Associations between patterns of modifiable risk factors in midlife to late life and longevity: 36 year prospective cohort study

Ming Ding 1,2, Garrett M Fitzmaurice 3,4,5, Mariel Arvizu 1,6, Walter C Willett 1,6,7, JoAnn E Manson, 6,7, Kathryn M Rexrode, 9, Frank B Hu 1,6,7, Jorge E Chavarro 1,6,7

ABSTRACT
OBJECTIVE To examine the associations between patterns of mid-life to late life modifiable risk factors and longevity.

DESIGN Prospective cohort study.

SETTING Data collected from the Nurses’ Health Study starting in 1984 and the Health Professionals Follow-up Study starting in 1986.

PARTICIPANTS 85,346 participants from the Nurses’ Health Study and the Health Professionals Follow-up Study.

MAIN OUTCOME MEASURES Death from any cause by 31 October 2020 for the Nurses’ Health Study and Health Professionals Follow-up Study. Risk factors investigated were body mass index, physical activity, alcohol intake, smoking status, and quality of diet. Trajectories of each risk factor and trajectories of changes in the risk factor were identified from baseline with smoothing mixture models, and the joint group memberships of participants was used to most efficiently capture patterns of the factor over time. For each risk factor, three trajectories (patterns with high, medium, and lower values) and three trajectories of change in the risk factor (patterns with increase, no change, and decrease in the factor from baseline) were assumed, giving nine joint patterns: high-stable, high-increase, high-decrease, medium-stable, medium-increase, medium-decrease, low-stable, low-increase, and low-decrease. Associations between patterns of modifiable risk factors and longevity (age at death ≥85 years) and life expectancy were examined with logistic regression and accelerated failure time models, respectively.

RESULTS The analysis included 85,346 participants, with 46,042 participants achieving longevity and 25,322 participants achieving healthy longevity (those who did not have a diagnosis of cardiovascular disease, type 2 diabetes, or cancer). Mean age at baseline was 56 years (standard deviation 5 years). Maximum longevity was achieved in participants with a low-stable pattern for body mass index (compared with a medium-stable pattern, odds ratio of longevity of 1.05, 95% confidence interval 1.00 to 1.10); those with a medium-increase pattern for physical activity (compared with a medium-stable pattern, odds ratio 1.08, 1.01 to 1.15); those with a medium-stable pattern for alcohol intake (high-increase v medium-stable pattern, odds ratio 0.83, 0.74 to 0.93); those who never smoked (low-stable v medium-stable pattern, odds ratio 3.09, 2.84 to 3.37); and those who with a high-increase pattern for quality of diet (compared with a medium-stable pattern, odds ratio 1.09, 1.01 to 1.18). The associations between each factor and life expectancy and healthy longevity (no diagnosis of cardiovascular disease, type 2 diabetes, or cancer) were similar to those for longevity.

CONCLUSIONS During mid-life and late life, maximum longevity was achieved in participants who maintained a normal body mass index, never smoked, ate a healthy diet, and who had physical activity levels and alcohol consumption that met public health recommendations.

WHAT IS ALREADY KNOWN ON THIS TOPIC
⇒ Modifiable risk factors are known to influence mortality
⇒ Obesity, smoking, physical inactivity, and low quality diet have been linked to greater risks of premature morbidity and mortality but few studies have examined the associations over the course of a lifetime

WHAT THIS STUDY ADDS
⇒ During mid-life and late life, maximum longevity was achieved in participants who maintained a normal body mass index, never smoked, ate a healthy diet, and who had physical activity levels and alcohol consumption that met public health recommendations

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE, OR POLICY
⇒ This study provides important evidence that maintaining healthy behaviours should be recommended to individuals not only at young ages, but also through mid-life to late adulthood

Introduction
Life expectancy has increased considerably worldwide.1 In the US, life expectancy was 78.9 years in 2019, lower than other high income countries, which had an average life expectancy of 81.3 years.2 Prolonged life expectancy has seen a growth in age related diseases, which impair quality of life and impose a substantial medical burden on society.3 Thus, how to achieve longevity, and longevity free from chronic diseases, in the US population are important questions.

Modifiable risk factors, including those that can be modified by specific lifestyle factors, such as obesity, are known to have a major effect on mortality. For example, obesity, smoking, physical inactivity, and a low quality diet have been linked to higher risks of premature morbidity and mortality but further investigation is needed.4–7 Firstly, studies on lifestyle and...
mortality usually use baseline or cumulative data, and the role of changes in behaviour over time can be less clear. Secondly, methodological advances to characterise trajectories in behavioural factors over time, including the recently developed smoothing mixture model, are under used. These methods could be used creatively to efficiently capture patterns of modifiable risk factors. Thirdly, although research has been conducted on mortality, few studies have taken into account morbidity status, such as cardi-vascular disease and cancer, which is important in terms of the effect on public health.

In this study, we examined associations between patterns of mid-life to late life modifiable risk factors and longevity, with data from the Nurses’ Health Study and the Health Professionals Follow-up Study. These studies have more than 30 years of follow-up, multiple repeated measures of risk factors, and longevity status for most participants. Figure 1 shows the visual abstract.

Methods
Study population
The Nurses’ Health Study began enrolling participants in 1976, when 121700 female registered nurses, aged 30-55 years, residing in 11 US states were recruited to complete a baseline questionnaire with information on lifestyle and medical history. The Health Professionals Follow-up Study recruited 51529 male healthcare professionals in 1986 (dentists, pharmacists, veterinarians, optometrists, osteopathic physicians, and podiatrists), aged 40-75 years at baseline. All participants returned a self-administered baseline questionnaire with a detailed medical history, lifestyle factors, and usual diet. In both cohorts, questionnaire data were collected at baseline and then every two years, to update information on modifiable risk factors and follow-up on the occurrence of chronic diseases.

In this analysis, we used modifiable risk factor data collected from the 1984 and 1986 cycles as baseline for the Nurses’ Health Study and Health Professionals Follow-up Study, respectively. We excluded participants who reported cardiovascular disease, cancer, or type 2 diabetes at baseline, and participants with extreme and implausible daily energy intakes (3500 kcal for women and 4200 kcal for men; 1 kcal=4.18 kJ).

Assessment of modifiable risk factors
Modifiable risk factors measured were body mass index, smoking status, alcohol intake, quality of diet, and physical activity. Participants were asked to self-report body weight and number of cigarettes currently smoked in questionnaires completed every two years. Physical activity was assessed by previously validated questionnaires to self-report the amount of time spent each week on different physical activities: walking, jogging, running, bicycling, calisthenics, aerobics, aerobic dance, use of rowing machine, lap swimming, playing tennis, and playing squash or racquet ball. Weekly energy expenditure in metabolic equivalent task hours (MET hours/week) was calculated. An 131 item food frequency questionnaire was given every four years to update dietary information from participants in the Nurses’ Health Study and the Health Professionals Follow-up Study. Participants were asked how often (from “never or less than once per month” to “six or more times per day”), on average, they consumed a standard portion size of each food item during the previous year. Questions about consumption of alcoholic beverages (beer, wine, and liquor) were included in each questionnaire.

We used the Alternative Healthy Eating Index 2010 to measure the quality of the diet, based on intake levels of 10 components: fruit, vegetables, whole grains, long chain omega 3 fats, nuts and legumes, polyunsaturated fatty acids, sugar sweetened beverages, red and processed meat, trans fat, and sodium. The total score of the Alternative Healthy Eating Index 2010 ranged from 0 to 100, with a higher score indicating a better quality of diet. Online supplemental file 1 details information on how the modifiable risk factors were assessed. Online supplemental table S1 shows the time of the assessments of the modifiable risk factors in the Nurses’ Health Study and the Health Professionals Follow-up Study.
Assessment of mortality and longevity
Our primary endpoint was death from any cause by 31 October 2020 for the Nurses’ Health Study and the Health Professionals Follow-up Study. In both cohorts, mortality data were collected from a systematic search of state vital records and the national death index database. The search was supplemented by reports from family members and postal authorities. These methods provided more than 98% of deaths in each cohort.12 A validation study conducted by physicians who were blinded to data on risk factors reviewed the death certificates and medical records to classify the cause of death according to ICD-8 and ICD-9 (international classification of diseases, eighth and ninth revisions).

Longevity was defined as survival to age ≥85 years, according to the life expectancy of the US population and the vital status of our data.9 We found that 66% of participants in the Nurses’ Health Study and 68% of participants in the Health Professionals Follow-up Study had achieved longevity (ie, participants were alive aged >85 years by 31 October 2020), and we excluded participants whose longevity status was not known at the end of follow-up. We also looked at longevity in participants who did not have a diagnosis of type 2 diabetes (ICD-8 code 250, ICD-9 code 2500), cardiovascular disease (ICD-8 codes 390-459 or 795, ICD-9 codes 390-4590 or 7950), or cancer (ICD-8 codes 140-207, ICD-9 codes 1400-2070).

Assessment of covariates
In the Nurses’ Health Study, age, race, use of aspirin, menopausal status, use of postmenopausal hormones, and family history of myocardial infarction were collected in the baseline questionnaire in 1984. Information on annual family income, use of multivitamins, and a family history of cancer or type 2 diabetes were collected in 1986. Information on education was asked in the questionnaire in 1992. In the Health Professionals Follow-up Study, age, race, work status, family history of myocardial infarction, cancer, or type 2 diabetes, and use of aspirin, antihypertensive agents (including β blockers, calcium channel blockers, and nitrates), and cholesterol lowering drugs were collected with the questionnaire at baseline in 1986, and use of multivitamins use was collected in 1988.

Statistical analysis
We identified patterns of risk factors with smooth mixed models, which describes trajectories with high flexibility with the use of smoothing functions of time and thus improves the accuracy of group classification.8 With the smoothing mixture model, we assumed smoothing functions for age and allowed for random effects of individuals classified in the same group. The R script of the smoothing mixture model was accessed in Github (https://github.com/lingding-hsph/Smoothing-mixture-model). We investigated the sources of variances for all risk factors, and found that the between-person variance was significantly higher than within-person variance for each risk factor We investigated the sources of variances for all risk factors, and found that the between-person variance was significantly higher than within-person variance for each risk factor (online supplemental table S2), indicating that patterns of risk factors identified would be largely driven by variations between individuals. Hence to capture patterns of a risk factor over time most efficiently, we simultaneously identified trajectories of the risk factor and trajectories of change in the risk factor (calculated as the difference in risk factor from baseline at each assessment during follow-up) with the smoothing mixture model, and used the joint group memberships to classify participants.

