methylation age assessed by the epigenetic clock. *Mol. Nutr. Food Res.* **62**, e1800092 (2018).
17. Obeid, R., Hübner, U., Bodis, M., Graeber, S. & Geisel, J. Effect of adding B-vitamins to vitamin D and calcium supplementation on CpG methylation of epigenetic aging markers. *Nutr. Metab. Cardiovasc. Dis.* **28**, 411–417 (2018).
18. Michels, K. B. & Binder, A. M. Impact of folic acid supplementation on the epigenetic profile in healthy unfortified individuals – a randomized intervention trial. *Epigenetics* **19**, 2293410 (2024).
19. Chen, L. et al. Effects of vitamin D3 supplementation on epigenetic aging in overweight and obese African Americans with suboptimal vitamin D status: a randomized clinical trial. *J. Gerontol. A Biol. Sci. Med. Sci.* **74**, 91–98 (2019).
20. McGee, K. C. et al. A combination nutritional supplement reduces DNA methylation age only in older adults with a raised epigenetic age. *Geroscience* **46**, 4333–4347 (2024).
21. García-García, I. et al. Examining nutrition strategies to influence DNA methylation and epigenetic clocks: a systematic review of clinical trials. *Front. Aging* **5**, 1417625 (2024).
22. Martin, M. Á & Ramos, S. Impact of cocoa flavanols on human health. *Food Chem. Toxicol.* **151**, 112121 (2021).
23. Sesso, H. D. et al. Effect of cocoa flavanol supplementation for the prevention of cardiovascular disease events: the COcoa Supplement and Multivitamin Outcomes Study (COSMOS) randomized clinical trial. *Am. J. Clin. Nutr.* **115**, 1490–1500 (2022).
24. Brickman, A. M. et al. Dietary flavanols restore hippocampal-dependent memory in older adults with lower diet quality and lower habitual flavanol consumption. *Proc. Natl Acad. Sci. USA* **120**, e2216932120 (2023).
25. Fang, M. Z. et al. Tea polyphenol (–)-epigallocatechin-3-gallate inhibits DNA methyltransferase and reactivates methylation-silenced genes in cancer cell lines. *Cancer Res.* **63**, 7563–7570 (2003).
26. Rist, P. M. et al. Design and baseline characteristics of participants in the COcoa Supplement and Multivitamin Outcomes Study (COSMOS). *Contemp. Clin. Trials* **116**, 106728 (2022).
27. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, R115 (2013).
28. Hannum, G. et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* **49**, 359–367 (2013).
29. Levine, M. E. et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)* **10**, 573–591 (2018).
30. Lu, A. T. et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging (Albany NY)* **11**, 303–327 (2019).
31. Belsky, D. W. et al. DunedinPACE, a DNA methylation biomarker of the pace of aging. *Elife* **11**, e73420 (2022).
32. Faul, J. D. et al. Epigenetic-based age acceleration in a representative sample of older Americans: associations with aging-related morbidity and mortality. *Proc. Natl Acad. Sci. USA* **120**, e2215840120 (2023).
33. Higgins-Chen, A. T. et al. A computational solution for bolstering reliability of epigenetic clocks: implications for clinical trials and longitudinal tracking. *Nat. Aging* **2**, 644–661 (2022).
34. Christopher, C. et al. Effects of randomized multivitamin supplementation on carotenoids and alpha-tocopherol in the COSMOS trial. *J. Acad. Nutr. Diet.* https://doi.org/10.1016/j.jand.2026.156299 (2026).
35. Baker, L. D. et al. Design and baseline characteristics of the cocoa supplement and multivitamin outcomes study for the mind: COSMOS-Mind. *Contemp. Clin. Trials* **83**, 57–63 (2019).
36. Li, S. et al. Effects of 2-year cocoa extract supplementation on inflammaging biomarkers in older US adults: findings from the COcoa Supplement and Multivitamin Outcomes Study randomised clinical trial. *Age Aging* **54**, afaf269 (2025).
37. Chervova, O. et al. Breaking new ground on human health and well-being with epigenetic clocks: a systematic review and meta-analysis of epigenetic age acceleration associations. *Ageing Res. Rev.* **102**, 102552 (2024).
38. O'Connor, E. A. et al. Vitamin and mineral supplements for the primary prevention of cardiovascular disease and cancer: updated evidence report and systematic review for the US Preventive Services Task Force. *JAMA* **327**, 2334–2347 (2022).
39. Loftfield, E. et al. Multivitamin use and mortality risk in 3 prospective US cohorts. *JAMA Netw. Open* **7**, e2418729 (2024).
40. Lopez-Otin, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. Hallmarks of aging: an expanding universe. *Cell* **186**, 243–278 (2023).
41. Xu, Q. et al. Multivitamin use and telomere length in women. *Am. J. Clin. Nutr.* **89**, 1857–1863 (2009).
42. Scott, A. J., Ellison, M. & Sinclair, D. A. The economic value of targeting aging. *Nat. Aging* **1**, 616–623 (2021).
43. Gensous, N. et al. One-year Mediterranean diet promotes epigenetic rejuvenation with country- and sex-specific effects: a pilot study from the NU-AGE project. *Geroscience* **42**, 687–701 (2020).
44. Oblak, L., van der Zaag, J., Higgins-Chen, A. T., Levine, M. E. & Boks, M. P. A systematic review of biological, social and environmental factors associated with epigenetic clock acceleration. *Ageing Res. Rev.* **69**, 101348 (2021).
45. Horvath, S. & Raj, K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat. Rev. Genet.* **19**, 371–384 (2018).
46. Yusipov, I., Kalyakulina, A., Trukhanov, A., Franceschi, C. & Ivanchenko, M. Map of epigenetic age acceleration: a worldwide analysis. *Ageing Res. Rev.* **100**, 102418 (2024).
47. Sehgal, R. et al. DNAm aging biomarkers are responsive: insights from 51 longevity interventional studies in humans. Preprint at *bioRxiv* https://doi.org/10.1101/2024.10.22.619522 (2024).
48. Borrus, D. S. et al. When to trust epigenetic clocks: avoiding false positives in aging interventions. Preprint at *bioRxiv* https://doi.org/10.1101/2024.10.22.619720 (2024).
49. Waziry, R. et al. Effect of long-term caloric restriction on DNA methylation measures of biological aging in healthy adults from the CALERIE trial. *Nat. Aging* **3**, 248–257 (2023).
50. Marioni, R. E. et al. Tracking the epigenetic clock across the human life course: a meta-analysis of longitudinal cohort data. *J. Gerontol. A Biol. Sci. Med. Sci.* **74**, 57–61 (2019).
51. Bischoff-Ferrari, H. A. et al. Individual and additive effects of vitamin D, omega-3 and exercise on DNA methylation clocks of biological aging in older adults from the DO-HEALTH trial. *Nat. Aging* **5**, 376–385 (2025).
52. Ellinger, S. & Stehle, P. Impact of cocoa consumption on inflammation processes – a critical review of randomized controlled trials. *Nutrients* **8**, 321 (2016).
53. Milenkovic, D. et al. Flavanol consumption in healthy men preserves integrity of immunological–endothelial barrier cell functions: nutri(epi)genomic analysis. *Mol. Nutr. Food Res.* **66**, e2100991 (2022).
54. Crescenti, A. et al. Cocoa consumption alters the global DNA methylation of peripheral leukocytes in humans with cardiovascular disease risk factors: a randomized controlled trial. *PLoS ONE* **8**, e65744 (2013).

55. Krolevets, M. et al. DNA methylation and cardiovascular disease in humans: a systematic review and database of known CpG methylation sites. *Clin. Epigenetics* **15**, 56 (2023).
56. Bell, C. G. et al. DNA methylation aging clocks: challenges and recommendations. *Genome Biol.* **20**, 249 (2019).
57. Bernabeu, E. et al. Refining epigenetic prediction of chronological and biological age. *Genome Med.* **15**, 12 (2023).
58. Sehgal, R. et al. Systems Age: a single blood methylation test to quantify aging heterogeneity across 11 physiological systems. *Nat. Aging* **5**, 1880–1896 (2025).

Sidong Li[1,2], Rikuta Hamaya[2], Haidong Zhu[3], Brian H. Chen[4], Alexandre C. Pereira[5], Kerry L. Ivey[5], Pamela M. Rist[2,6], JoAnn E. Manson[2,6,7], Yanbin Dong[3] & Howard D. Sesso[2,5,6] ✉

[1]Institute of Public Health Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China. [2]Division of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA. [3]Georgia Prevention Institute, Medical College of Georgia, Augusta University, Augusta, GA, USA. [4]California Pacific Medical Center Research Institute, San Francisco, CA, USA. [5]Division of Aging, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA. [6]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [7]Mary Horrigan Connors Center for Women's Health Research, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA. ✉e-mail: hsesso@bwh.harvard.edu

Methods

Study design and participants

The COSMOS study design, baseline characteristics and main results have been published previously[15,23,26]. In brief, COSMOS was a large, randomized, double-blind, placebo-controlled, 2 × 2 factorial trial testing the effects of a cocoa extract supplement (500 mg cocoa flavanols per day, including 80 mg (−)-epicatechin; provided by Mars Edge) and an MVM supplement (Centrum Silver; provided by Pfizer Consumer Healthcare) on CVD and cancer in middle-aged or older US adults (registration: NCT02422745). COSMOS included women aged ≥65 years and men aged ≥60 years, and recruitment for COSMOS included mass mailings to active Women's Health Initiative (WHI) Extension Study participants, to people who were contacted for but not randomized into the VITamin D and OmegA-3 TriaL (VITAL)[59], and through mass mailings, advertisements and media campaigns. Eligibility for the main COSMOS trial included: (1) no history of myocardial infarction or stroke; (2) no history of invasive cancer (except nonmelanoma skin cancer) in the past 2 years; (3) no serious illness precluding participation (including renal failure and dialysis); (4) willingness to forgo supplements of cocoa extract and multivitamins (chocolate intake was not restricted) and limit vitamin D to ≤1,000 IU d^{-1} and calcium to ≤1,200 mg d^{-1} from all supplemental sources during the trial; (5) no reported extreme sensitivity to cocoa products or caffeine; and (6) successful completion of at least a 2-month placebo run-in phase with greater than 75% study pill adherence. Between April 2016 and March 2018, 21,442 participants were randomized into one of four treatment groups: (1) active cocoa extract and an active daily MVM; (2) active cocoa extract and multivitamin placebo; (3) active multivitamin and cocoa extract placebo; or (4) both placebos, using a computer-generated permuted block approach blinded to investigators and stratified by sex (female, male), age (separate 5-year age blocks for females and males) and recruitment source (WHI, Brigham and Women's Hospital (BWH)) in blocks of 12. Participants in a household were randomly assigned to the same intervention, when possible, to reduce cross-contamination risk. BWH staff sent randomized participants calendar packs containing cocoa extract or placebo capsules and multivitamin or placebo tablets. The COSMOS intervention continued to 31 December 2020, with a median treatment period of 3.6 years. All participants provided written informed consent before enrollment in the trial, and trial activities were overseen by the Human Subjects Committee at BWH/Mass General Brigham (2014P002768). There was no participant or public involvement in the design, conduct and reporting of the COSMOS trial.