To allow for meaningful interpretation, for each risk factor, we assumed three trajectories for the risk factor (patterns with high, medium, and low values) and three trajectories of change in the risk factor (patterns with increase, no change, and decrease in the factor from baseline), which gave nine joint patterns: high-stable, high-increase, high-decrease, medium-stable, medium-increase, medium-decrease, low-stable, low-increase, and low-decrease. Because the smoothing mixture model is a new method, we also identified trajectories of risk factors and trajectories of change in risk factors with group based trajectory analysis.13 The online supplemental file provides detailed information on the smoothing mixture model, variance decomposition, joint group membership of trajectories of risk factors and trajectories of change in risk factors, and group based trajectory analysis.

To minimise the possibility of reverse causation, we censored risk factors reported after a diagnosis of cardiovascular disease, type 2 diabetes, or cancer, and risk factors reported after age 85 years. We also excluded individuals with less than two measurements for derivation of trajectories (online supplemental figures S1, S2). To remove the effects of temporal trends in dietary intake and physical activity,14 we standardised physical activity and dietary factors (alcohol intake and Alternative Healthy Eating Index score) during each follow-up cycle,15 and mapped the standardised values to real values based on the mean and standard deviation over the whole follow-up period with the inverse cumulative distribution function.

For each risk factor, we examined associations between patterns of risk factor and longevity with logistic regression models, and associations between patterns of risk factor and healthy longevity (defined as individuals who did not have a diagnosis of cardiovascular disease, type 2 diabetes, or cancer) with multinomial logistic models. The models were adjusted for age (continuous), race (white v
non-white), family history of cancer (yes, no), myocardial infarction (yes, no), type 2 diabetes (yes, no), multivitamin use (yes, no), menopausal status (yes, no, women only), postmenopausal hormone use (yes, no, women only), education (registered nurse, bachelor degree, master degree and higher, women only), social economic status (annual family income (four groups) for women and work status (disabled, retired, part time, full time) for men), and the other four factors at baseline (continuous variables).

In models that assessed the association between patterns of change in risk factor and longevity, we also adjusted for the pattern of that factor (categorical). We examined associations between patterns of risk factor and life expectancy with the accelerated failure time model. We modelled survival time as a risk factor and life expectancy with the accelerated failure time model. We examined associations between patterns of risk factor and life expectancy with the accelerated failure time model. We modelled survival time as a risk factor and life expectancy with the accelerated failure time model.

In each cohort, we identified trajectories of risk factors as high, medium, and low patterns (online supplemental figures S3 and S4) and trajectories of change in risk factors as increase, no change, and decrease patterns (online supplemental figures S5 and S6). We classified nine joint patterns: high-stable, high-increase, high-decrease, medium-stable, medium-increase, medium-decrease, low-stable, low-increase, and low-decrease (online supplemental figures S7 and S8). The shapes of the curves were similar for women and men and so we combined data for the two cohorts to display the patterns (figure 2).

For body mass index, the low-decrease, low-stable, and low-increase patterns were found mainly in participants who had a normal body weight (body mass index <25) during follow-up; most participants in the three medium patterns were overweight (body mass index 25.0–29.9) in mid-life to late life; and most participants in the three high patterns were obese (body mass index ≥30).

For smoking, we identified a low-stable pattern mainly in never smokers. Participants tended to reduce the number of cigarettes smoked or stop smoking during follow-up, as reflected by the steep curves in the patterns for high-decrease, medium-decrease, and low-decrease. Even for the high-stable and high-increase patterns, participants reduced the intensity of smoking in late life.

We classified trajectories for alcohol intake as high, medium, and low patterns. For example, participants were labelled as having a high pattern if the mean amount of intake of that group was the highest, and were labelled as low pattern if the mean amount of intake of that group was the lowest. To quantitatively describe the amount of intake for each group, mean values in the low and medium patterns were within the range of moderate alcohol intake recommended by the 2020-25 dietary guidelines for Americans (<30 g/day for men and <15 g/day for women).16 For physical activity, our study population was generally physically active: mean values for the three low patterns were close to the minimum level (7.5 MET hours/week) recommended by the 2008 physical activity guidelines for Americans;17 the mean values for the three medium patterns were slightly above the recommended level for more benefits (15 MET hours/week); and the mean values for the high-stable pattern were six times the recommended minimum level.
Online supplemental table S4 shows the associations between trajectories of risk factors and trajectories of change in risk factors and longevity in women and men separately. We found that the associations for all risk factors in women and men were similar. Hence we pooled the two cohorts; table 1 shows the pooled associations between trajectories of risk factors and trajectories of change in risk factors and longevity, figure 3 shows the pooled associations between joint patterns and longevity and healthy longevity (ie, no diagnosis of cardiovascular disease, type 2 diabetes, or cancer), and figure 4 shows the pooled associations between joint patterns and life expectancy. (Here and below, we report our findings according to risk factors rather than order of figures and tables. For each factor, we reported our findings on longevity and healthy longevity in table 1 and figure 3 and life expectancy in figure 4.)

We found that maximum longevity was achieved in those with a low-stable pattern for body mass index: compared with the medium-stable pattern, the odds ratio was 1.05 (95% confidence interval 1.00 to 1.10) of achieving longevity and 1.19 (1.12 to 1.25) of achieving healthy longevity (ie, no diagnosis of cardiovascular disease, type 2 diabetes, or cancer). Participants with the three high body mass index patterns, as well as those with a low-decrease pattern, were less likely to achieve longevity and were associated with lower life expectancy compared with those with the medium-stable pattern.

For patterns of physical activity, participants with the three low patterns were less likely to achieve longevity compared with those with a medium pattern. We found the highest longevity benefits among participants with a medium-increase pattern (compared with a medium-stable pattern, the odds ratio for longevity was 1.08, 95% confidence interval 1.01 to 1.15; odds ratio for healthy longevity 1.17, 1.08 to 1.25; difference in life expectancy 5.07%, 95% confidence interval 1.36% to 8.92%). No further gain in longevity was found for the high-stable pattern for physical activity, and a decrease in longevity was found for the high-decrease pattern.
For alcohol intake, the medium-stable pattern seemed to be most likely associated with achieving longevity; compared with the medium-stable pattern, the odds ratio was 0.83 (95% confidence interval 0.74 to 0.93) higher for achieving longevity and 0.83 (0.72 to 0.94) higher for achieving healthy longevity, and 7.14% (1.13% to 12.79%) lower for life expectancy for the high-increase pattern. We identified strong dose-response relations between smoking patterns and longevity. Those with a low-stable pattern (mainly never smokers) achieved the highest longevity, followed by a low-decrease pattern (light smokers who stopped smoking during follow-up), a low-increase pattern (light smokers with a long duration of smoking), and participants in the medium patterns. Participants with a high-increase pattern (heavy smokers with a long duration of smoking) were least likely to achieve longevity; compared with those with a medium-stable pattern, a low-stable pattern was associated with an odds ratio

---

**Table 1** | Associations between trajectories of modifiable risk factors and trajectories of change in risk factors from baseline and odds ratio (95% confidence interval) of achieving longevity by pooling data from the Nurses’ Health Study and the Health Professionals Follow-up Study

| Risk factor | No of participants | Model 1 or 3 | Model 2 or 4 |
|-------------|--------------------|--------------|--------------|
| **Body mass index** | | | |
| No of participants who achieved longevity/Total No of participants | 18 655/32 857 | 22 105/38 051 | 3591/8754 |
| 25 095/38 695 | 4071/9809 | 11 615/22 748 |
| No change | 1.00 | 1.05 (1.02 to 1.08) | 0.52 (0.50 to 0.55) |
| Increase | 1.00 | 0.43 (0.41 to 0.45) | 0.56 (0.54 to 0.58) |
| Decrease | 1.00 | 0.72 (0.68 to 0.76) | 0.56 (0.54 to 0.59) |

| Smoking | | | |
| No of participants/Total No of participants | 3105/8817 | 40 637/67 990 | 1095/4388 |
| 39 925/67 287 | 527/1647 | 2336/6772 |
| No change | 1.00 | 2.93 (2.80, 3.07) | 0.60 (0.56 to 0.66) |
| Increase | 1.00 | 0.49 (0.43 to 0.55) | 1.31 (1.20 to 1.44) |
| Decrease | 1.00 | 0.63 (0.55 to 0.72) | 1.26 (1.14 to 1.40) |

| Diet quality (AHEI) | | | |
| No of participants/Total No of participants | 20 439/35 470 | 12 395/25 111 | 11 808/18 972 |
| 20 018/32 453 | 11 144/18 895 | 9951/17 717 |
| No change | 1.00 | 0.72 (0.69 to 0.74) | 1.22 (1.18 to 1.26) |
| Increase | 1.00 | 1.01 (0.96 to 1.05) | 0.66 (0.63 to 0.69) |
| Decrease | 1.00 | 0.63 (0.55 to 0.72) | 1.26 (1.14 to 1.40) |

| Physical activity | | | |
| No of participants/Total No of participants | 14 600/23 867 | 28 071/54 329 | 26 297/45 777 |
| 16 862/30 364 | 23 055/40 387 | 1259/2570 |
| No change | 1.00 | 0.68 (0.66 to 0.70) | 0.86 (0.80 to 0.91) |
| Increase | 1.00 | 1.01 (0.97 to 1.06) | 0.56 (0.50 to 0.62) |
| Decrease | 1.00 | 0.71 (0.69 to 0.80) | 1.00 (1.05 to 1.15) |

| Alcohol intake | | | |
| No of participants/Total No of participants | 10 418/18 068 | 31 402/55 486 | 2872/6076 |
| 453/2570 | 3680/6602 | 1973/4106 |
| No change | 1.00 | 0.94 (0.91 to 0.98) | 0.66 (0.62 to 0.70) |
| Increase | 1.00 | 0.78 (0.74 to 0.83) | 0.64 (0.58 to 0.69) |
| Decrease | 1.00 | 0.91 (0.87 to 0.94) | 0.77 (0.72 to 0.82) |

AHEI=Alternative Healthy Eating Index.
Model 1 is the univariate analysis.
Model 2 adjusted for baseline age (continuous), race (white, black, Asian, and other), family history of cancer (yes, no), myocardial infarction (yes, no), or type 2 diabetes (yes, no), multivitamin use (yes, no), menopausal status (yes, no, women only), postmenopausal hormone use (yes, no, women only), cohort, education (registered nurse, bachelor degree, master degree, and higher, women only), social economic status (annual family income (four groups) for women and work status (disabled, retired, part time, full time) for men), aspirin use (yes, no), use of antihypertensive agents (yes, no), and other four risk factors at baseline as continuous (annual family income (four groups) for women and work status (disabled, retired, part time, full time) for men), aspirin use (yes, no), use of antihypertensive agents (yes, no), use of cholesterol lowering drugs (yes, no), and the other four risk factors at baseline as continuous variables.
Model 3 applies to changes in risk factors and is adjusted for all variables included in model 1, as well as group membership of trajectories of the risk factor (categorical).
Model 4 applies to changes in risk factors and is adjusted for all variables included in model 2, as well as group membership of trajectories of the risk factor (categorical).
of 3.09 (95% confidence interval 2.84 to 3.37) of achieving longevity, 3.52 (3.17 to 3.91) of achieving healthy longevity, and 59.27% (52.57% to 66.26%) higher life expectancy. Conversely, an inverse dose-response relation was found for adherence to the Alternative Healthy Eating Index score and longevity.
status. Compared with those with a medium-stable pattern for the Alternative Healthy Eating Index, a high-increase pattern had 1.09 (95% confidence interval 1.01 to 1.18) times the odds of achieving longevity, and 5.08% (0.59% to 9.78%) higher life expectancy.

We assigned a score to the joint pattern of each factor, ranging from 1 to 9, with a higher score indicating a healthier lifestyle (online supplemental table S5). We then created a score summing up the score for each risk factor and we found that a healthier lifestyle across adulthood seemed more likely to achieve longevity (table 2). Compared with those with the most healthy lifestyle, participants with the least healthy lifestyle were 70% (95% confidence interval 68% to 71%) less likely to achieve longevity, 77% (76% to 78%) less likely to achieve healthy longevity, and had a 53% (50% to 56%) lower life expectancy.

Because the smoothing mixture model is a newly developed model, we also used group based trajectory analysis to identify trajectories of risk factors and trajectories of change in risk factors. Although not as flexible, the shapes of trajectories were similar to those with the smoothing mixture model (online supplemental figures S9-S12). Also, group memberships classified with the group based trajectory analysis and with the smoothing mixture model were highly correlated (online supplemental table S6).

The pattern of risk factors had trajectories with different start ages, and therefore we examined whether age at baseline modified the associations between joint patterns of risk factors and longevity. Online supplemental table S7 shows that adhering to a healthy lifestyle, including maintaining a normal body mass index, never smoking, eating a healthy diet, being physically active, and having moderate
alcohol consumption had significantly more beneficial effects on longevity among participants who were younger at baseline (P for interaction <0.001 for all risk factors), suggesting that participants should adhere to a healthy lifestyle at an early age.

Online supplemental table S8 shows the percentage of missing risk factors at each assessment. Missing rates for body mass index, smoking, and physical activity were low, and missing rates for quality of diet and alcohol intake were moderate. Online supplemental tables S9 and S10 show the distributions for total number of assessments during the follow-up period. Although censoring risk factors reduced the total number of measurements, most participants had at least two measurements, even after censoring risk factors. We further imputed missing data by carrying forward the values measured most closely before a diagnosis of cardiovascular disease, type 2 diabetes, or cancer. We found that the shape of the joint patterns of risk factors and the associations with longevity were similar to our main findings (online supplemental figure S13). We conducted two sensitivity analyses to evaluate whether exclusion of these participants influenced our findings. In the first analysis, we included those who were excluded, used mortality as a surrogate for longevity, and examined the associations between patterns of risk factors and risk of mortality (online supplemental table S14).

In our main analysis, we excluded participants whose longevity was unknown (ie, participants who were aged <51 years in the Nurses' Health Study and <49 years in the Health Professionals Follow-up Study at baseline and still alive at the end of follow-up). The population characteristics of participants who were excluded were similar to the main study population, indicating that excluding those participants would cause minimal selection bias (online supplemental table S13). We conducted two sensitivity analyses to evaluate whether exclusion of these participants influenced our findings. In the first analysis, we included those who were excluded, used mortality as a surrogate for longevity, and examined the associations between patterns of risk factors and risk of mortality (online supplemental table S14).

In the second analysis, we restricted the population to those aged ≥51 years in the Nurses' Health Study and ≥49 years in the Health Professionals Follow-up Study at baseline, because longevity status was known for all of the participants in this population (online supplemental table S15). The findings from both analyses were similar to the main findings, suggesting that our main findings are relatively robust to selection of participants.

**Table 2 | Associations between combination of joint patterns of the five modifiable risk factors and longevity (odds ratios and 95% confidence intervals) by pooling data from the Nurses' Health Study and the Health Professionals Follow-up Study**

|                | Group 1 (most healthy) | Group 2 | Group 3 | Group 4 | Group 5 (least healthy) |
|----------------|------------------------|---------|---------|---------|-------------------------|
| Total No of participants | 15 357                | 18 139  | 15 984  | 13 601  | 16 171                  |
| Longevity: No of participants who achieved longevity | 10 401            | 11 671  | 9 350   | 7 040   | 5 756                   |
| Model 1 | 1.00                   | 0.86 (0.82 to 0.90) | 0.67 (0.64 to 0.70) | 0.51 (0.49 to 0.54) | 0.26 (0.25 to 0.28) |
| Model 2 | 1.00                   | 0.84 (0.80 to 0.88) | 0.67 (0.64 to 0.70) | 0.52 (0.50 to 0.55) | 0.30 (0.29 to 0.32) |
| Healthy longevity*: No of participants who achieved healthy longevity | 6474            | 6962    | 5378    | 3724    | 2733                    |
| Model 1 | 1.00                   | 0.82 (0.78 to 0.87) | 0.62 (0.59 to 0.65) | 0.44 (0.41 to 0.46) | 0.20 (0.19 to 0.21) |
| Model 2 | 1.00                   | 0.80 (0.76 to 0.85) | 0.61 (0.58 to 0.65) | 0.45 (0.42 to 0.47) | 0.23 (0.22 to 0.24) |
| Life expectancy (%): Model 1 | 1.00               | -5 (-6 to -4) | -10 (-11 to -9) | -14 (-15 to -13) | -25 (-26 to -24) |
| Model 2 | 1.00                   | -10 (-12 to -7) | -20 (-22 to -17) | -29 (-32 to -26) | -53 (-56 to -50) |

Model 1 is the univariate analysis. Model 2 adjusted for baseline age (continuous), race (white, black, Asian, and other), family history of cancer (yes, no), myocardial infarction (yes, no), or type 2 diabetes (yes, no), multivitamin use (yes, no), menopausal status (yes, no, women only), postmenopausal hormone use (yes, no women only), cohort, education (registered nurse, bachelor degree, master degree and higher, women only), social economic status (annual family income (four groups) for women and work status (disabled, retired, part time, full time) for men), and aspirin use (yes, no), use of antihypertensive agents (yes, no), and use of cholesterol lowering drugs (yes, no).

Logistic model (longevity), multinomial logistic model (longevity free from diagnosed CVD, type 2 diabetes, and cancer), and accelerated failure time model (life expectancy) were used.

*Participants who did not have a diagnosis of cardiovascular disease, type 2 diabetes, or cancer.
Discussion
Principal findings
Our analysis followed 85,346 men and women over a period of 36 years, and we found that 46,042 participants achieved longevity (alive at 85 years) and 25,322 participants achieved healthy longevity (i.e., no diagnosis of cardiovascular disease, type 2 diabetes, or cancer). During mid-life and late life, we found that maximum longevity was achieved in participants maintaining a normal body mass index, who never smoked, ate a healthy diet, and had physical activity levels and alcohol consumption that met public health recommendations. The similarity of the associations between risk factors and longevity for both women and men further strengthens our findings. Previous studies have shown that modifiable risk factors, including not smoking and being physically active, were associated with longer survival.18 19 Our study adds new evidence to the topic from the perspective of the course of a lifetime.

We found that participants with a low-stable pattern for body mass index were most likely to achieve longevity and healthy longevity, and this finding is consistent with previous studies involving millions of participants that showed that body mass index in the range 21-25 was associated with the lowest risk of mortality.20-24 Moreover, our study suggested no further gain in longevity for the low-increase, medium-increase, and high-increase patterns in body mass index, in agreement with previous studies showing that even moderate weight gain during mid-life was associated with a much higher risk of mortality and major chronic diseases.25 Our findings for participants with low-decrease, medium-decrease, and high-decrease patterns indicated that participants with weight loss were less likely to achieve longevity. The reason might be reverse causation, because participants with underlying medical conditions might lose weight before the symptoms of disease appeared. Although we censored body mass index before a diagnosis of cardiovascular disease, cancer, or type 2 diabetes, reverse causation bias is a plausible explanation, and previous research showed that unintentional weight loss was associated with higher mortality rates.26

Comparison with other studies
The association between physical activity with mortality has been extensively studied,27 and the 2008 physical activity guidelines for Americans recommends a minimum of 7.5 MET hours/week of physical activity for health benefits (equivalent to 150 minutes/week of moderate intensity activity or 75 minutes/week of vigorous physical activity), and 15.0 MET hours/week for more benefits.17 However, whether greater health benefits can be achieved with physical activity levels >15.0 MET hours/week or whether a threshold of physical activity exists for mortality benefits is uncertain. Previous studies showed a non-linear association, with the lowest risk of mortality at 21-35 MET hours/week.28 29 We also found that a medium-increase pattern (20-30 MET hours/week) in mid-life to late life was associated with the highest probability of longevity and healthy longevity, and that the benefit to longevity persisted for the high patterns, although no further gains in longevity were found. Moreover, our study showed that an increase in physical activity during a lifetime was associated with greater longevity, but the benefits to longevity in mid-life were lost for participants who reduced their physical activity levels. Our findings are consistent with the EPIC (European Prospective Investigation into Cancer) Norfolk study,30 which looked at three measures of physical activity and assumed linear associations for participants’ trajectories. One explanation for reduced life expectancy in those with a decrease pattern could be reverse causation, because participants with chronic diseases, such as cardiovascular disease and cancer, might reduce their amount of physical activity. Another explanation is that recent physical activity might have a more important role in affecting health.

We detected a dose-response relation between patterns for quality of diet and longevity, with those with a high-increase pattern achieving highest longevity and healthy longevity. Our findings are in line with previous studies showing that a higher quality diet was associated with a lower risk of mortality31 32 and that improving the quality of your diet was associated with a lower risk of death.6 Randomised clinical trials have shown that a healthy diet promotes weight loss,33 substantially lowers blood pressure,34 35 and efficiently prevents cardiovascular disease.16 A plausible mechanism is that the beneficial effects might be mediated through improvement in endothelial function, inflammatory markers, and insulin resistance,37 and the foods and nutrients promoted by the Alternative Healthy Eating Index (fruits, vegetables, whole grain, and unsaturated fatty acids) might account for the effect. Overall, our study adds to the evidence that individuals should adhere to a healthy diet in mid-life to late life to boost longevity.