In the COSMOS Blood subcohort, 1,829 participants provided blood samples at baseline, year 1 and year 2 of follow-up. Blood biospecimens were collected either through mailed biospecimen kits returned by the participant or by home-based visits by Examination Management Services, Inc. All samples were returned overnight to our biorepository, where they were immediately processed and stored at −170 °C.

This ancillary study (NCT05510375) was prespecified before the data unlock and completion of the parent COSMOS trial. A total of 998 participants were randomly selected for DNAm assessment, with oversampling of participants of color and balanced by biological sex. We originally based our analyses upon 600 participants through one funded project plus participants through additionally obtained support. Because we were specifically interested in the effects of the MVM and cocoa extract interventions among healthy older adults, all participants selected were free of major chronic conditions at baseline and did not experience CVD (myocardial infarction, stroke, coronary revascularization) or invasive cancer between the date of baseline blood drawn and 2 years after the blood collection period. This minimized the influence of major changes in health status on our longitudinal epigenetic measurements. Participants included in the current study were generally representative of the overall COSMOS Blood subcohort and the overall COSMOS trial[26].

DNAm assessment

Physical human buffy coat samples and an accompanying sample manifest were shipped from BWH to Neogen Corporation. The laboratory extracted DNA from the frozen buffy coat samples. Samples were randomly arranged on 96-well plates before extraction to minimize biases due to any batch effects downstream. The DNA samples were bisulfite converted using the EZ-96DNA Methylation kit (Zymo Research). After bisulfite conversion and whole-genome amplification, DNA was hybridized onto a custom-designed Infinium EPIC+ Array (Illumina), which quantifies DNAm levels at >850,000 CpG sites across all known genes, regions and key regulatory regions. The methylation assay was performed according to the manufacturer's explicit specifications, using the Freedom EVO (Tecan) for the staining steps. Probe-level pairwise comparison between 42 technical duplicates from 14 participants showed excellent reliability for DNAm measurements (mean value: 0.992 ± 0.008 for Pearson correlation; 0.996 ± 0.016 for R square; 0.002 ± 0.002 for mean absolute error; 0.027 ± 0.013 for mean squared error).

Quality control and normalization analyses were performed using the R Bioconductor package SeSAMe (R v.4.3.0; Bioconductor v.3.17; SeSAMe v.1.18.4)[60]. Normalization to eliminate systematic dye bias in two-channel probes was carried out using the SeSAMe default pipeline. Probes were considered missing in a sample if they had detection P value >0.05 and were excluded from the analysis if they were missing in >5% of sample. Sample-level quality control excluded participants with mismatched sex estimation (estimated by a custom Python function developed to mirror the getSex function of the R Bioconductor package minfi)[61], >5% of CpGs with detection P values >0.05, identified as outliners by multidimensional scaling primary medoid (>3 s.d.) or with disagreement in genetic fingerprint. Following quality control and normalization, DNAm data for 822,759 CpGs were available for 2,815 samples (baseline, n = 958; year 1, n = 934; year 2, n = 923). Missing beta values were imputed using K nearest neighbor imputation, and batch effects were additionally identified using singular value decomposition analysis and then corrected using Combat in sva R package[62,63].

DNAm measures of biological aging

DNAm clocks are algorithms that combine information from DNAm measurements across the genome to quantify variation in biological age as epigenetic clocks[41]. For this study, we focused on five common DNAm measures of biological aging, including the first-generation clocks predicting chronological age (Horvath[27] and Hannum[28]), second-generation clocks predicting mortality outcomes (PhenoAge[29] and GrimAge[30]) and third-generation clocks of epigenetic metrics developed to predict system integrity and Pace of Aging (DunedinPACE)[31]. All of these epigenetic clocks have been strongly associated with major morbidity and mortality[30–32]. To ensure that values of the first two generations of clocks were highly reproducible, we used the PC versions using the method developed previously[33]. Epigenetic age deviation (also sometimes referred as age acceleration) was estimated by residualizing these clocks for chronological, with positive values meaning accelerated aging and vice versa[5,27]. Because DunedinPACE was initially created using CpGs with good replication performance to achieve high technical reliability, it was directly calculated using the method developed in ref. 31. DunedinPACE corresponds to the number of biological years per chronological year, with values above 1 indicating a faster pace of aging. The within-individual stability across repeated measurements was measured using intra-class correlation coefficients, with 0.937 for PCHorvath, 0.910 for PCHannum, 0.919 for PCGrimAge, 0.877 for PCPhenoAge and 0.818 for DunedinPACE.

Other covariates

All participants completed baseline questionnaires with self-reports of demographics, lifestyle, clinical and dietary factors via a

semi-quantitative food-frequency questionnaire. Baseline characteristics included age, sex (male, female), race or ethnicity (white, people of color), smoking status (never, former, current), alcohol intake (rarely or never, monthly, weekly or daily), hypertension, diabetes and use of cholesterol-lowering medications. Diet quality was assessed using AHEI, a well-validated score with a higher score representing a healthier diet[64].

Statistical analysis

The current study followed the ITT principle and included all participants with available DNAm data at baseline and at least one follow-up time point. To assess the effect of the intervention on changes in each epigenetic clock, we used linear mixed-effect models to estimate the yearly change in epigenetic clock values in active and placebo groups, respectively, and their between-group differences. Models were adjusted for chronological age at baseline, sex, baseline biological aging for that clock, recruitment source and the other intervention as fixed covariates with an interaction between randomized intervention arm and time (year 0, 1 or 2). Standardized effect sizes (Cohen's *d*) were estimated by scaling the repeated measures of age-difference values (that is, epigenetic clock age − chronological age) using the standard deviation of the difference between clock age and chronological age at baseline for the first two generations of epigenetic clocks. For DunedinPACE, we directly scaled the values using the standard deviation at baseline. In addition, the between-group differences at year 1 and year 2 were estimated using mixed models for repeated measurements with an unstructured covariance matrix[34].

Meanwhile, we explored effect modification using the following prespecified baseline covariates: age (<70 or ≥70 years), sex (male, female), race (white or people of color), BMI (<30 or ≥30 kg m^{-2}), AHEI (tertiles), smoking status (ever or never), hypertension (yes or no), diabetes (yes or no), use of cholesterol-lowering medications (yes or no), baseline levels of biological aging from that epigenetic clock (accelerated or decelerated) and the other intervention arm. Coefficients of the three-way interaction terms of intervention arm, time and each effect modifier tested the statistical significance of the effect measure modifications.

In sensitivity analysis, we also assessed the effect on the age deviation residuals for the first two generations of age clocks. In addition to the ITT analysis, we conducted a per-protocol analysis by adjusting for selection bias due to lost follow-up and nonadherence. A participant was nonadherent when they discontinued trial pills, started using nonstudy supplements and/or took <75% of study pills. Per-protocol effects of dietary supplementation on changes in epigenetic clocks were estimated using inverse probability weight-adjusted linear mixed-effects models. Considering the epigenetic clocks were highly intercorrelated, multiple comparison adjustments were not applied in this study to avoid the risk of excluding true positives. We focused on identifying consistent patterns in which yearly changes had 95% CI values that did not include zero. Two-sided $P < 0.05$ was considered statistically significant. Analyses were performed using R v.4.4.2 (R Foundation).

Mediation analyses and other post-hoc analyses

To determine the clinical relevance of our findings of MVM on epigenetic clocks, we conducted post-hoc analyses to explore the mediation effects through epigenetic clocks for the MVM intervention on cognition and inflammaging biomarkers among participants overlapped with two other COSMOS ancillary studies. COSMOS-Mind (NCT03035201) assessed cognition by telephone among 2,262 participants at baseline and annually during 3 years of follow-up[11,35]. The primary outcome of COSMOS-Mind was a global cognition composite (reported as a *z*-score based on baseline data from the study cohort), with higher scores reflecting better performance. Key secondary outcomes included an episodic memory composite and an executive function composite. The effects of MVM (intervention) on 3-year changes in cognitive function (outcomes) that were mediated through 2-year changes in epigenetic measures (mediators) were examined based on 117 participants with complete data on cognitive function and epigenetic measures at both baseline and follow-up. This COSMOS Inflammaging ancillary study (NCT05510375) measured five inflammaging biomarkers among a randomly selected sample of 598 participants at baseline, year 1 and year 2[36]. The effects of a MVM (exposure) on 2-year changes in inflammaging biomarkers (outcomes) that were mediated through 2-year changes in epigenetic measures (mediators) were examined among 573 participants with complete data on inflammaging biomarkers and epigenetic measures at both baseline and year 2. Inflammaging biomarkers were log-transformed and then standardized. Mediation analyses for an MVM were performed based on a counterfactual framework by the imputation-based approach using the R package medflex[65]. Models were adjusted for age, sex, recruitment source, baseline outcome measures (cognition or inflammaging biomarkers), smoking status, diet quality, hypertension, diabetes, BMI and randomized cocoa extract assignment. We also evaluated the association of epigenetic measures with inflammaging biomarkers or cognitive function measures at baseline and during 2 years of follow-up using linear mixed-effect models, adjusting for the same covariates as aforementioned. Meanwhile, to explain the possible underlying reason for the protective effect of MVM on epigenetic clocks, we also explored the longitudinal associations of nutritional biomarkers with epigenetic aging measures and then tested the effect of MVM on nutritional biomarkers stratified by baseline epigenetic acceleration aging status among a subset of 166 COSMOS participants with available nutritional biomarkers (carotenoids, vitamin B_{12}, folate, vitamin D and alpha-tocopherol)[34]. There was no formal adjustment to the *P* values for multiple testing in these exploratory analyses; thus, results should be interpreted with caution.

Epigenetic clocks of biological aging

The first-generation epigenetic clocks were developed by comparing samples from individuals of different chronological ages to predict chronological age[45]. These clocks were highly accurate in predicting the chronological age and also showed some capacity for predicting differences in mortality risk, although effect sizes tend to be small and inconsistent across studies[66]. The Horvath multi-tissue clock and the Hannum blood-based clock were the most well-known first-generation clocks[27,28].

The second-generation epigenetic clocks were developed with the goal of improving the quantification of biological aging by focusing on differences in mortality risk instead of on differences in chronological age. They were more predictive of morbidity and mortality compared with the first-generation clocks and were proposed to have improved potential for testing the impacts of interventions to slow aging[67,68]. The age values computed by the second-generation clocks correspond to the age at which predicted mortality risk would be approximately normal in the reference population used to develop the clock. We analyzed the second-generation PhenoAge clock and GrimAge clock[29,30]. These clocks include an intermediate step in which DNAm data are fitted to clinical measures of physiological parameters.

The PhenoAge clock was trained to predict an estimate of 'Phenotypic Age', which was developed based upon nine blood clinical markers (albumin, creatinine, serum glucose, C-reactive protein, lymphocyte percent, mean (red) cell volume, red cell distribution width, alkaline phosphatase, white blood cell count) and was significantly associated with mortality in the US National Health and Nutrition Examination Surveys ($n = 9{,}926$ participants aged 18 years and older; 23 years of mortality follow-up)[29]. The epigenetic-based biomarker, PhenoAge, was trained using DNAm and blood chemistry data from the Invecchiare in Chianti (InCHIANTI) study ($n = 912$ participants aged 21–100 years) and validated across multiple cohorts.

The GrimAge clock was developed to directly predict mortality based upon DNAm-based surrogate biomarkers (telomere length, smoking pack-years, adrenomedullin levels, beta-2 microglobulin, cystatin C, GDF-15, leptin, plasminogen activation inhibitor 1 and tissue inhibitor metalloproteinase 1) as well as chronological age and sex in the Framingham Heart Study Offspring and Gen3 cohorts (n = 2,751 participants aged 24–92 years)[30]. The DNAm-based mortality risk estimates are then normalized and linearly transformed into an age estimate (in units of years).

A limitation of the first two generations of epigenetic clocks is that, when residualized for chronological age, epigenetic estimates across technical replicates show only moderate test–retest reliability, which may introduce technical noise into the changes from pre-intervention to post-intervention and make it difficult to distinguish the intervention effects. A new computational method that retrained DNAm clocks using DNAm PCs was developed previously to improve technical reliability for repeated measures[33]. The resulting 'PC clocks' demonstrate exceptional test–retest reliability across technical replicates. Thus, in the current analysis, we applied the PC versions of the first two generations of age clocks.

A third generation of epigenetic clocks, referred to as the Pace of Aging measures, aims to quantify how fast the process of aging-related deterioration of system integrity is proceeding. Pace of Aging estimates the rate of biological aging, defined as the rate of decline in overall system integrity. Pace of Aging values correspond to the years of biological aging experienced during a single calendar year. A value of 1 represents the typical Pace of Aging in a reference population; values above 1 indicate faster Pace of Aging; values below 1 indicate slower Pace of Aging. The newest Pace of Aging measure, DunedinPACE (Pace of Aging Computed from the Epigenome), was developed by predicting within-individual multi-system physiological change across four time points in same-age individuals in the Dunedin Study 1972–1973 birth cohort (n = 817 participants examined at ages 26, 32, 38 and 45 years)[31]. DunedinPACE was developed from analysis of a Pace of Aging composite of slopes of aging-related change in 19 blood chemistry and organ function test metrics of system integrity from four repeated measurements collected over a period of two decades: ApoB100/ApoA1 ratio, BMI, blood urea nitrogen, high-sensitivity C-reactive protein, cardiorespiratory fitness, dental caries experience, total cholesterol, forced expiratory volume in 1 s, forced expiratory volume in 1 s or fixed vital capacity ratio, estimated glomerular filtration rate, hemoglobin A1C, high-density lipoprotein cholesterol, leptin, lipoprotein(a), mean arterial pressure, mean periodontal attachment loss, triglycerides, waist-to-hip ratio and white blood cell count. The DunedinPACE DNAm algorithm was derived from elastic net regression of the physiological Pace of Aging composite on Illumina EPIC array DNAm data derived from blood samples collected at age 45. The set of CpG sites included in the DNAm dataset used to develop the DunedinPACE algorithm was restricted to those showing acceptable test–retest reliability.

Statistics and reproducibility

Sample size calculations for this ancillary study were based upon our original grant for 600 participants, after which additional in-kind support from FOXO Technologies expanded our total sample size to 958 participants. The study flowchart is presented in Fig. 1 that outlines exclusions based on sample availability and data quality. A preliminary power calculation, based on 600 consenting participants with baseline, year 1, and year 2 blood samples, indicated that this sample size would provide 80% power for detecting modest intervention effects as observed in this study.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Due to participant confidentiality and privacy concerns, COSMOS data are available upon request, with formal applications submitted to the COSMOS research committee, to ensure data security and ethical considerations. Please refer to https://cosmostrial.org/for-investigators/ for various forms, information and resources to propose a manuscript/abstract, or to request data using COSMOS resources. Proposals to access and analyze the data will be reviewed by the COSMOS research committee based on whether the hypothesis has been tested, data availability and cleaning requirements, and the programming time required to prepare the data. A response to a request for access is typically provided within two weeks. De-identified data may be subsequently made available to external researchers subject to a controlled-access system and a data-sharing agreement.

References

59. Bassuk, S. S. et al. Baseline characteristics of participants in the VITamin D and OmegA-3 TriaL (VITAL). *Contemp. Clin. Trials* **47**, 235–243 (2016).
60. Zhou, W., Triche, T. J. Jr., Laird, P. W. & Shen, H. SeSAMe: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. *Nucleic Acids Res.* **46**, e123 (2018).
61. Aryee, M. J. et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369 (2014).
62. Tian, Y. et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics* **33**, 3982–3984 (2017).
63. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).
64. McCullough, M. L. et al. Diet quality and major chronic disease risk in men and women: moving toward improved dietary guidance123. *Am. J. Clin. Nutr.* **76**, 1261–1271 (2002).
65. Steen, J., Loeys, T., Moerkerke, B. & Vansteelandt, S. medflex: an R package for flexible mediation analysis using natural effect models. *J. Stat. Softw.* **76**, 1–46 (2017).
66. Chen, B. H. et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging (Albany NY)* **8**, 1844–1865 (2016).
67. Fahy, G. M. et al. Reversal of epigenetic aging and immunosenescent trends in humans. *Aging Cell* **18**, e13028 (2019).
68. Levine, M. E. Assessment of epigenetic clocks as biomarkers of aging in basic and population research. *J. Gerontol. A Biol. Sci. Med. Sci.* **75**, 463–465 (2020).

Acknowledgements

We are deeply indebted to the COSMOS study participants and specifically acknowledge the COSMOS Research Group for their scientific (Brigham and Women's Hospital (BWH), Fred Hutchinson Cancer Research Center (FHCRC), Women's Health Initiative (WHI), Data Safety and Monitoring Board (DSMB), Mars Edge) and logistical (BWH, FHCRC, DSMB, Mars Edge, Contract Pharmacal Corp, Pfizer Consumer Healthcare (now Haleon)) contributions to the overall COSMOS trial. This work was supported by the National Institutes of Health (NIH, grant no. HL157665 to Y.D. and H.D.S.) and investigator-initiated grants from both FOXO Technologies and Massachusetts Life Sciences Center (H.D.S.). Charitable donations made possible by Sutter Health's California Pacific Medical Center Foundation provided salary support for B.H.C. COSMOS is supported by an investigator-initiated grant from Mars Edge, a segment of Mars dedicated to nutrition research and products, which included infrastructure support and the donation of study pills and packaging. Pfizer Consumer Healthcare (now Haleon) provided support through the partial provision of study pills and packaging. COSMOS is also

supported in part by NIH grants nos. AG050657, AG071611 and EY025623. Funding sources were not involved in the trial design or conduct, data collection, data analysis, or manuscript preparation or review.

Author contributions

Y.D. and H.D.S. conceived and/or designed the study and obtained study funding. H.Z., B.H.C., P.M.R., J.E.M., Y.D. and H.D.S. contributed to the acquisition of data. S.L., R.H. and H.D.S. contributed to data analysis. S.L., R.H., H.Z., B.H.C., A.C.P., K.L.I., P.M.R., J.E.M., Y.D. and H.D.S. contributed to data interpretation. S.L. drafted the initial manuscript. All authors provided substantive contributions to the revisions and approved the final version of the submitted manuscript.

Competing interests

J.E.M. and H.D.S. received investigator-initiated grants from Mars Edge, a segment of Mars Incorporated dedicated to nutrition research and products, for infrastructure support and donation of COSMOS study pills and packaging, and Pfizer Consumer Healthcare (now Haleon) for donation of COSMOS study pills and packaging during the conduct of the study. H.D.S. additionally received investigator-initiated grants from FOXO Technologies, Massachusetts Life Sciences Center, Pure Encapsulations, American Pistachio Growers and Haleon, and honoraria and/or travel for lectures from the Council for Responsible Nutrition, BASF, Haleon and NIH during the conduct of the study. P.M.R. has received in-kind support (specifically donations of study pills and packaging) from Mars Edge to be used in an NIH-funded, investigator-initiated trial (U01 AT012611). B.H.C. was formerly an employee of FOXO Technologies, who provided in-kind donations to generate and preprocess the DNA methylation data. S.L. received the EPI Early Career Travel Grant sponsored by the Council on Epidemiology and Prevention's Early Career Committee, American Heart Association. The other authors declare no competing interests.