The health benefits of alcohol intake are controversial, particularly for moderate intake of alcohol. Our study showed that individuals with the medium pattern were more likely to achieve longevity and healthy longevity than those with the low and high patterns, supporting the 2020-25 dietary guidelines for Americans of no more than two drinks (~30 g) a day for men and no more than one drink (~15 g) a day for women.16 Our findings are in agreement with previous studies. One meta-analysis involving more than a million participants showed a non-linear association between intake of alcohol and total mortality, with the lowest risk at 1-2 drinks/day for women and 2-4 drinks/day for men.38 Prospective studies found that moderate intake of alcohol was
associated with the lowest risk of total mortality and the highest probability of achieving longevity.\textsuperscript{39, 40} In the Global Burden of Disease study, however, with data pooled from 195 countries, the risk of mortality increased with higher levels of consumption, including moderate intake of alcohol.\textsuperscript{43} The reason for the inconsistency is because of the varied distribution of causes for cause specific mortality between developing and developed countries. In fact, the Global Burden of Disease study found that moderate intake of alcohol was associated with a lower risk of cardiovascular disease and type 2 diabetes, the main drivers of mortality in developed countries.\textsuperscript{53} Longevity of moderate intake of alcohol can have benefits; alcohol raises levels of high density lipoprotein and improves insulin sensitivity, which protect against cardiovascular disease and type 2 diabetes.\textsuperscript{52} Alcohol intake with family and friends has social and psychological benefits that might also contribute to health and wellbeing.\textsuperscript{53} We found that high patterns of alcohol intake were less likely to be associated with longevity, consistent with previous studies.\textsuperscript{58–40} and evidence has been convincing that alcohol consumption increases the risk of breast cancer.\textsuperscript{44}

Smoking was the leading modifiable risk factor for mortality worldwide in 2015.\textsuperscript{45} and 11.5% of global deaths were attributed to smoking.\textsuperscript{55} In our study, smoking showed the strongest association with longevity compared with other risk factors, highlighting the importance of stopping smoking in the promotion of longevity. We found that an increase in intensity of smoking shortens the lifespan and a decrease in smoking extends life. Although the prevalence of smoking has been decreasing over the past 30 years, the rate is still high (ie, 25.0% for men and 5.4% for women worldwide).\textsuperscript{56} Our study supports an important public health message that it is never too late to stop smoking and gain benefits for healthy longevity.

**Strengths and limitations of the study**

Our study benefited from appropriate use of advanced statistical analyses and strict sensitivity analyses. Firstly, we creatively captured individual’s risk factor patterns by classifying participants based on joint group membership of trajectories of risk factors and change in risk factors. We used the smoothing mixture model to derive patterns of risk factors, which models trajectories with high flexibility and reduces the probability of misclassification of trajectories.\textsuperscript{5} Secondly, although participants who achieved longevity tended to have more repeated measures of risk factors than those who died before age 85 years, classification of risk factor patterns was independent of the number of measurements and thus independent of outcome.\textsuperscript{5} Thirdly, because censoring time is dependent on longevity as an outcome, implementing survival analysis or competing risk analysis to examine longevity can result in selection bias.\textsuperscript{47} Thus the use of logistic regression and multinomial regression dealt with the possibility of biased estimates. To leverage censoring time, however, we also used the accelerated failure time model to examine associations between risk factor patterns and life expectancy. Fourthly, because the start age of each individual’s trajectory varied, we examined whether age at baseline affected the associations between joint patterns of risk factors and longevity, and found that participants should adhere to a healthy lifestyle at an earlier age to increase the odds of longevity.

Our study had several limitations. Firstly, the Nurses’ Health Study and Health Professionals Follow-up Study included predominantly white healthcare professionals, which could limit the generalisability of our findings to other nationalities and races or ethnicities. However, participants’ health related occupations were an advantage that allowed us to collect high quality data with self-reported questionnaires and enhance the internal validity of the study by reducing confounding. Secondly, risk factors were self-reported by questionnaire, and measurement error was inevitable. Our food frequency questionnaires have been extensively validated against diet records.\textsuperscript{48–51} however, and body mass index and physical activity have been validated against standard measurements.\textsuperscript{52} Thirdly, given the observational design of the study, we could not directly establish causal relations between risk factor patterns and longevity. Fourthly, we censored risk factors after a diagnosis of cardiovascular disease, type 2 diabetes, or cancer to avoid reverse causation, but this method might result in missing data. We imputed missing data and found that the shape of joint patterns of risk factors and the associations with longevity were similar to our main findings, indicating the robustness of our main findings. Fifthly, we excluded participants whose longevity was unknown, which might cause selection bias. The main reason for the exclusion was that these participants were younger at baseline and were alive without reaching the age of 85 years at the end of follow-up. The population characteristics of the excluded participants and the main study population were similar, however, and we believe that the excluded population would likely give similar findings to our main population. Our sensitivity analyses (online supplemental tables S14 and S15) suggest that exclusion of these participants had minimal effects on our findings.

**Conclusion**

We found that maximum longevity was achieved in participants maintaining a normal body mass index, who never smoked, ate a healthy diet, and had physical activity levels and alcohol intake that...
met public health recommendations through mid-life and late life.

AUTHOR AFFILIATIONS

1Department of Nutrition, Harvard T H Chan School of Public Health, Boston, MA, USA
2Department of Emergency Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
3Laboratory for Psychiatric Biostatistics, McLean Hospital, Harvard Medical School, Boston, MA, USA
4Department of Biostatistics, Harvard T H Chan School of Public Health, Boston, MA, USA
5Department of Psychiatry, Harvard Medical School, Boston, MA, USA
6Department of Epidemiology, Harvard T H Chan School of Public Health, Boston, MA, USA
7Channing Division of Network Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA
8Division of Preventive Medicine, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA
9Division of Women’s Health, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA

Contributors. MD, GF, MA, WW, JEM, KMR, FH, and JEC designed the study and collected the data. MD mainly conducted the analyses and wrote the manuscript. JEC supervised the data analysis and reviewed and edited the manuscript. All authors contributed substantially to the interpretation of the data and the drafting or critical revision of the manuscript for important intellectual content. The authors assume full responsibility for analyses and interpretation of these data. MD and JEC are the guarantors. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. Transparency: The lead author (the guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Funding. This work was supported by grants R01AG060027, P30-DK46200, R37 DA402347, R01 HL134934, R01 HL066921, U01 CA16792, R01 HL43643, and UM1 CA68107 from the National Institutes of Health. The funders had no role in considering the study design or in the collection, analysis, interpretation of data, writing of the report, or decision to submit the article for publication.

Competing interests. All authors have completed the ICMJE uniform disclosure form at www.icmje.org. A complete disclosure is available at https://bmjmedicine.bmj.com/content/1/1/e000098.

Ethics approval. This study involves human participants, and the study protocol was approved by the institutional review boards of the Brigham and Women’s Hospital and Harvard T H Chan School of Public Health, and those of participating registries as required (IRB protocol No 999P011114, 1999P003389, core C 10161). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review. Not commissioned; externally peer reviewed.

Data availability statement. Data are available upon reasonable request.

Supplemental material. This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access. This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs

Ming Ding http://orcid.org/0000-0001-9497-3446
Mariel Anzu http://orcid.org/0000-0003-1860-7726
Jorge E Chavarro http://orcid.org/0000-0002-4436-9630

REFERENCES

1. GBD 2019 Demographics Collaborators. Global age-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950-2019: a comprehensive demographic analysis for the global burden of disease study 2019. Lancet 2020;396:1160–203. doi:10.1016/S0140-6736(20)30777-6
2. Drau V, Inouye SK, Rowe JW, et al. Enabling healthful aging for all - The National Academy of Medicine Grand Challenge in Healthy Longevity. N Engl J Med 2019;381:1699–701. doi:10.1056/NEJMoa1907288
3. Carter BD, Abnet CC, Feskanich D, et al. Smoking and mortality beyond established causes. N Engl J Med 2015;372:631–40. doi:10.1056/NEJMoa150211
4. Ding M, Satija A, Bhupathiraju SN, et al. Association of coffee consumption with total and cause-specific mortality in 3 large prospective cohorts. Circulation 2015;132:2305–15. doi:10.1161/CIRCULATIONAHA.115.013744
5. Ding M, Li J, QL, et al. Associations of dairy intake with risk of mortality in women and men: three prospective cohort studies. BMJ 2019;367:l6204. doi:10.1136/bmj.l6204
6. Sotos-Prietó M, Bhupathiraju SN, Matti J, et al. Association of changes in diet quality with total and cause-specific mortality. N Engl J Med 2017;377:143–53. doi:10.1056/NEJMoa163502
7. Li Y, Schoufoul J, Wang DD, et al. Healthy lifestyle and life expectancy free of cancer, cardiovascular disease, and type 2 diabetes: prospective cohort study. BMJ 2020;368:m1666. doi:10.1136/bmj.m1666
8. Ding M, Chavarro JE, Fitzmaurice GM. Development of a mixture model allowing for smoothing functions of longitudinal trajectories. Stor Methods Med Res 2021;30:569–62. doi:10.1177/0962280220966019
9. Wolf AM, Hunter DJ, Colditz GA, et al. Reproducibility and validity of a self-administered physical activity questionnaire. Int J Epidemiol 1994;23:991–9. doi:10.1093/ije/23.5.991
10. Ainsworth BE, Haskell WL, Leon AS, et al. Compendium of physical activities: classification of energy costs of human physical activities. Med Sci Sports Exerc 1995;275–71. doi:10.1249/00005768-199511000-00020
11. Chiuse SE, Fung TT, Rimm EB, et al. Alternative dietary indices both strongly predict risk of chronic disease. J Nutr 2012;142:1009–18. doi:10.3945/jn.111.157222
12. Rich-Edwards JW, Corsano KA, Stampfer MJ. Test of the National Death Index and Equifax Nationwide Death Search. Am J Epidemiol 1994;140:1046–9. doi:10.1093/oxfordjournals.aje.a117191
13. Nagan DS, Tremblay RE. Analyzing developmental trajectories of distinct but related behaviors: a group-based method. Psychol Methods 2001;6:18–34. doi:10.1037/1082-9894.6.1.18
14. Wang DD, Leung CW, Li Y, et al. Trends in dietary quality among adults in the United States, 1999 through 2010. JAMA Intern Med 2014;174:1587–95. doi:10.1001/jamainternmed.2014.3422
15. Bernstein AM, Rosner BA, Willett WC. Cereal fiber and coronary heart disease: a comparison of modeling approaches for repeated dietary measurements, intermediate outcomes, and long follow-up. Eur J Epidemiol 2011;26:877–86. doi:10.1007/s10654-011-9562-x
16. U.S. Department of Agriculture and U.S. Department of Health and human services. Dietary guidelines for Americans, 2020-2025, 9th Edition, 2020. https://www.dietaryguidelines.gov/sites/default/files/2020-12/Dietary_Guidelines_for_Americans_2020-2025.pdf
17. U.S. Department of Health and Human Services. 2008 physical activity guidelines for Americans. Available: https://health.gov/sites/default/files/2019-09/pguide.pdf
18. Rizzuto D, Orsin N, Qi C, et al. Lifestyle, social factors, and survival after age 75: population based study. BMJ 2012;345:e5568. doi:10.1136/bmj.e5568
19. Wilhelmsen L, Svardsudd K, Eriksson H, et al. Factors associated with reaching 90 years of age: a study of men born in 1913 in Gothenburg, Sweden. J Intern Med 2011;269:441–51. doi:10.1111/j.1365-2796.2010.02313.x
cardiovascular disease with a Mediterranean diet supplemented and mortality in elderly European men. The role of lifestyle behaviors. Circ Res 2017;120:1927–37. doi:10.1161/CIRCRESAHA.116.309443
44 Smith-Warner SA, Spiegelman D, Yuán SS et al. Alcohol and breast cancer in women: a pooled analysis of cohort studies. JAMA 1998;279:535–40. doi:10.1001/jama.279.7.535
45 GBD 2015 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016;388:1651–78. doi:10.1016/S0140-6736(16)31678-9
46 GBD 2015 Tobacco Collaborators. Smoking prevalence and attributable disease burden in 195 countries and territories, 1990–2015: a systematic analysis from the Global Burden of Disease Study 2015. Lancet 2017;389:1888–906. doi:10.1016/S0140-6736(17)30819-X
47 Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;94:496–509. doi:10.1080/01621459.1999.1047444
48 Feskanchich D, Rimm EB, Giovannucci EL et al. Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. Am J Epidemiol 1999;150:793–800. doi:10.1093/aje/150.5.793
49 Rimm EB, Giovannucci EL, Stampfer MJ et al. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. Am J Epidemiol 1992;135:1114–26. discussion 27–36. doi:10.1093/aje/135.5.1114
50 Saelens SB, Hunter DJ, Sampson L et al. Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. Int J Epidemiol 1989;18:858–67. doi:10.1093/ije/18.4.858
51 Willett WC, Sampson L, Stampfer MJ et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 1985;122:51–5. doi:10.1093/aje/122.1.51
52 Rimm EB, Stampfer MJ, Colditz GA et al. Validity of self-reported waist and hip circumferences in men and women. Epidemiology 1990;1:486–73. doi:10.1093/epi/1.4.486
Supplemental material

Assessment of modifiable risk factors

BMI. Body weight and height were asked in biennial questionnaires from 1984 to 2016 in the NHS and from 1986 to 2016 in the HPFS, and BMI was calculated as weight (kg) over height (m) squared. The correlation between measured and self-reported weight was 0.97 for both NHS and HPFS participants. We have 17 and 16 repeated measures of BMI in the NHS and HPFS, respectively.

Smoking. Participants self-reported their smoking status (past, current, never smoker) and the number of cigarettes smoked (1-4, 5-14, 15-24, 25-34, 35-44, and ≥45 cigarettes/day) in biennial questionnaires from 1984 to 2016 in the NHS and from 1986 to 2016 in the HPFS. We categorized smoking status into 13 groups, namely never, former (categorized into 1-4, 5-14, 15-24, 25-34, 35-44, ≥45 cigarettes/day), current (categorized in to 1-4, 5-14, 15-24, 25-34, 35-44, ≥45 cigarettes/day), and assigned numbers 0-12 for the 13 groups. We have 17 and 16 repeated measures of smoking in the NHS and HPFS, respectively.

Alcohol intake. In 1984, a 116-item food frequency questionnaire (FFQ) was administered to the NHS participants to obtain information on usual intake of food and beverages. Starting in 1986, an expanded 131-item FFQ was administered every 4 years to update diet from the NHS and the HPFS participants. In FFQs, participants were asked how often (from “never or less than once per month” to “6 or more times per day”) on average they consumed a standard portion size of each food item during the previous year. The FFQ has shown good validity and reproducibility, and the results have been described in detail elsewhere. Questions about the consumption of alcoholic
beverages (including beer, wine, and liquor) were included in each questionnaire. Alcohol intake was assessed until 2010 in the NHS and 2014 in the HPFS, and we have 8 repeated measures in both cohorts.

**Diet quality: The Alternative Health Eating Index (AHEI).** In each 4-year cycle of diet assessment via an extensively validated food frequency questionnaire, food items were combined into food groups, and a dietary pattern score AHEI-2010 was derived to reflect diet quality. In our analysis, scoring for the AHEI-2010 was based on intake levels of 10 components, excluding alcohol, which were chosen on the basis of their association with chronic disease and mortality risk in observational and interventional studies. The score emphasized higher intakes of fruit, vegetables, whole grains, long-chain omega-3 fats, nuts and legumes, and polyunsaturated fatty acids and lower intakes of sugar-sweetened beverages, red and processed meat, trans fat, and sodium. Each component was scored from 0 (unhealthy) to 10 (healthiest), and the total score ranged from 0 (non-adherence) to 100 (perfect adherence). We have 8 and 6 repeated measures of diet in the NHS and HPFS, respectively.

**Physical activity.** Physical activity was assessed by previously validated questionnaires. Beginning in 1986 participants from both cohorts self-reported amount of time spent per week on each of the following physical activities: walking; jogging; running; bicycling; calisthenics, aerobics, aerobic dance, or rowing machine use; lap swimming; playing tennis; and playing squash or racquet ball. From this information, weekly energy expenditure in metabolic equivalent task-hours (MET-hours) was calculated. The reproducibility and validity of the physical activity questionnaire has been described elsewhere. The correlation for physical activity between two
repeated questionnaires administered two years apart was 0.59, and correlation between physical activity reported in diaries and that reported on questionnaire was 0.62. Physical activity was assessed 11 times in the NHS, and 7 times in the HPFS.

**Smoothing mixture model (SMM)**

We used SMM to identify patterns of each risk factor, which adopted a modified expectation-maximization (EM) algorithm to delineate trajectories with smoothing functions of age.\(^9\)

Initially, we divided participants into \( k \) groups according to the mean value of a risk factor across follow-up period. The group assignment and estimation of the smooth trajectories for each group were achieved by iterating the expectation step (E step) and maximization step (M step).

In M step, for individual \( i \) that was classified into the group \( m \), the model was \( E(Y_i) = f_m(t_i) + b_{mi} \), where the vector \( Y_i \) was the multiple repeated measurements of the risk factor over time \( t_i \) for individual \( i \), \( f_m(.) \) was a non-parametric penalized smoothing function of age, and \( b_{mi} \) was random effect of individual \( i \). Then we obtained mean predicted value \( \hat{Y}_{i(1)}, \hat{Y}_{i(2)}, \ldots, \hat{Y}_{i(k)} \) estimated from the model fitted in the 1st, 2nd, \( \ldots, k \)th groups, respectively.

In the E step, we obtained the log likelihood contributions of individual \( i \)’s trajectory of responses belonging to the 1st, 2nd, \( \ldots, k \)th groups, and reassigned individual \( i \) to the group with the largest log likelihood. The E and M steps were iterated until the model converged, which was determined when the sum of the largest log likelihood for all individuals remained the same. The R script of the SMM is available in Github (https://github.com/mingding-hsph/Smoothing-mixture-model).
**Total, between-person, and within-person sum of squares**

Total sum of squares: We first calculated the difference between each value and the grant mean, which was the mean of all values from all participants. Then we summed the squared differences across all participants.

Between-person sum of squares: For each participant, we first computed the difference between its group mean and the grand mean. Then we summed the squared differences across all participants.

Within-person sum of squares: For each participant, we first computed the difference between each value and its group mean. Then we summed the squared differences across all participants.

Total sum of squares would be equal to the sum of between-person sum of squares and within-person sum of squares.

**Joint membership of trajectories of risk factors and trajectories of change in risk factors**

First, we identified trajectories of risk factors during follow-up and classified participants into three groups (high, medium, low) using SMM. Within each group, we obtained the mean predicted value of risk factor with age (fixed effects) and the predicted value of risk factor with age for each participant (random effects) from the output of SMM. We plotted the mean and 95% CI of predicted trajectories of each group and predicted trajectories of individuals within each group using the ‘ggplot’ command in R 3.5.0. The identification of trajectories of risk factors was conducted in the NHS and the HPFS, separately.

Second, we reclassified these participants using SMM from a different dimension—trajectories of change in risk factors from baseline. The participants were classified into three groups: increase, stable, and decrease. Within each group, we obtained the mean predicted value
of change in risk factor (fixed effects) and the predicted value of change in risk factor for each participant (random effects) from the output of SMM. We plotted the mean and 95% CI of predicted trajectories of each group and predicted trajectories of individuals within each group using the ‘ggplot’ command in R 3.5.0. The identification of trajectories of risk factors was conducted in the NHS and the HPFS, separately.

Finally, based on the group memberships of trajectories of risk factors and trajectories of change in risk factors, we jointly classified participants into nine groups: high–stable, high–increase, high–decrease, medium–stable, medium–increase, medium–decrease, low–stable, low–increase, and low–decrease. After obtaining the joint group membership, we pooled the data of the NHS and the HFPS. Within each joint group, we calculated mean value of risk factor with age and used ‘loess.smooth’ function in R to obtain the trajectories of risk factors with age. The trajectories of risk factors based on joint group memberships were plotted using ‘plot’ command in R 3.5.0.

**Group-based trajectory analysis**

Proposed by Nagin et al, group-based trajectory analysis is a widely used method to classify participants within a population with heterogeneous longitudinal trajectories. It allows different sets of parameter values for mixture components corresponding to different unobserved subgroups of individuals, and captures latent trajectory classes with different growth curves by using an expectation-maximization (EM) algorithm. Group-based trajectory analysis assumes individuals within groups are homogenous and uses polynomials to generate flexible trajectories. We modeled trajectories with cubic polynomials, fitted group-based trajectory model using “proc traj” command using SAS version 9.2 for UNIX (SAS Institute, Cary, NC), and plotted the trajectories using “trajplot” command.
Table S1. Time (year) at assessment of modifiable risk factors in the Nurses’ Health Study (NHS) and the Health Professionals Follow-up Study (HPFS).

|                  | 84 | 86 | 88 | 90 | 92 | 94 | 96 | 00 | 02 | 04 | 06 | 08 | 10 | 12 | 14 | 16 |
|------------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| BMI, smoking     | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |
| NHS              | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |
| HPFS             |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| Diet quality     | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |
| NHS              | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |
| HPFS             |    | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |
| Physical activity| X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |
| NHS              | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |
| HPFS             |    | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |
| Alcohol intake   | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |
| NHS              | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |
| HPFS             |    | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  | X  |
Table S2. Variance components of body mass index (BMI), smoking, alcohol intake, diet quality, and physical activity in the Nurses’ Health Study and the Health Professionals Follow-up Study.

| Risk factors     | NHS                        | HPFS                       |
|------------------|---------------------------|---------------------------|
|                  | Between person sum of squares (SS) | Within person SS | Total SS | Proportion of between person SS over total SS, % | Between person sum of squares (SS) | Within person SS | Total SS | Proportion of between person SS over total SS, % |
| BMI              | 10650346                  | 1533622                   | 12183968 | 87                                                   | 2381276                  | 275831           | 2657107 | 90                                                   |
| Smoking          | 22907934                  | 9006215                   | 31914149 | 72                                                   | 6508684                  | 2263251          | 8771935 | 74                                                   |
| Alcohol intake   | 8218491                   | 2793784                   | 11012275 | 75                                                   | 9266668                  | 2681234          | 11947902 | 78                                                   |
| AHEI             | 15671920                  | 9071161                   | 24743080 | 63                                                   | 8416905                  | 2650293          | 11067198 | 76                                                   |
| Physical activity| 27163279                  | 28920211                  | 56083489 | 48                                                   | 26252022                 | 18104344         | 44356366 | 59                                                   |
Figure S1. A flow diagram of sample selection in the Nurses’ Health Study (NHS).