Additional information

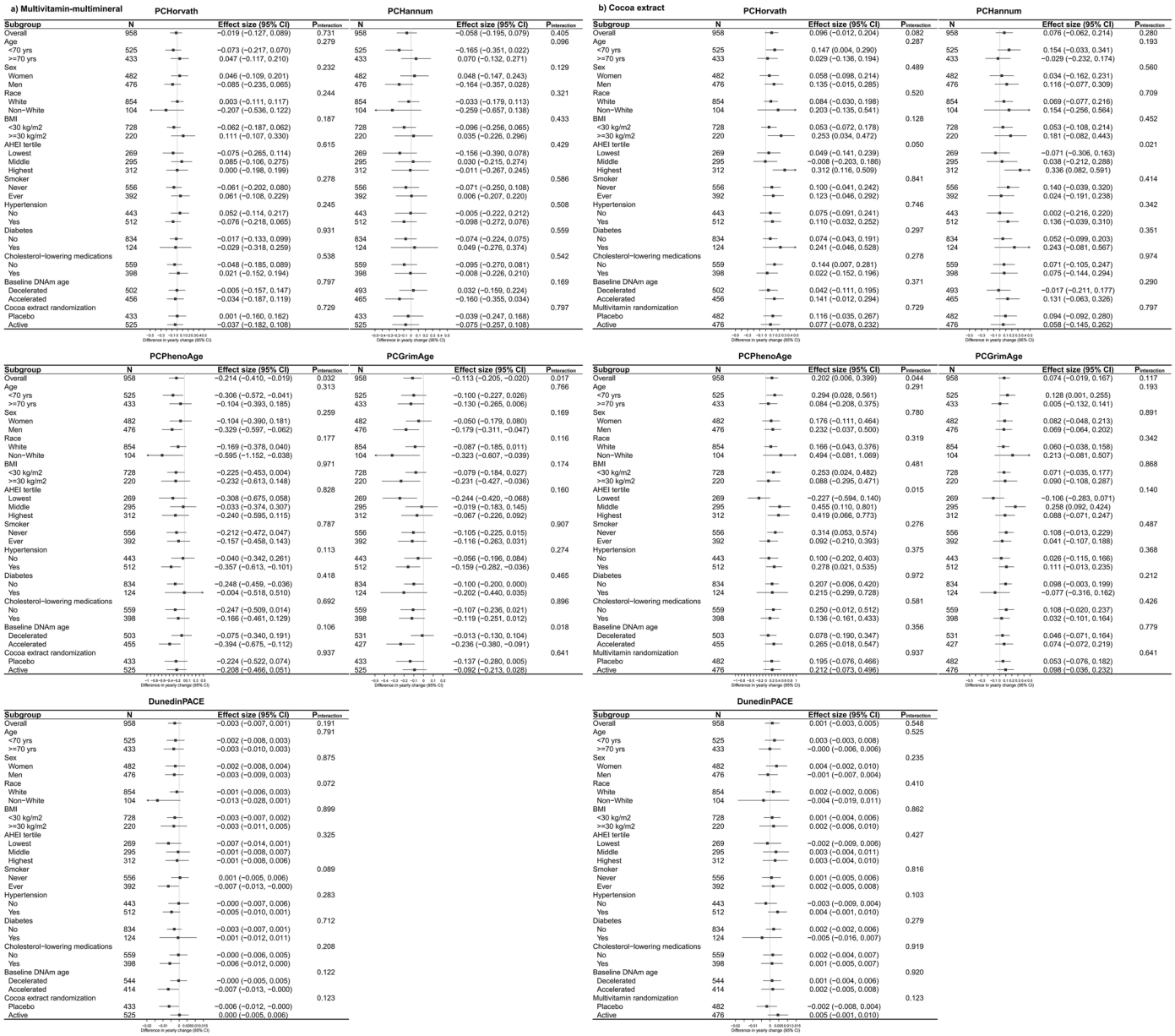
Extended data is available for this paper at https://doi.org/10.1038/s41591-026-04239-3.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41591-026-04239-3.

Correspondence and requests for materials should be addressed to Howard D. Sesso.

Peer review information *Nature Medicine* thanks Daniel Belsky, Albert Higgins-Chen and Jesse Poganik for their contribution to the peer review of this work. Primary Handling Editor: Michael Basson, in collaboration with the *Nature Medicine* team.

https://doi.org/10.1038/s41591-026-04239-3



Extended Data Fig. 1 | Yearly change in epigenetic clocks according to active and placebo Groups by baseline subgroups. a, Between-group differences in yearly change in epigenetic clocks and 95% CI according to multivitamin-multimineral supplementation and placebo (N = 482 for placebo group; N = 476 for active group) by baseline subgroups. **b**, Between-group differences in yearly change in epigenetic clocks and 95% CI according to multivitamin-multimineral supplementation (N = 433 for placebo group; N = 525 for active group) and placebo by baseline subgroups. The sample size for each subgroup is presented in the Figure. Estimates were determined from linear mixed-effects models, using changes in epigenetic clocks as the outcome. The models were adjusted for chronological age at baseline, sex, baseline biological aging for that clock, recruitment source, the other randomized arm, and a random intercept for each individual. Statistical significance for the consistency between subgroups, as determined by two-tailed P values, was assessed through the three-way interactions of intervention, time, and the stratified variables, without adjustment for multiple comparisons. Note, baseline accelerated DNAm age was defined as participants with corresponding age deviation residuals>0 at baseline for the first two generations of epigenetic clocks. For DunedinPACE, accelerated DNAm age was defined as participants with DunedinPACE >1 at baseline.

https://doi.org/10.1038/s41591-026-04239-3

Extended Data Table 1 | Mean values for epigenetic clocks at baseline, year 1, and year 2 stratified by COSMOS interventions

Mean (SD)	Year	Multivitamin		Cocoa extract	
		Placebo (N=482)	Active (N=476)	Placebo (N=433)	Active (N=525)
PCHorvath	0	62.41 (5.52)	62.25 (5.45)	62.41 (5.45)	62.27 (5.51)
	1	63.31 (5.69)	62.96 (5.54)	63.15 (5.56)	63.13 (5.66)
	2	63.9 (5.61)	63.78 (5.67)	63.85 (5.58)	63.83 (5.69)
PCHorvath AgeDev	0	-0.14 (4.63)	-0.17 (4.57)	0.08 (4.58)	-0.35 (4.61)
	1	0.07 (4.73)	-0.15 (4.68)	0.13 (4.64)	-0.18 (4.75)
	2	0.16 (4.71)	0.12 (4.80)	0.28 (4.65)	0.02 (4.83)
PCHannum	0	69.38 (5.61)	69.21 (5.35)	69.33 (5.62)	69.27 (5.37)
	1	70.3 (5.88)	69.82 (5.63)	70.07 (5.82)	70.06 (5.72)
	2	70.97 (5.76)	70.73 (5.65)	70.86 (5.79)	70.85 (5.64)
PCHannum AgeDev	0	-0.06 (4.47)	-0.10 (4.36)	0.13 (4.56)	-0.26 (4.28)
	1	0.09 (4.65)	-0.24 (4.59)	0.11 (4.62)	-0.23 (4.62)
	2	0.21 (4.62)	0.05 (4.57)	0.28 (4.61)	0.02 (4.58)
PCPhenoAge	0	65.55 (6.72)	65.45 (6.48)	65.52 (6.58)	65.48 (6.62)
	1	66.8 (6.93)	66.29 (6.73)	66.49 (6.83)	66.60 (6.84)
	2	67.74 (6.89)	67.27 (6.87)	67.39 (6.90)	67.60 (6.87)
PCPhenoAge AgeDev	0	-0.24 (5.24)	-0.18 (5.35)	0.02 (5.24)	-0.40 (5.32)
	1	0.10 (5.42)	-0.23 (5.60)	0.09 (5.40)	-0.19 (5.60)
	2	0.38 (5.54)	0.01 (5.64)	0.25 (5.47)	0.16 (5.69)
PCGrimAge	0	76.59 (5.00)	76.66 (4.67)	76.40 (4.78)	76.81 (4.88)
	1	77.74 (5.00)	77.63 (4.72)	77.41 (4.90)	77.92 (4.82)
	2	78.54 (4.99)	78.44 (4.78)	78.26 (4.95)	78.68 (4.83)
PCGrimAge AgeDev	0	-0.29 (2.79)	-0.07 (2.91)	-0.20 (2.75)	-0.17 (2.94)
	1	-0.04 (2.80)	0.03 (2.93)	-0.07 (2.80)	0.05 (2.91)
	2	0.12 (2.85)	0.11 (2.90)	0.04 (2.83)	0.17 (2.91)
DunedinPACE	0	0.987 (0.104)	0.987 (0.106)	0.988 (0.106)	0.987 (0.105)
	1	0.991 (0.107)	0.986 (0.108)	0.987 (0.108)	0.990 (0.107)
	2	0.992 (0.107)	0.986 (0.103)	0.988 (0.106)	0.990 (0.104)

AgeDev denotes age deviation, which was estimated by residualizing epigenetic clocks with respect to chronological age.

https://doi.org/10.1038/s41591-026-04239-3

Extended Data Table 2 | Beta coefficients for the associations of selected epigenetic clocks with nutritional biomarkers among a subset of 166 COSMOS participants

Nutrients	β coefficients ± Standard errors (P value)					
	PCPhenoAge deviation residuals		PCGrimAge deviation residuals		DunedinPACE	
	Baseline	Change	Baseline	Change	Baseline	Change
Total carotenoids	-0.13±0.08 (0.092)	-0.023±0.03 (0.429)	**-0.159±0.07 (0.021)**	-0.022±0.03 (0.400)	**-0.341±0.07 (<0.001)**	**-0.074±0.03 (0.013)**
Lycopene	**-0.156±0.08 (0.042)**	-0.011±0.02 (0.645)	**-0.149±0.07 (0.032)**	0.003±0.02 (0.875)	**-0.25±0.07 (<0.001)**	-0.036±0.02 (0.137)
β-carotene	-0.083±0.08 (0.290)	-0.011±0.03 (0.726)	-0.135±0.07 (0.056)	-0.02±0.03 (0.485)	**-0.341±0.07 (<0.001)**	-0.059±0.03 (0.070)
Lutein	-0.113±0.08 (0.134)	-0.041±0.03 (0.171)	-0.111±0.07 (0.103)	**-0.056±0.03 (0.035)**	**-0.268±0.07 (<0.001)**	-0.057±0.03 (0.067)
Zeaxanthin	-0.096±0.08 (0.205)	-0.006±0.02 (0.775)	-0.059±0.07 (0.394)	-0.003±0.02 (0.870)	**-0.279±0.07 (<0.001)**	-0.026±0.02 (0.249)
β-cryptoxanthin	-0.046±0.08 (0.544)	0.01±0.03 (0.772)	-0.066±0.07 (0.334)	-0.03±0.03 (0.316)	**-0.182±0.07 (0.009)**	-0.059±0.03 (0.089)
α-carotene	-0.087±0.08 (0.260)	0.028±0.03 (0.376)	**-0.187±0.07 (0.007)**	-0.014±0.03 (0.613)	**-0.271±0.07 (<0.001)**	-0.01±0.03 (0.766)
Vitamin E	-0.046±0.08 (0.560)	-0.013±0.02 (0.494)	-0.035±0.07 (0.625)	-0.011±0.02 (0.521)	-0.139±0.07 (0.054)	-0.012±0.02 (0.566)
Folate	-0.005±0.08 (0.950)	-0.03±0.02 (0.153)	-0.106±0.07 (0.127)	-0.037±0.02 (0.054)	**-0.152±0.07 (0.032)**	-0.003±0.02 (0.877)
Vitamin B12	0.144±0.08 (0.067)	**-0.057±0.03 (0.031)**	0.017±0.07 (0.814)	-0.038±0.02 (0.110)	-0.040±0.07 (0.577)	0.002±0.03 (0.945)
Vitamin D	-0.02±0.08 (0.796)	**-0.068±0.02 (0.007)**	**-0.139±0.07 (0.044**）	-0.039±0.02 (0.082)	-0.078±0.07 (0.274)	-0.026±0.03 (0.328)