121,706 participants at baseline

Exclusions (n=70264)
8002 participants who were dead or reported cardiovascular disease, cancer, or type 2 diabetes at baseline.
35,999 participants with extreme and implausible daily energy intakes at baseline
53 participants with missing date of birth
26,210 participants with unknown longevity status (participants were alive and did not reach 85 years old at the end of follow up)

51,442 participants with 29,016 achieving longevity, including 17,448 participants achieving healthy longevity

Censor lifestyle factors reported after development of CVD, type 2 diabetes, or cancer or lifestyle factors reported after age 85

Exclude participants with less than two measurements of lifestyles or change in lifestyles to derive patterns

Patterns of lifestyles:
- 49,293 participants on BMI
- 49,967 participants on smoking
- 49,869 participants on alcohol intake
- 49,869 participants on diet quality
- 49,069 participants on physical activity

Patterns of change in lifestyles:
- 44,866 participants on BMI
- 48,341 participants on smoking
- 46,062 participants on alcohol intake
- 46,062 participants on diet quality
- 44,059 participants on physical activity
Figure S2. A flow diagram of sample selection in the Health Professionals Follow-up Study (HPFS).

51,530 participants at baseline

Exclusions (n=17,626)
- 722 participants who were dead or reported cardiovascular disease, cancer, or type 2 diabetes at baseline.
- 1573 participants with extreme and implausible daily energy intakes at baseline.
- 1 participants with missing date of birth.
- 15,330 participants with unknown longevity status (participants were alive and did not reach 85 years old at the end of follow up).

33,904 participants with 17,026 achieving longevity, including
- 7874 participants achieving healthy longevity

Censor lifestyle factors reported after development of CVD, type 2 diabetes, or cancer or lifestyle factors reported after age 85

Excluding participants with less than two measurements of lifestyles or change in lifestyles to derive patterns

Patterns of lifestyles:
- 30,369 participants on BMI
- 31,228 participants on smoking
- 29,761 participants on alcohol intake
- 29,684 participants on diet quality
- 33,704 participants on physical activity

Patterns of change in lifestyles:
- 26,386 participants on BMI
- 27,365 participants on smoking
- 23,144 participants on alcohol intake
- 23,003 participants on diet quality
- 29,242 participants on physical activity
Table S3. Baseline characteristics of participants by trajectories of modifiable risk factors including body mass index (BMI), smoking, alcohol intake, diet quality, and physical activity in the Nurses’ Health Study (NHS, women) and the Health Professionals Follow-up Study (HPFS, men).

|                  | NHS (1984) |               | HPFS (1986) |               |
|------------------|------------|---------------|-------------|---------------|
|                  | Low (n=23,856) | Medium (n=19,658) | High (n=5,810) | P value | Low (n=39,773) | Medium (n=67,399) | High (n=34,879) | P value |
| Age              | 54.79 (5.18) | 54.08 (5.34) | 52.15 (5.88) | <0.001 | 54.36 (5.29) | 53.11 (5.78) | 54.50 (5.66) | <0.001 |
| BMI (kg/m²)      | 21.96 (1.95) | 26.44 (2.61) | 34.06 (4.73) | <0.001 | 25.43 (4.76) | 24.09 (4.13) | 24.47 (4.38) | <0.001 |
| Alcohol intake (g)| 8.73 (12.71) | 6.706 (11.348) | 4.09 (9.078) | <0.001 | 6.49 (10.68) | 9.814 (13.879) | 13.08 (17.94) | <0.001 |
| Physical activity (MET-h/week) | 15.79 (10.52) | 44.36 (19.63) | 10.15 (10.1) | <0.001 | 14.66 (19.71) | 42.13 (9.71) | 9.91 (20.78) | <0.001 |
| AHEI             | 44.33 (10.52) | 13.06 (10.1) | 43.53 (10.1) | <0.001 | 44.94 (10.36) | 12.29 (19.71) | 40.49 (9.37) | <0.001 |
| Current smoker, %| 29 (10) | 24 (9) | 20 (8) | <0.001 | 8 (5) | 97 (9) | 99 (9) | <0.001 |
| Caucasian, %     | 98 (9) | 98 (9) | 97 (8) | <0.001 | 98 (9) | 98 (9) | 99 (9) | <0.001 |
| Menopausal status, % | 70 (21) | 69 (17) | 68 (11) | <0.001 | 68 (19) | 74 (16) | 75 (14) | <0.001 |
| Postmenopausal hormone use, %§ | 21 (43) | 17 (43) | 11 (42) | 0.78 | 43 (35) | 40 (36) | 39 (37) | 0.43 |
| Family history of cancer, % | 35 (40) | 36 (38) | 37 (35) | 0.004 | 36 (39) | 36 (34) | 37 (34) | <0.001 |
| Family history of cardiovascular disease, % | 20 (40) | 25 (38) | 30 (35) | <0.001 | 23 (39) | 21 (34) | 20 (34) | <0.001 |
| Multivitamin use, % | 40 (10) | 38 (12) | 35 (16) | <0.001 | 39 (15) | 34 (16) | 34 (16) | <0.001 |
| Physical activity (MET-h/week) | Low (n=16,152) | Medium (n=22395) | High (n=11,355) | P value | Low (n=32,110) | Medium (n=14251) | High (n=2739) | P value |
|-------------------------------|----------------|-----------------|-----------------|---------|----------------|-----------------|---------------|---------|
| Age                           | 53.47 (5.64)   | 54.33 (5.28)    | 55.04 (5.06)    | <0.001  | 54.29 (5.43)   | 54.17 (5.24)    | 53.41 (5.41)  | <0.001  |
| BMI (kg/m²)                   | 25.50 (4.64)   | 25.27 (4.64)    | 24.66 (4.32)    | <0.001  | 25.66 (4.92)   | 24.33 (3.94)    | 23.67 (3.83)  | <0.001  |
| Alcohol intake (g)            | 7.87 (13.07)   | 7.447 (11.73)   | 6.57 (10.31)    | <0.001  | 7.24 (12.18)   | 7.601 (11.291)  | 8.04 (11.70)  | <0.001  |
| Physical activity (MET-h/week) | 9.98 (14.23)   | 44.83 (6.67)    | 19.62 (26.38)   | <0.001  | 7.35 (8.62)    | 46.2 (10.38)    | 51.81 (51.80) | <0.001  |
| AHEI                          | 35.19 (6.31)   | 14.08 (20.04)   | 55.96 (8.06)    | <0.001  | 43.03 (10.01)  | 21.72 (18.25)   | 48.82 (10.81) | <0.001  |
| Current smoker, %             | 33 (25)        | 25 (17)         | 17 (9)          | <0.001  | 28 (22)        | 22 (17)         | 21 (8)        | <0.001  |
| Caucasian, %                  | 99 (98)        | 98 (97)         | 97 (98)         | <0.001  | 98 (98)        | 98 (98)         | 98 (98)       | 0.03    |
| Menopausal status, %          | 70 (70)        | 70 (69)         | 69 (70)         | <0.001  | 70 (69)        | 69 (70)         | 70 (70)       | 0.09    |
| Postmenopausal hormone use, % | 16 (19)        | 19 (21)         | 21 (18)         | <0.001  | 18 (19)        | 19 (19)         | 19 (19)       | <0.001  |
| Family history of cancer, %   | 43 (43)        | 43 (43)         | 43 (43)         | 0.24    | 43 (43)        | 44 (43)         | 43 (43)       | 0.02    |
| Family history of cardiovascular disease, % | 36 | 36 | 36 | 0.26 | 36 | 35 | 36 | 0.52 |
|--------------------------------------------|--|--|--|--|--|--|--|--|
| Family history of diabetes, %              | 24 | 23 | 23 | 0.23 | 23 | 23 | 21 | 0.12 |
| Multivitamin use, %                        | 34 | 40 | 45 | <0.001 | 37 | 42 | 44 | <0.001 |

| HPFS (1986)          | Low (n=8978) | Medium (n=13082) | High (n=7628) | Low (n=22,244) | Medium (n=9625) | High (n=1839) |
|----------------------|--------------|------------------|---------------|----------------|-----------------|---------------|
| Age                  | 58.20 (7.88) | 59.47 (7.69)     | 60.61 (7.62)  | 60.01 (8.05)   | 58.75 (7.76)    | 57.84 (7.63)  |
| BMI (kg/m²)          | 25.41 (5.17) | 25.06 (4.90)     | 24.46 (4.90)  | 25.23 (5.54)   | 24.62 (4.7)     | 24.38 (4.37)  |
| Alcohol intake (g)   | 14.23 (18.55)| 12.07 (15.64)    | 9.60 (13.42)  | 11.83 (16.55)  | 12.47 (10.75)   | 12.35 (15.83) |
| Physical activity (MET-h/week) | 15.5 (22.85) | 48.31 (29.67)    | 25.59 (29.67) | 9.60 (10.20)   | 49.49 (10.75)   | 73.03 (65.82) |
| AHEI                 | 36.84 (6.41) | 19.54 (26.73)    | 59.96 (7.43)  | 46.77 (10.83)  | 32.45 (22.39)   | 50.68 (11.22) |
| Current smoker, %    | 16 8        | 9 4              | 4 7           | 13 8           | 7 9              | 7 9           |
| Caucasian, %         | 92 91       | 91 91            | 91 91         | 91 91          | 91 91            | 91 91         |
| Family history of cancer, % | 9 9 | 10 0.01 | 9 9 | 9 1.00 |
| Family history of cardiovascular disease, % | 33 35 | 39 <0.001 | 36 34 | 35 0.04 |
| Family history of diabetes, %              | 15 16 | 16 0.01 | 15 15 | 15 0.56 |
| Multivitamin use, %                           | 32 35 | 40 <0.001 | 32 35 | 36 <0.001 |
## Alcohol intake

|                | NHS (1984) | HPFS (1986) |
|----------------|------------|-------------|
|                | Low (n=36,728) | Medium (n=10025) | High (n=3149) | P value |
| Age            | 54.23 (5.41)  | 54.2 (5.29)  | 54.02 (5.32) | <0.001 |
| BMI (kg/m²)    | 25.69 (4.96)  | 23.91 (3.54) | 23.81 (3.70) | <0.001 |
| Alcohol intake (g) | 2.46 (3.84)  | 16.27 (10.82) | 36.39 (17.98) | <0.001 |
| Physical activity (MET-h/week) | 13.28 (19.50) | 44.63 (10.01) | 14.42 (19.91) | <0.001 |
| AHEI           | 44.27 (10.45) | 16.52 (23.68) | 42.76 (9.74) | <0.001 |
| Current smoker, % | 23 (10) | 32 (10) | 45 (10) | <0.001 |
| Caucasian, %   | 97 (10) | 99 (10) | 99 (10) | <0.001 |
| Menopausal status, % | 69 (10) | 70 (10) | 71 (10) | 0.15 |
| Postmenopausal hormone use, % | 18 (10) | 20 (10) | 19 (10) | <0.001 |
| Family history of cancer, % | 43 (10) | 43 (10) | 42 (10) | 0.28 |
| Family history of cardiovascular disease, % | 36 (10) | 36 (10) | 36 (10) | 0.60 |
| Family history of diabetes, % | 25 (10) | 20 (10) | 17 (10) | <0.001 |
| Multivitamin use, % | 39 (10) | 40 (10) | 38 (10) | 0.72 |