Beta coefficients were estimated using linear mixed-effect models using standardized epigenetic measures as outcome variables. Each nutritional biomarker (that is nutrient) was log-transformed and standardized as exposure variables. Models included nutrients at baseline (beta coefficients denote the association between nutrients and clocks at baseline, and associated two-tailed P values estimated in the linear mixed-effect models were used to determine statistical significance), baseline nutrient×time, changes in nutrients from baseline×time (beta coefficients denote the association between longitudinal change in epigenetic measures with changes in nutrients from baseline, and associated two-tailed P values estimated in the linear mixed-effect models were used to determine statistical significance), chronological age, sex, and an individual-level random intercept. Statistical significance with two-tailed P values was not adjusted for multiple comparisons. Values are beta coefficients (P values), and estimates with P value < 0.05 are marked in bold type.

https://doi.org/10.1038/s41591-026-04239-3

Extended Data Table 3 | Effects of daily multivitamin-multimineral use on longitudinal changes in nutritional biomarkers stratified by baseline epigenetic age acceleration status among a subset of 166 COSMOS participants

Nutrient	Difference in yearly change in nutritional biomarkers between MVM and placebo groups (95% CI); P for interaction								
	PCPhenoAge			PCGrimAge			DunedinPACE		
	Decelerated (n=89)	Accelerated (n=77)	P	Decelerated (n=99)	Accelerated (n=67)	P	Decelerated (n=96)	Accelerated (n=70)	P
Total carotenoids	0.15 (0.01, 0.29)	0.15 (-0.02, 0.32)	0.99	0.15 (0.01, 0.3)	0.13 (-0.03, 0.29)	0.85	0.09 (-0.03, 0.2)	0.22 (0.02, 0.42)	0.23
Lycopene	0.10 (-0.09, 0.29)	0.10 (-0.12, 0.31)	0.97	0.14 (-0.05, 0.32)	0.04 (-0.18, 0.26)	0.51	0.04 (-0.13, 0.2)	0.19 (-0.06, 0.43)	0.31
β-carotene	0.25 (0.13, 0.37)	0.20 (0.05, 0.35)	0.59	0.19 (0.06, 0.32)	0.27 (0.14, 0.41)	0.42	0.19 (0.09, 0.29)	0.27 (0.09, 0.45)	0.44
Lutein	0.07 (-0.08, 0.23)	0.21 (0.05, 0.37)	0.22	0.07 (-0.09, 0.22)	0.24 (0.09, 0.39)	0.14	**0.02 (-0.12, 0.15)**	**0.29 (0.11, 0.46)**	**0.02**
Zeaxanthin	0.01 (-0.13, 0.14)	0.13 (-0.11, 0.36)	0.39	0.05 (-0.14, 0.24)	0.10 (-0.05, 0.26)	0.66	0.03 (-0.13, 0.2)	0.11 (-0.1, 0.32)	0.54
β-cryptoxanthin	-0.02 (-0.15, 0.11)	-0.09 (-0.25, 0.07)	0.53	-0.01 (-0.15, 0.12)	-0.12 (-0.26, 0.02)	0.33	-0.11 (-0.24, 0.02)	0.01 (-0.15, 0.17)	0.24
α-carotene	0.03 (-0.12, 0.17)	-0.10 (-0.24, 0.04)	0.22	0.02 (-0.11, 0.15)	-0.12 (-0.28, 0.04)	0.16	-0.03 (-0.16, 0.09)	-0.04 (-0.2, 0.12)	0.92
Vitamin E	0.51 (0.31, 0.71)	0.47 (0.24, 0.69)	0.77	0.56 (0.35, 0.77)	0.40 (0.18, 0.61)	0.30	0.48 (0.3, 0.66)	0.51 (0.25, 0.76)	0.88
Folate	0.45 (0.26, 0.64)	0.59 (0.34, 0.83)	0.37	**0.38 (0.18, 0.58)**	**0.73 (0.5, 0.96)**	**0.03**	0.54 (0.34, 0.74)	0.48 (0.25, 0.71)	0.72
Vitamin B12	0.41 (0.26, 0.56)	0.54 (0.37, 0.7)	0.27	0.48 (0.32, 0.64)	0.44 (0.29, 0.59)	0.71	0.43 (0.29, 0.56)	0.52 (0.34, 0.7)	0.41
Vitamin D	0.38 (0.22, 0.54)	0.54 (0.37, 0.71)	0.19	0.38 (0.24, 0.51)	0.57 (0.36, 0.78)	0.1	0.52 (0.38, 0.66)	0.37 (0.17, 0.57)	0.22

Accelerated biological aging was defined as participants with corresponding age deviation residuals>0 at baseline for PCPhenoAge and PCGrimAge or participants with DunedinPACE >1 at baseline for DunedinPACE. Estimates were from the linear mixed-effect models adjusted for chronological age at baseline, sex, recruitment source, randomized cocoa extract assignment, and a random intercept of each individual. Statistical significance for the consistency between subgroups, as determined by two-tailed P values, was assessed through the three-way interactions of intervention, time, and the stratified variables, without adjustment for multiple comparisons.

https://doi.org/10.1038/s41591-026-04239-3

Extended Data Table 4 | Yearly change in epigenetic clocks according to Cocoa Extract and Multivitamin-multimineral Interventions

	Baseline	Year 1	Year 2	Between-group differences in yearly change	Cohen's d	P value
PCHorvath						
Placebo+Placebo (N=219)	62.51 (5.58)	63.30 (5.73)	63.94 (5.60)	Ref	Ref	
Placebo+MVM (N=214)	62.30 (5.32)	63.00 (5.40)	63.76 (5.56)	0.001 (-0.159, 0.162)	0 (-0.03, 0.03)	0.985
CF+Placebo (N=263)	62.32 (5.47)	63.32 (5.67)	63.86 (5.63)	0.115 (-0.037, 0.268)	0.021 (-0.008, 0.05)	0.139
CF+MVM (N=262)	62.22 (5.56)	62.93 (5.66)	63.80 (5.77)	0.078 (-0.074, 0.231)	0.015 (-0.014, 0.043)	0.315
PCHorvath AgeDev						
Placebo+Placebo (N=219)	0.05 (4.80)	0.13 (4.89)	0.25 (4.76)	Ref	Ref	
Placebo+MVM (N=214)	0.11 (4.36)	0.13 (4.39)	0.31 (4.55)	0.001 (-0.159, 0.161)	0 (-0.035, 0.035)	0.992
CF+Placebo (N=263)	-0.30 (4.49)	0.02 (4.60)	0.09 (4.67)	0.112 (-0.04, 0.265)	0.024 (-0.009, 0.058)	0.149
CF+MVM (N=262)	-0.41 (4.74)	-0.38 (4.90)	-0.04 (5.00)	0.078 (-0.075, 0.231)	0.017 (-0.016, 0.05)	0.318
PCHannum						
Placebo+Placebo (N=219)	69.47 (5.68)	70.30 (5.90)	71.03 (5.82)	Ref	Ref	
Placebo+MVM (N=214)	69.19 (5.56)	69.84 (5.74)	70.68 (5.77)	-0.039 (-0.242, 0.164)	-0.008 (-0.049, 0.033)	0.707
CF+Placebo (N=263)	69.31 (5.56)	70.30 (5.88)	70.92 (5.72)	0.094 (-0.1, 0.288)	0.018 (-0.021, 0.057)	0.342
CF+MVM (N=262)	69.22 (5.19)	69.81 (5.56)	70.77 (5.57)	0.019 (-0.175, 0.213)	0.004 (-0.035, 0.042)	0.849
PCHannum AgeDev						
Placebo+Placebo (N=219)	0.12 (4.66)	0.18 (4.75)	0.32 (4.68)	Ref	Ref	
Placebo+MVM (N=214)	0.14 (4.47)	0.04 (4.51)	0.23 (4.55)	-0.040 (-0.243, 0.164)	-0.009 (-0.055, 0.037)	0.702
CF+Placebo (N=263)	-0.21 (4.30)	0.02 (4.58)	0.13 (4.58)	0.091 (-0.103, 0.285)	0.021 (-0.023, 0.065)	0.358
CF+MVM (N=262)	-0.30 (4.26)	-0.48 (4.65)	-0.10 (4.59)	0.018 (-0.176, 0.213)	0.004 (-0.04, 0.048)	0.852
PCPhenoAge						
Placebo+Placebo (N=219)	65.73 (6.61)	66.83 (6.90)	67.84 (7.01)	Ref	Ref	
Placebo+MVM (N=214)	65.30 (6.55)	66.15 (6.77)	66.94 (6.77)	-0.224 (-0.514, 0.066)	-0.041 (-0.093, 0.012)	0.131
CF+Placebo (N=263)	65.39 (6.82)	66.77 (6.97)	67.65 (6.79)	0.196 (-0.081, 0.472)	0.034 (-0.016, 0.084)	0.166
CF+MVM (N=262)	65.57 (6.42)	66.41 (6.70)	67.55 (6.96)	-0.013 (-0.29, 0.264)	-0.002 (-0.052, 0.047)	0.928
PCPhenoAge AgeDev						
Placebo+Placebo (N=219)	0.05 (5.31)	0.23 (5.44)	0.54 (5.64)	Ref	Ref	
Placebo+MVM (N=214)	-0.01 (5.19)	-0.05 (5.37)	-0.05 (5.30)	-0.225 (-0.515, 0.065)	-0.043 (-0.097, 0.012)	0.129
CF+Placebo (N=263)	-0.48 (5.18)	-0.01 (5.41)	0.25 (5.47)	0.192 (-0.085, 0.468)	0.036 (-0.016, 0.089)	0.174
CF+MVM (N=262)	-0.32 (5.48)	-0.38 (5.80)	0.06 (5.91)	-0.013 (-0.29, 0.264)	-0.003 (-0.055, 0.05)	0.925
PCGrimAge						
Placebo+Placebo (N=219)	76.48 (4.89)	77.57 (5.00)	78.47 (5.04)	Ref	Ref	
Placebo+MVM (N=214)	76.32 (4.67)	77.26 (4.79)	78.04 (4.86)	**-0.137 (-0.275, 0)**	**-0.042 (-0.083, 0)**	**0.050**
CF+Placebo (N=263)	76.69 (5.10)	77.88 (4.99)	78.60 (4.95)	0.053 (-0.078, 0.184)	0.014 (-0.025, 0.054)	0.427
CF+MVM (N=262)	76.93 (4.65)	77.95 (4.64)	78.77 (4.70)	-0.04 (-0.171, 0.091)	-0.012 (-0.052, 0.027)	0.548
PCGrimAge AgeDev						
Placebo+Placebo (N=219)	-0.29 (2.78)	-0.11 (2.85)	0.10 (2.87)	Ref	Ref	
Placebo+MVM (N=214)	-0.10 (2.71)	-0.04 (2.75)	-0.02 (2.79)	**-0.138 (-0.275, -0.001)**	**-0.048 (-0.097, 0)**	**0.048**
CF+Placebo (N=263)	-0.28 (2.80)	0.02 (2.75)	0.13 (2.84)	0.049 (-0.081, 0.18)	0.017 (-0.028, 0.063)	0.458
CF+MVM (N=262)	-0.05 (3.08)	0.08 (3.07)	0.22 (2.98)	-0.041 (-0.172, 0.09)	-0.014 (-0.06, 0.032)	0.543
DunedinPACE						
Placebo+Placebo (N=219)	0.993 (0.103)	0.995 (0.107)	1.001 (0.107)	Ref	Ref	
Placebo+MVM (N=214)	0.982 (0.109)	0.979 (0.109)	0.975 (0.103)	**-0.006 (-0.012, 0)**	**-0.059 (-0.117, -0.002)**	**0.043**
CF+Placebo (N=263)	0.981 (0.105)	0.987 (0.107)	0.985 (0.106)	-0.002 (-0.008, 0.004)	-0.018 (-0.073, 0.037)	0.514
CF+MVM (N=262)	0.992 (0.104)	0.992 (0.107)	0.995 (0.103)	-0.002 (-0.007, 0.004)	-0.016 (-0.071, 0.039)	0.56