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s).
|                                | Group 1 | Group 2 | Group 3 | p-value |
|--------------------------------|---------|---------|---------|---------|
| Physical activity (MET-h/week) | 48.31   | 47.98   | 44.46   | <0.001  |
|                                | (11.05) | (10.56) | (10.37) |         |
| AHEI                           | 18.94   | 21.83   | 19.48   | <0.001  |
|                                | (26.18) | (27.33) | (27.07) |         |
| Current smoker, %              | 8       | 11      | 21      | <0.001  |
| Caucasian, %                   | 91      | 92      | 93      | <0.001  |
| Family history of cancer, %    | 10      | 9       | 10      | 0.54    |
| Family history of cardiovascular disease, % | 35 | 35 | 37 | 0.30 |
| Family history of diabetes, %  | 17      | 14      | 15      | <0.001  |
| Multivitamin use, %            | 35      | 35      | 37      | 0.21    |

MET-h/week: metabolic equivalent hours per week; AHEI: Alternate healthy eating index
For continuous variables, data are mean values and standard deviations.
§Percentage of current postmenopausal hormone use among total women.

P values were obtained using linear regression for continuous variables and logistic regression for categorical variables.
Figure S3. Plots of patterns of modifiable risk factors with 95% confidence interval in the Nurses’ Health Study.
Solid line: the mean of predicted trajectories of each group; Dash line: 95% confidence interval (CI) of predicted trajectories of each group.

For AHEI-2010, physical activity, and alcohol intake, we classified trajectories using standardized values and mapped the standardized values to real values using the inverse cumulative distribution function when plotting the trajectories.

We identified patterns of risk factors using smoothing mixture models (SMM). We obtained the mean and 95% CI of predicted trajectories of each group and predicted trajectories of each participant within each group.
Figure S4. Plots of patterns of modifiable risk factors with 95% confidence interval in the Health Professionals Follow-up Study.
Solid line: the mean of predicted trajectories of each group; Dash line: 95% confidence interval (CI) of predicted trajectories of each group.

For AHEI-2010, physical activity, and alcohol intake, we classified trajectories using standardized values and mapped the standardized values to real values using the inverse cumulative distribution function when plotting the trajectories.

We identified patterns of risk factors using smoothing mixture models (SMM). We obtained the mean and 95% CI of predicted trajectories of each group and predicted trajectories of each participant within each group.
Figure S5. Plots of patterns of change in modifiable risk factors from baseline with 95% confidence interval in the NHS.
Solid line: the mean of predicted trajectories of each group; Dash line: 95% confidence interval (CI) of predicted trajectories of each group.

For change in AHEI-2010, physical activity, and alcohol intake, we plotted the trajectories in the scale of standardized values.

We identified patterns of risk factors using smoothing mixture models (SMM). We obtained the mean and 95% CI of predicted trajectories of each group and predicted trajectories of each participant within each group.
Figure S6. Plots of patterns of change in modifiable risk factors from baseline with 95% confidence interval in the Health Professionals Follow-up Study.
Solid line: the mean of predicted trajectories of each group; Dash line: 95% confidence interval (CI) of predicted trajectories of each group.

For change in AHEI-2010, physical activity, and alcohol intake, we plotted the trajectories in the scale of standardized values. We identified patterns of risk factors using smoothing mixture models (SMM). We obtained the mean and 95% CI of predicted trajectories of each group and predicted trajectories of each participant within each group.
Figure S7. Plots of joint patterns of modifiable risk factors and change in the factor from baseline in the Nurses’ Health Study (women).
We identified patterns of risk factor and patterns of change in the risk factor using smoothing mixture models, classified participants according to joint group membership, and plotted the mean values of the risk factor with age within each category using loess.smooth function in R.
Figure S8. Plots of joint patterns of modifiable risk factors and change in the factor from baseline in the Health Professionals Follow-up Study (men).

Body mass index (BMI)

Smoking

 doi: 10.1136/bmjmed-2021-000098
We identified patterns of risk factor and patterns of change in the risk factor using smoothing mixture models, classified participants according to joint group membership, and plotted the mean values of the risk factor with age within each category using loess.smooth function in R.
Table S4. Associations of trajectories of modifiable risk factors and trajectories of change in the factors with odds ratios (OR) of achieving longevity in the Nurses’ Health Study (women) and the Health Professionals Follow-up Study (men).

| Risk factors | Medium | Low | High | Change in risk factors | No change | Increase | Decrease |
|--------------|--------|-----|------|------------------------|-----------|----------|----------|
| BMI          |        |     |      |                        |           |          |          |
| **NHS**      |        |     |      |                        |           |          |          |
| Cases/Participants | 11739/19645 | 14106/23841 | 2544/5807 | Cases/Participants | 15870/23961 | 2972/6555 | 7493/14350 |
| Model 1      | 1.00   | 0.98 (0.94, 1.01) | 0.53 (0.49, 0.56) | Model 3      | 1.00      | 0.47 (0.44, 0.50) | 0.55 (0.53, 0.58) |
| Model 2      | 1.00   | 0.96 (0.92, 1.01) | 0.58 (0.54, 0.62) | Model 4      | 1.00      | 0.77 (0.71, 0.82) | 0.55 (0.52, 0.58) |
| **HPFS**     |        |     |      |                        |           |          |          |
| Cases/Participants | 6916/13212 | 7999/14210 | 1047/2947 | Cases/Participants | 9225/14734 | 1099/3254 | 4122/8398 |
| Model 1      | 1.00   | 1.17 (1.12, 1.23) | 0.50 (0.46, 0.54) | Model 3      | 1.00      | 0.35 (0.32, 0.38) | 0.57 (0.54, 0.60) |
| Model 2      | 1.00   | 1.04 (0.98, 1.09) | 0.62 (0.57, 0.68) | Model 4      | 1.00      | 0.63 (0.57, 0.69) | 0.59 (0.56, 0.63) |
| Smoking      |        |     |      |                        |           |          |          |
| **NHS**      |        |     |      |                        |           |          |          |
| Cases/Participants | 2497/6731 | 25206/39752 | 889/3484 | Cases/Participants | 25781/41782 | 322/932 | 2004/5627 |
| Model 1      | 1.00   | 2.94 (2.79, 3.10) | 0.58 (0.53, 0.64) | Model 3      | 1.00      | 0.55 (0.48, 0.64) | 1.28 (1.16, 1.42) |
| Model 2      | 1.00   | 2.80 (2.63, 2.98) | 0.54 (0.48, 0.60) | Model 4      | 1.00      | 0.72 (0.60, 0.85) | 1.24 (1.10, 1.39) |
| **HPFS**     |        |     |      |                        |           |          |          |
| Cases/Participants | 608/2086 | 15431/28238 | 206/904 | Cases/Participants | 14144/25505 | 205/715 | 332/1145 |
| Model 1      | 1.00   | 2.93 (2.66, 3.23) | 0.72 (0.60, 0.86) | Model 3      | 1.00      | 0.36 (0.29, 0.44) | 1.59 (1.28, 1.98) |
| Model 2      | 1.00   | 2.68 (2.41, 2.97) | 0.61 (0.51, 0.75) | Model 4      | 1.00      | 0.49 (0.39, 0.60) | 1.50 (1.19, 1.89) |
| AHEI         |        |     |      |                        |           |          |          |
| **NHS**      |        |     |      |                        |           |          |          |
| Cases/Participants | 13237/22388 | 8102/16135 | 7316/11346 | Cases/Participants | 13321/21554 | 7696/12937 | 6536/11571 |
| Model 1      | 1.00   | 0.70 (0.67, 0.73) | 1.25 (1.20, 1.32) | Model 3      | 1.00      | 1.02 (0.97, 1.08) | 0.67 (0.64, 0.71) |
| Model 2      | 1.00   | 0.83 (0.79, 0.87) | 1.06 (1.01, 1.12) | Model 4      | 1.00      | 1.07 (1.00, 1.13) | 0.82 (0.77, 0.87) |
| **HPFS**     |        |     |      |                        |           |          |          |
| Cases/Participants | 7202/13082 | 4293/8976 | 4492/7626 | Cases/Participants | 6697/10899 | 3415/6146 | 3418/5958 |
| Model 1      | 1.00   | 0.75 (0.71, 0.79) | 1.17 (1.11, 1.24) | Model 3      | 1.00      | 0.98 (0.91, 1.05) | 0.64 (0.59, 0.69) |
| Model 2      | 1.00   | 0.88 (0.83, 0.93) | 1.02 (0.96, 1.09) | Model 4      | 1.00      | 0.97 (0.89, 1.05) | 0.78 (0.71, 0.84) |
|                      | Physical activity |                      | Physical activity |                      |
|----------------------|-------------------|----------------------|-------------------|----------------------|
| **NHS**              | Cases/Participants | 9126/14243          | 17522/32088       | 1671/2738            |
|                      | Model 1           | 1.00                 | 0.67 (0.65, 0.70) | 0.88 (0.81, 0.96)   |
|                      | Model 2           | 1.00                 | 0.75 (0.72, 0.79) | 0.90 (0.82, 0.99)   |
| **HPFS**             | Cases/Participants | 5474/9624           | 10549/22241       | 958/1839             |
|                      | Model 1           | 1.00                 | 0.68 (0.65, 0.72) | 0.82 (0.75, 0.91)   |
|                      | Model 2           | 1.00                 | 0.65 (0.62, 0.69) | 0.88 (0.79, 0.98)   |
| **NHS**              | Cases/Participants | 5862/10019          | 21292/36703       | 1501/3147            |
|                      | Model 1           | 1.00                 | 0.98 (0.94, 1.02) | 0.65 (0.60, 0.70)   |
|                      | Model 2           | 1.00                 | 0.99 (0.93, 1.04) | 0.78 (0.71, 0.85)   |
| **HPFS**             | Cases/Participants | 4556/8049           | 10110/18783       | 1371/2929            |
|                      | Model 1           | 1.00                 | 0.89 (0.85, 0.94) | 0.67 (0.62, 0.73)   |
|                      | Model 2           | 1.00                 | 0.84 (0.79, 0.89) | 0.74 (0.68, 0.81)   |

Model 1 is univariate analysis.