Estimates were from the linear mixed-effect models using change in epigenetic measures of biological aging (original clocks or AgeDev measures) as the outcome. Models were adjusted for chronological age at baseline, sex, baseline biological aging measures from that clock, recruitment source, and a random intercept of each individual. For original clock values, Cohen's *d* was estimated by scaling the repeated measures of age-difference values (that is epigenetic clock age - chronological age) using the standard deviation of age-difference values at baseline for the first two generations of epigenetic clocks; DunedinPACE was directly scaled using the standard deviation at baseline. For AgeDev measures, Cohen's *d* was estimated by standardizing repeated measures of epigenetic age deviation residuals using the mean and standard deviation at baseline. Statistical significance with two-tailed P values was determined through the interactions of intervention and time in the linear mixed-effect models, without adjustment for multiple comparisons. AgeDev, age deviation; MVM, multivitamin-multimineral; CF, cocoa flavanols.

https://doi.org/10.1038/s41591-026-04239-3

Extended Data Table 5 | Differences in yearly changes in epigenetic clocks between intervention and placebo groups in sensitivity analysis by using age acceleration measures of the first two generations of age clocks as outcome

Epigenetic age accelerations	Yearly change from baseline		Difference (Multivitamin versus Placebo)	Cohen's *d*	P value
	Placebo	Active			
Method 1: using a single linear model for all three time points					
Multivitamin					
PCHorvath AgeDev	0.142 (0.066,0.218)	0.124 (0.031,0.217)	-0.018 (-0.126,0.09)	-0.004 (-0.027,0.02)	0.745
PCHannum AgeDev	0.120 (0.024,0.217)	0.063 (-0.055,0.182)	-0.057 (-0.194,0.08)	-0.013 (-0.044,0.018)	0.415
PCPhenoAge AgeDev	0.300 (0.162,0.437)	0.087 (-0.082,0.256)	-0.213 (-0.409,-0.017)	-0.040 (-0.077,-0.003)	0.033
PCGrimAge AgeDev	0.216 (0.151,0.281)	0.105 (0.025,0.185)	-0.111 (-0.204,-0.019)	-0.039 (-0.072,-0.007)	0.018
Cocoa extract					
PCHorvath AgeDev	0.081 (0.001,0.161)	0.176 (0.081,0.271)	0.095 (-0.013,0.203)	0.021 (-0.003,0.044)	0.086
PCHannum AgeDev	0.051 (-0.05,0.153)	0.126 (0.005,0.247)	0.075 (-0.063,0.212)	0.017 (-0.014,0.048)	0.288
PCPhenoAge AgeDev	0.085 (-0.061,0.23)	0.285 (0.113,0.458)	0.201 (0.004,0.397)	0.038 (0.001,0.075)	0.045
PCGrimAge AgeDev	0.121 (0.053,0.19)	0.194 (0.113,0.276)	0.073 (-0.02,0.166)	0.026 (-0.007,0.058)	0.124
Method 2: using separate linear models for three time points					
Multivitamin					
PCHorvath AgeDev	-0.006 (-0.082,0.07)	-0.024 (-0.117,0.07)	-0.017 (-0.125,0.09)	-0.004 (-0.027,0.02)	0.752
PCHannum AgeDev	0.012 (-0.084,0.109)	-0.043 (-0.161,0.075)	-0.055 (-0.192,0.082)	-0.013 (-0.044,0.019)	0.429
PCPhenoAge AgeDev	0.097 (-0.041,0.234)	-0.116 (-0.285,0.053)	-0.212 (-0.408,-0.017)	-0.04 (-0.077,-0.003)	0.034
PCGrimAge AgeDev	0.069 (0.004,0.134)	-0.042 (-0.122,0.038)	-0.111 (-0.204,-0.019)	-0.039 (-0.071,-0.007)	0.018
Cocoa extract					
PCHorvath AgeDev	-0.066 (-0.146,0.014)	0.028 (-0.068,0.123)	0.093 (-0.015,0.202)	0.02 (-0.003,0.044)	0.091
PCHannum AgeDev	-0.054 (-0.155,0.048)	0.017 (-0.104,0.137)	0.07 (-0.067,0.208)	0.016 (-0.015,0.047)	0.317
PCPhenoAge AgeDev	-0.117 (-0.263,0.028)	0.082 (-0.091,0.254)	0.199 (0.003,0.395)	0.038 (0,0.075)	0.047
PCGrimAge AgeDev	-0.025 (-0.094,0.043)	0.047 (-0.035,0.128)	0.072 (-0.021,0.165)	0.025 (-0.007,0.058)	0.128
Method 3: using linear mixed models adjusting for individual-level random intercepts					
Multivitamin					
PCHorvath AgeDev	0.076 (0.014,0.138)	0.058 (-0.018,0.134)	-0.018 (-0.106,0.070)	-0.018 (-0.106,0.069)	0.682
PCHannum AgeDev	0.092 (0.013,0.171)	0.031 (-0.066,0.128)	-0.062 (-0.174,0.051)	-0.049 (-0.137,0.04)	0.282
PCPhenoAge AgeDev	0.249 (0.136,0.362)	0.033 (-0.106,0.172)	-0.216 (-0.377,-0.055)	-0.122 (-0.212,-0.031)	0.009
PCGrimAge AgeDev	0.164 (0.111,0.217)	0.051 (-0.015,0.116)	-0.113 (-0.189,-0.038)	-0.134 (-0.223,-0.044)	0.003
Cocoa extract					
PCHorvath AgeDev	0.015 (-0.051,0.08)	0.111 (0.033,0.189)	0.096 (0.008,0.185)	0.096 (0.008,0.184)	0.033
PCHannum AgeDev	0.019 (-0.064,0.103)	0.097 (-0.002,0.196)	0.078 (-0.035,0.191)	0.061 (-0.028,0.151)	0.176
PCPhenoAge AgeDev	0.031 (-0.088,0.151)	0.234 (0.091,0.376)	0.202 (0.041,0.364)	0.114 (0.023,0.205)	0.014
PCGrimAge AgeDev	0.067 (0.011,0.124)	0.141 (0.074,0.208)	0.074 (-0.002,0.15)	0.087 (-0.003,0.177)	0.057

Considering there is no consistent way to estimate age deviation (AgeDev) measures in longitudinal studies, we used three different methods to calculate the residuals of regressing epigenetic clocks for chronological age: 1) Directly regressing epigenetic clocks on chronological age for all samples: this method is the most commonly used but it did not take the clustering effect within each participant into consideration; 2) Estimating residuals separately for different time points: this methods may ignore the natural changes in age acceleration residuals within each participant; 3) Using linear mixed-effect models: this methods addressed the clustering effect within each participant but also removed the individual-level biological variations in age acceleration residuals. Estimates were from the mixed linear model using change in age deviation residuals as the outcome. Models were adjusted for chronological age at baseline, sex, baseline age deviation residuals from that clock, recruitment source, the other intervention arm, and a random intercept of each individual. Cohen's *d* was estimated by standardizing repeated measures of epigenetic age deviation residuals using the mean and standard deviation at baseline. Statistical significance with two-tailed P values was determined through the interactions of intervention and time in the linear mixed-effect models, without adjustment for multiple comparisons. AgeDev, age deviation.

https://doi.org/10.1038/s41591-026-04239-3

Extended Data Table 6 | Differences in yearly changes in epigenetic clocks between intervention and placebo groups in per-protocol analysis

	Yearly change from baseline		Difference (Active versus placebo)	Cohen's *d*	P value
	Placebo	Active			
Multivitamin					
PCHorvath	0.769 (0.691,0.847)	0.753 (0.657,0.848)	-0.016 (-0.127,0.094)	-0.002 (-0.023,0.018)	0.773
PCHannum	0.826 (0.727,0.925)	0.753 (0.631,0.875)	-0.073 (-0.214,0.067)	-0.014 (-0.042,0.014)	0.307
PCPhenoAge	1.14 (0.999,1.281)	0.912 (0.739,1.085)	-0.228 (-0.428,-0.028)	-0.04 (-0.076,-0.004)	0.026
PCGrimAge	1.032 (0.965,1.098)	0.911 (0.829,0.993)	-0.120 (-0.215,-0.025)	-0.035 (-0.064,-0.007)	0.013
DunedinPACE	0.004 (0.001,0.007)	0.001 (-0.003,0.004)	-0.003 (-0.007,0.001)	-0.027 (-0.067,0.012)	0.176
Cocoa extract					
PCHorvath	0.698 (0.617,0.779)	0.796 (0.699,0.892)	0.097 (-0.012,0.207)	0.018 (-0.003,0.039)	0.081
PCHannum	0.726 (0.623,0.829)	0.814 (0.691,0.936)	0.087 (-0.052,0.227)	0.017 (-0.011,0.045)	0.221
PCPhenoAge	0.878 (0.733,1.024)	1.113 (0.94,1.287)	0.235 (0.037,0.432)	0.042 (0.006,0.077)	0.020
PCGrimAge	0.917 (0.848,0.987)	1.002 (0.919,1.085)	0.085 (-0.009,0.179)	0.025 (-0.003,0.053)	0.078
DunedinPACE	0.001 (-0.002,0.004)	0.003 (-0.001,0.006)	0.002 (-0.003,0.006)	0.015 (-0.024,0.054)	0.447

Estimates were from linear mixed-effect models using the change in epigenetic clocks as the outcome. Models were adjusted for chronological age at baseline, sex, baseline biological aging for that clock, recruitment source, the other randomization arm, and a random intercept of each individual. Cohen's *d* was estimated by scaling the repeated measures of age-difference values (that is epigenetic clock age - chronological age) using the standard deviation of age-difference values at baseline for the first two generations of epigenetic clocks; DunedinPACE was directly scaled using the standard deviation at baseline. Statistical significance with two-tailed P values was determined through the interactions of intervention and time in the linear mixed-effect models, without adjustment for multiple comparisons.

https://doi.org/10.1038/s41591-026-04239-3

Extended Data Table 7 | Differences in yearly changes in different components of PCGrimAge between multivitamin and placebo

DNAm measures	Yearly change from baseline		Difference (Active versus placebo)	P value
	Placebo	Multivitamin		
Multivitamin				
PCDNAmTL	-0.115 (-0.128,-0.101)	-0.093 (-0.109,-0.076)	0.022 (0.003,0.042)	0.027
PCPACKYRS	0.071 (0.054,0.087)	0.053 (0.032,0.073)	-0.018 (-0.042,0.006)	0.139
PCADM	0.125 (0.114,0.137)	0.114 (0.1,0.129)	-0.011 (-0.028,0.005)	0.189
PCB2M	0.202 (0.19,0.214)	0.174 (0.159,0.189)	-0.028 (-0.045,-0.01)	0.002
PCCystatinC	0.211 (0.2,0.223)	0.184 (0.17,0.199)	-0.027 (-0.044,-0.011)	0.001
PCPAI1	0.066 (0.047,0.085)	0.06 (0.036,0.083)	-0.007 (-0.034,0.02)	0.628
PCGDF15	0.216 (0.204,0.228)	0.187 (0.172,0.202)	-0.029 (-0.046,-0.012)	0.001
PCLeptin	0.01 (0.005,0.015)	0.016 (0.01,0.022)	0.006 (-0.001,0.013)	0.098
PCTIMP1	0.218 (0.206,0.229)	0.202 (0.187,0.216)	-0.016 (-0.033,0)	0.057
Cocoa extract				
PCDNAmTL	-0.102 (-0.116,-0.087)	-0.105 (-0.123,-0.088)	-0.004 (-0.023,0.016)	0.710
PCPACKYRS	0.057 (0.04,0.075)	0.065 (0.044,0.086)	0.008 (-0.016,0.032)	0.507
PCADM	0.11 (0.098,0.122)	0.128 (0.113,0.143)	0.018 (0.002,0.035)	0.033
PCB2M	0.18 (0.167,0.192)	0.195 (0.18,0.211)	0.016 (-0.001,0.033)	0.074
PCCystatinC	0.191 (0.179,0.204)	0.203 (0.189,0.218)	0.012 (-0.005,0.029)	0.152
PCPAI1	0.061 (0.041,0.081)	0.065 (0.041,0.089)	0.004 (-0.023,0.031)	0.770
PCGDF15	0.194 (0.181,0.206)	0.208 (0.193,0.223)	0.014 (-0.003,0.031)	0.102
PCLeptin	0.013 (0.008,0.019)	0.013 (0.007,0.019)	0 (-0.007,0.007)	0.997
PCTIMP1	0.201 (0.189,0.214)	0.217 (0.202,0.231)	0.015 (-0.001,0.032)	0.073

PCGrimAge included 9 DNAm based surrogates for telomere length (DNAmTL), smoking pack-years, adrenomedullin levels (ADM), beta-2 microglobulin (B2M), cystatin C, growth differentiation factor 15 (GDF-15), leptin, plasminogen activation inhibitor 1 (PAI-1), tissue inhibitor metalloproteinase 1 (TIMP-1). As the DNAm-based biomarkers have different scales, we standardized measurements using mean and standard deviations of corresponding measures at baseline. Estimates were from mixed linear models using change in DNAm measures as the outcome. Models were adjusted for chronological age at baseline, sex, baseline measurement of corresponding DNAm-based biomarkers, recruitment source, the other randomization arm, and a random intercept of each individual. Statistical significance with two-tailed P values was determined through the interactions of intervention and time in the linear mixed-effect models, without adjustment for multiple comparisons.

https://doi.org/10.1038/s41591-026-04239-3

Extended Data Table 8 | Beta coefficients for the associations of selected epigenetic clocks with inflammaging biomarkers and cognitive function among subsets of participants that were part of other COSMOS ancillary studies

	β coefficients ± Standard errors (P value)					
	PCGrimAge AgeDev		PCPhenoAge AgeDev		DunedinPACE	
	Baseline	Change	Baseline	Change	Baseline	Change
Inflammmaging ancillary study (N=591)						
hsCRP	**0.175±0.046 (<0.001)**	**0.085±0.028 (0.002)**	**0.12±0.039 (0.002)**	**0.119±0.024 (<0.001)**	**0.272±0.044 (<0.001)**	**0.132±0.025 (<0.001)**
IL6	**0.254±0.043 (<0.001)**	**0.17±0.033 (<0.001)**	**0.164±0.038 (<0.001)**	**0.126±0.029 (<0.001)**	**0.264±0.041 (<0.001)**	**0.166±0.028 (<0.001)**
TNFa	**0.101±0.048 (0.034)**	0.037±0.03 (0.215)	**0.175±0.041 (<0.001)**	0.02±0.026 (0.431)	**0.147±0.046 (0.002)**	0.02±0.026 (0.426)
IL10	**0.118±0.045 (0.010)**	**0.157±0.032 (<0.001)**	0.076±0.04 (0.056)	**0.138±0.028 (<0.001)**	**0.088±0.044 (0.047)**	**0.088±0.028 (0.002)**
IFNg	0.033±0.047 (0.482)	0.04±0.036 (0.264)	**0.11±0.041 (0.007)**	**0.067±0.031 (0.032)**	**0.101±0.046 (0.027)**	0.024±0.031 (0.437)
COSMOS-Mind (N=127)						
Global cognition	-0.107±0.079 (0.179)	0.065±0.047 (0.171)	**-0.165±0.071 (0.022)**	0.043±0.04 (0.279)	-0.016±0.079 (0.837)	0.071±0.036 (0.051)
Episodic memory	-0.099±0.091 (0.277)	0.09±0.065 (0.170)	-0.085±0.083 (0.306)	0.011±0.055 (0.845)	0.001±0.09 (0.990)	0.048±0.05 (0.340)
Executive function	-0.068±0.081 (0.404)	0.039±0.05 (0.436)	**-0.189±0.073 (0.011)**	0.061±0.042 (0.148)	-0.048±0.08 (0.552)	0.049±0.039 (0.207)

COSMOS inflammaging ancillary study measured five inflammaging biomarkers (including high-sensitivity C-reactive proteins [hsCRP], interleukin 6 [IL-6], tumor necrosis factor alpha [TNF-α], IL-10, and Interferon gamma [IFNγ]) among a randomly selected sample of 598 participants at baseline, year 1, and year 2. COSMOS-Mind assessed cognition by telephone at baseline and annually during 3-years of follow-up among 2262 participants. The primary outcome of COSMOS-Mind was a global cognition composite (reported as a z-score based on baseline data from the study cohort), with higher scores reflecting better performance. Key secondary outcomes included an episodic memory composite and an executive function composite. Epigenetic measures were standardized as exposure variables. The association between baseline and changes in epigenetic measures from baseline with longitudinal changes in inflammaging biomarkers (log-transformed and then standardized) or cognitive function measures (standardized z scores) within 2 years of follow-up using linear mixed-effect models. Models included epigenetic deviation residuals at baseline (beta coefficients denote the association between epigenetic deviation residuals and outcomes at baseline, and associated two-tailed P values estimated in the linear mixed-effect models were used to determine statistical significance), baseline epigenetic deviation residuals×time, changes in epigenetic deviation residuals from baseline×time (beta coefficients denote the association between change in epigenetic measures from baseline with longitudinal changes in outcomes, and associated two-tailed P values estimated in the linear mixed-effect models were used to determine statistical significance), age, sex, recruitment source, smoking status, diet quality, hypertension, diabetes, body mass index, and an individual-level random intercept. Statistical significance with two-tailed P values was not adjusted for multiple comparisons. Values are beta coefficients (P values) and estimates with P value < 0.05 are marked in bold type.

https://doi.org/10.1038/s41591-026-04239-3

Extended Data Table 9 | Effects of a MVM on inflammaging biomarkers and cognitive function mediated through selected epigenetic clocks among subsets of participants that were part of other COSMOS ancillary studies

Outcome	Total effect		Natural indirect effect								
	β±SD	P	PCGrimAge AgeDev			PCPhenoAge AgeDev			DunedinPACE		
			β±SD	% Med	P	β±SD	% Med	P	β±SD	% Med	P
Inflammmaging ancillary study (N=573)											
hsCRP	0.05±0.064	0.43	-0.01±0.007	-19.4	0.19	-0.02±0.011	-40.8	0.059	-0.015±0.013	-30.6	0.23
IL6	-0.079±0.076	0.30	-0.02±0.013	25.1	0.12	-0.022±0.013	28.0	0.084	-0.014±0.012	18.1	0.24
TNFa	0.012±0.065	0.85	-0.001±0.006	-9.9	0.84	-0.001±0.007	-6.0	0.92	0±0.004	0.5	0.99
IL10	-0.035±0.07	0.62	-0.02±0.012	57.9	0.093	-0.025±0.013	70.0	0.064	-0.006±0.007	16.8	0.38
IFNg	-0.081±0.077	0.29	-0.008±0.008	9.5	0.32	-0.014±0.01	17.8	0.15	0.002±0.005	-2.5	0.71
COSMOS-Mind (N=117)											
Global cognition	-0.016±0.123	0.90	-0.004±0.011	26.6	0.70	-0.002±0.008	10.2	0.83	-0.006±0.02	39.1	0.76
Episodic memory	0.087±0.16	0.59	0.004±0.012	4.1	0.76	0.005±0.011	5.2	0.67	-0.004±0.021	-4.5	0.85
Executive function	-0.025±0.122	0.84	-0.01±0.014	39.8	0.48	-0.003±0.009	11.2	0.75	-0.003±0.009	12.1	0.73

COSMOS inflammaging ancillary study measured five inflammaging biomarkers among a random selected sample of 600 participants at baseline, year 1, and year 2. The effects of MVM (exposure) on 2-year changes in inflammaging biomarkers (outcomes) that were mediated through 2-year changes in epigenetic measures (mediators) were examined based on 573 participants with complete data on inflammaging biomarkers and epigenetic measures at both baseline and year 2. COSMOS-Mind assessed cognition by telephone among 2262 participants at baseline and annually during 3 years of follow-up. The primary outcome of COSMOS-Mind was a global cognition composite (reported as a z-score based on baseline data from the study cohort), with higher scores reflecting better performance. Key secondary outcomes included an episodic memory composite and an executive function composite. The effects of MVM (intervention) on 3-year changes in cognitive function (outcomes) that were mediated through 2-year changes in epigenetic measures (mediators) were examined based on 117 participants with complete data on cognitive function and epigenetic measures at both baseline and follow-up. Total effect (TE) indicates the overall effect of exposure on outcome, and the corresponding p-value denotes the significance of the effect of MVM on the outcome. The natural indirect effect (NIE) represents the estimated effect of MVM on 2-year changes in outcomes when controlling for MVM while changing epigenetic measures from the level it would have been in the placebo group. The proportion of the effects of MVM on outcomes that is mediated through epigenetic measures (% Med) was estimated by dividing the beta coefficients for the NIE by the beta coefficient for the TE, and the corresponding two-tailed p-value denotes the significance of the mediation effect through epigenetic measures, without adjustment for multiple comparisons. Mediation analyses were performed based on a counterfactual framework using methods developed by VanderWeele et al. and by the imputation-based approach using the R package medflex. Models were adjusted for age, sex, recruitment source, baseline cognitive function, smoking status, diet quality, hypertension, diabetes, body mass index, and randomized cocoa extract assignment.

nature portfolio

Corresponding author(s): Howard D. Sesso

Last updated by author(s): Jan 5, 2026

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Data collection was conducted using REDCap (Research Electronic Data Capture).
Data analysis	Data analysis was performed with R (version 4.4.2). Quality control and normalization analyses were performed using the R Bioconductor package SeSAMe (R version 4.3.0; Bioconductor version 3.17; SeSAMe version 1.18.4). Batch effects were additionally identified using singular value decomposition analysis and then corrected using Combat in sva R package. PC versions of the Horvath, and Hannum, GrimAge, PhenoAge epigenetic clocks were computed according to the method described by Higgins-Chen et al. using the R code hosted on GitHub (https://github.com/MorganLevineLab/PC-Clocks) using R (version 4.4.2). DunedinPACE was calculated according to the method described by Belsky et al. using the R code hosted on GitHub (https://github.com/danbelsky/DunedinPACE/) using R (version 4.4.2). Statistical analyses were performed using lmer, lmerTest, mmrm, and emmeans R packages.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Due to participant confidentiality and privacy concerns, COSMOS data are available upon request, with formal applications submitted to the COSMOS research committee, to ensure data security and ethical considerations. Please refer to https://cosmostrial.org/for-investigators/ for various forms, information, and resources to propose a manuscript/abstract, or to request data using COSMOS resources. Proposals to access and analyze the data will be reviewed by the COSMOS research committee based on whether the hypothesis has been tested, data availability and cleaning requirements, and the programming time required to prepare the data. A response to a request for access is typically provided within two weeks. De-identified data may be subsequently made available to external researchers subject to a controlled-access system and a data-sharing agreement.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	We collected the biological sex of participants, which was self-reported. All study participants provided written informed consent. As predefined in the study protocol, we tested whether the treatment effects varied by sex (Extended Data Figure 1). Since the interaction for the main findings was not statistically significant, the findings of the present study apply to all sexes.
Reporting on race, ethnicity, or other socially relevant groupings	Race/ethnicity were self-reported by participants. As predefined in the study protocol, we collected and tested whether the treatment effects varied by race.
Population characteristics	Table 1 presents the baseline characteristics in the overall study population and stratified by randomized multivitamin assignment. The mean age was 70.2±5.6 years. 482 (50.3%) were female and 854 (89.1%) were white participants.
Recruitment	Recruitment included mailings to 71,521 active Women's Health Initiative (WHI) Extension Study participants and mailings by Brigham and Women's Hospital (BWH) to 237,736 women and men contacted for, but not randomly assigned to, the VITamin D and OmegA-3 TriaL (VITAL); mass mailings to 2,616,343 US men and women; and mailings to 3380 volunteers through other sources, including responses to media stories and advertisements. Of 191,796 participants completing a brief screening questionnaire, 120,629 were initially eligible and willing; then 63,025 completed another questionnaire with written informed consent before enrollment with oversight by the Human Subjects Committee at BWH. Participants were required to forego cocoa supplements (chocolate intake was not restricted) and multivitamins during the trial (vitamin D was limited to ≤1000 IU/d and calcium to ≤1200 mg/d from all supplements). Safety exclusions included renal failure or dialysis, cirrhosis, other serious conditions that precluded participation, and reported extreme sensitivity to caffeine, as the cocoa extract supplement had modest theobromine (~50 mg/d) and caffeine (~15 mg/d) content. A total of 35,669 eligible, willing, and consenting participants began at least a 2-mo placebo run-in to eliminate poor compliers (defined as taking <75% of the study pills) before randomization to increase study power. At the end of the run-in, participants returned a final compliance, eligibility, and risk factor questionnaire and semi-quantitative food-frequency questionnaire. Finally, a total of 21,442 US adults, including 12,666 women aged ≥65 y and 8776 men aged ≥60 y were included in COSMOS. Baseline biospecimens were obtained during the run-in period from 6867 (32.0%) of 21,442 randomized participants, of whom 1,829 participants provided bloods at baseline, Year 1, and Year 2 of follow-up. In this ancillary study, we randomly selected a total of 998 participants for DNA methylation assessment, with oversampling of non-White participants and balanced by biological sex. We originally based our analyses upon 600 participants through one funded project, then added participants through additionally obtained support. Participants included in the current study were representative of the overall COSMOS Blood subcohort and the overall COSMOS trial.
Ethics oversight	COSMOS trial activities were overseen by the Human Subjects Committee at BWH/Mass General Brigham (2014P002768).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	In this ancillary study, we randomly selected a total of 998 participants for DNA methylation assessment, with oversampling of non-White participants and balanced by biological sex. We originally based our analyses upon 600 participants through one funded project plus participants through additionally obtained support. Participants included in the current study were representative of the overall COSMOS Blood subcohort and the overall COSMOS trial.
Data exclusions	Data were excluded from the analysis due to low quality of DNA extraction, quality control (mismatched sex estimation, poor quality samples, and disagreement in genetic fingerprint), and absence of baseline and follow-up measurements. More details can be found in Figure 1.
Replication	For quality control purpose, we tested 42 blinded split samples from 14 participants as technical replicates. The probe-level pairwise comparison showed excellent reliability for DNA methylation measurements (Mean value: 0.992±0.008 for Pearson correlation; 0.996±0.016 for R square; 0.002±0.002 for mean absolute error; 0.027±0.013 for mean sqaured error).
Randomization	From April 2016 to March 2018, 21,442 participants meeting all eligibility criteria were randomized according to a schedule prepared and implemented by WHI staff to 1 of 4 arms in equal proportions: 1) active cocoa extract and active multivitamin, 2) active cocoa extract and multivitamin placebo, 3) active multivitamin and cocoa extract placebo, or 4) both placebos, using a computer-generated permuted block approach blinded to investigators and stratified by sex (women, men), age (separate 5-y age blocks for women and men), and recruitment source (WHI, BWH) in blocks of 12. Participants within a household were randomly assigned to the same intervention, when possible, to reduce cross-contamination risk.
Blinding	Participants, investigators, staff dispensing study pills and collecting outcomes, and data analysts were masked to group assignment.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	NCT02422745; NCT05510375
Study protocol	The study protocol has been previously published in Rist PM, Sesso HD, Johnson LG, Aragaki AK, Wang L, Rautiainen S, Hazra A, Tobias DK, LeBoff MS, Schroeter H, Friedenberg G, Copeland T, Clar A, Tinker LF, Hunt RP, Bassuk SS, Sarkissian A, Smith DC, Pereira E, Carrick WR, Wion ES, Schoenberg J, Anderson GL, Manson JE; COSMOS Research Group. Design and baseline characteristics of participants in the COcoa Supplement and Multivitamin Outcomes Study (COSMOS). Contemp Clin Trials. 2022 May;116:106728. doi: 10.1016/j.cct.2022.106728. Epub 2022 Mar 12. PMID: 35288332; PMCID: PMC9133193. The full study protocol can be found in the website: https://cdn.clinicaltrials.gov/large-docs/45/NCT02422745/Prot_000.pdf.
Data collection	The COSMOS randomized a total of 21,442 older adults across US from April 2016 to March 2018, and the randomized treatments continued through 31 December 2020, ending as scheduled, with a median (IQR) treatment period of 3.6 (3.2, 4.2) y.
Outcomes	We focused on five common DNAm measures of biological aging, including the first-generation clocks predicting chronological age (Horvath and Hannum), second-generation clocks predicting mortality outcomes (PhenoAge and GrimAge), and third-generation clocks of epigenetic metrics developed to predict system integrity and pace of aging (DunedinPACE). To ensure that values of the first two generations of clocks were highly reproducible, we used the principal component (PC) versions using the method developed by Higgins-Chen et al. Then we residualized these clocks for chronological age to estimate epigenetic age acceleration at each time point, with positive values meaning accelerated aging and vice versa. As DunedinPACE was initially created using CpGs with good

replication performance to achieve high technical reliability, it was directly calculated using the method developed by Belsky et al. DunedinPACE corresponds to the number of biologic years per chronological year, with values above 1 indicating faster pace of aging.

Plants

Seed stocks: N/A

Novel plant genotypes: N/A

Authentication: N/A