Model 2 adjusted for baseline age (continuous), race (White, Black, Asian, and other), family histories of cancer (yes, no), myocardial infarction (yes, no), and type 2 diabetes (yes, no), multivitamin use (yes, no), menopausal status (yes, no, women only), postmenopausal hormone use (yes, no, women only), cohort, education (registered nurse, bachelor degree, master degree and higher, women only), social economic status (annual family income [quartiles] for women and work status [disabled, retired, part-time, full-time] for men), aspirin use (yes, no), use of antihypertensive medications (yes, no), and use of cholesterol lowering medications (yes, no), and the other four risk factors at baseline as continuous variables.

Model 3 adjusted for all variables included in model 1, as well as group membership of trajectories of the risk factor (categorical).

Model 4 adjusted for all variables included in model 2, as well as group membership of trajectories of the risk factor (categorical).
Table S5. Assignment of score (1-9) to each risk factor based on the joint patterns.

| Joint patterns  | BMI | Smoking | AHEI | Physical activity | Alcohol intake |
|-----------------|-----|---------|------|-------------------|----------------|
| Low, decrease   | 3   | 2       | 9    | 9                 | 6              |
| Low, stable     | 1   | 1       | 8    | 8                 | 4              |
| Low, increase   | 5   | 3       | 7    | 7                 | 3              |
| Medium, decrease| 4   | 4       | 6    | 5                 | 5              |
| Medium, stable  | 2   | 5       | 5    | 3                 | 1              |
| Medium, increase| 6   | 6       | 4    | 2                 | 2              |
| High, decrease  | 8   | 7       | 3    | 6                 | 8              |
| High, stable    | 7   | 8       | 2    | 4                 | 7              |
| High, increase  | 9   | 9       | 1    | 1                 | 9              |
Figure S9. Plots of patterns of modifiable risk factors with 95% confidence interval in the Nurses’ Health Study using group-based trajectory analysis.
Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s).

Ding M, et al. BMJMED 2022; 1:e000098. doi: 10.1136/bmjmed-2021-000098

### Alcohol Intake Over Time (Age)

| Time (Age) | Alcohol Intake (grams/day) |
|------------|----------------------------|
| 40         | 20                         |
| 50         | 15                         |
| 60         | 10                         |
| 70         | 5                          |
| 80         | 0                          |
| 90         | 0                          |

- **Curve 1**: 18.6% of participants
- **Curve 2**: 70.3% of participants
- **Curve 3**: 11.1% of participants
Figure S10. Plots of patterns of modifiable risk factors with 95% confidence interval in the Health Professionals Follow-up Study using group-based trajectory analysis.
Figure S11. Plots of patterns of change in modifiable risk factors from baseline with 95% confidence interval in the Nurses’ Health Study using group-based trajectory analysis.

Note: We did not obtain valid model output for change in alcohol intake and physical activity due to the ERROR “Floating Point Zero Divide”.

40
Figure S12. Plots of patterns of change in modifiable risk factors from baseline with 95% confidence interval in the Health Professionals Follow-up Study using group-based trajectory analysis.
Change in alcohol intake over time (age) represented in the diagram. The x-axis represents time or age, ranging from 40 to 90, and the y-axis represents the change in alcohol intake. Three different lines indicate different percentages: 1. 4.1%, 2. 91.6%, and 3. 4.3%.
Table S6. Comparison of classification of trajectories of risk factors and change in risk factors using smoothing mixture model and group-based trajectory model by pooling the Nurses’ Health Study and the Health Professionals Follow-up Study.

| Classification of trajectories of risk factors | Classification of trajectories of change in risk factors |
|----------------------------------------------|-------------------------------------------------------|
| | Group-based trajectory model | Spearman correlation | | Group-based trajectory model | Spearman correlation |
| **BMI** | | | | | |
| Low | Medium | High | Low | Medium | High | | Decrease | Stable | Increase | |
| 22636 | 704 | 137 | 21580 | 10371 | 170 |
| Smoking | | | | | |
| Low | Medium | High | Low | Medium | High | | Decrease | Stable | Increase | |
| 63550 | 10 | 3 | 63562 | 6127 | 8 |
| AHEI-2010 | | | | | |
| Low | Medium | High | Low | Medium | High | | Decrease | Stable | Increase | |
| 22231 | 488 | 2 | 17578 | 16946 | 152 |
| Physical activity | | | | | |
| Low | Medium | High | Low | Medium | High | | Decrease | Stable | Increase | |
| 27647 | 7072 | 461 | 2039 | 11870 | 0 |
| Alcohol intake | | | | | |
| Low | Medium | High | Low | Medium | High | | Decrease | Stable | Increase | |
| 49826 | 564 | 1373 | 7 | 2131 | 1971 |
| BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s).
Table S7. Stratified analysis by age at baseline on associations of joint patterns of modifiable risk factors and change in the factor from baseline with odds ratios (OR) of achieving longevity by pooling the Nurses’ Health Study and the Health Professionals Follow-up Study.

|                          | Age at baseline<65 years | Age at baseline≥65 years | P for interaction* |
|--------------------------|--------------------------|--------------------------|--------------------|
|                          | Probability of achieving longevity* | Odds ratio of achieving longevity# | Probability of achieving longevity* | Odds ratio of achieving longevity# |                   |
| **BMI**                  |                          |                          |                    |                                |                |
| High, decrease           | 0.40 (0.37, 0.43)        | 0.37 (0.33, 0.42)        | 0.42 (0.32, 0.53)  | 0.28 (0.18, 0.45)              |                |
| High, increase           | 0.53 (0.50, 0.55)        | 0.62 (0.55, 0.70)        | 0.78 (0.59, 0.90)  | 1.37 (0.57, 3.30)              |                |
| High, stable             | 0.52 (0.49, 0.54)        | 0.60 (0.54, 0.67)        | 0.60 (0.52, 0.67)  | 0.57 (0.41, 0.80)              |                |
| Low, decrease            | 0.51 (0.50, 0.52)        | 0.58 (0.55, 0.62)        | 0.61 (0.58, 0.65)  | 0.61 (0.53, 0.72)              |                |
| Low, increase            | 0.58 (0.53, 0.62)        | 0.77 (0.64, 0.92)        | 0.72 (0.42, 0.90)  | 0.99 (0.28, 3.56)              |                |
| Low, stable              | 0.67 (0.65, 0.68)        | 1.11 (1.05, 1.18)        | 0.71 (0.68, 0.74)  | 0.95 (0.82, 1.10)              |                |
| Medium, decrease         | 0.51 (0.49, 0.52)        | 0.58 (0.54, 0.62)        | 0.58 (0.54, 0.62)  | 0.53 (0.44, 0.64)              |                |
| Medium, increase         | 0.64 (0.62, 0.66)        | 0.98 (0.90, 1.08)        | 0.64 (0.51, 0.75)  | 0.67 (0.39, 1.16)              |                |
| Medium, stable           | 0.64 (0.63, 0.65)        | Ref                      | 0.72 (0.69, 0.75)  | Ref                             | <0.001          |
| **Smoking**              |                          |                          |                    |                                |                |
| High, decrease           | 0.24 (0.21, 0.27)        | 0.59 (0.49, 0.70)        | 0.35 (0.23, 0.48)  | 0.44 (0.22, 0.86)              |                |
| High, increase           | 0.26 (0.21, 0.31)        | 0.64 (0.49, 0.84)        | 0.32 (0.17, 0.52)  | 0.39 (0.15, 0.97)              |                |
| High, stable             | 0.25 (0.23, 0.28)        | 0.63 (0.54, 0.72)        | 0.40 (0.32, 0.49)  | 0.55 (0.34, 0.92)              |                |
| Low, decrease            | 0.49 (0.45, 0.53)        | 1.75 (1.46, 2.10)        | 0.50 (0.28, 0.73)  | NA                             |                |
| Low, increase            | 0.38 (0.31, 0.46)        | 1.15 (0.83, 1.61)        | 0.68 (0.66, 0.71)  | 0.83 (0.30, 2.33)              |                |
| Low, stable              | 0.64 (0.63, 0.65)        | 3.25 (2.96, 3.57)        | 0.48 (0.39, 0.57)  | 1.78 (1.23, 2.57)              |                |
| Medium, decrease         | 0.38 (0.36, 0.40)        | 1.12 (0.99, 1.26)        | 0.52 (0.42, 0.62)  | 0.76 (0.46, 1.27)              |                |
| Medium, increase         | 0.39 (0.35, 0.44)        | 1.21 (0.98, 1.49)        | 0.48 (0.42, 0.53)  | 0.91 (0.53, 1.55)              |                |
| Medium, stable           | 0.35 (0.33, 0.37)        | Ref                      | 0.55 (0.45, 0.64)  | Ref                             | <0.001          |
| **AHEI**                 |                          |                          |                    |                                |                |
| High, decrease           | 0.49 (0.35, 0.63)        | 0.88 (0.81, 0.96)        | 0.75 (0.70, 0.79)  | 0.87 (0.67, 1.13)              |                |
| High, increase           | 0.48 (0.34, 0.62)        | 1.13 (1.03, 1.23)        | 0.74 (0.69, 0.79)  | 0.85 (0.66, 1.09)              |                |
| High, stable             | 0.47 (0.33, 0.61)        | 1.04 (0.96, 1.12)        | 0.78 (0.75, 0.82)  | 1.06 (0.85, 1.33)              |                |
| Low, decrease            | 0.49 (0.35, 0.63)        | 0.72 (0.66, 0.78)        | 0.69 (0.62, 0.74)  | 0.65 (0.49, 0.86)              |                |
| Low, increase | 0.45 (0.32, 0.59) | 0.80 (0.73, 0.87) | 0.73 (0.67, 0.78) | 0.80 (0.60, 1.06) |
|---------------|-------------------|-------------------|-------------------|-------------------|
| Low, stable   | 0.43 (0.30, 0.57) | 0.84 (0.79, 0.90) | 0.74 (0.70, 0.78) | 0.84 (0.68, 1.04) |
| Medium, decrease | 0.54 (0.40, 0.68) | 0.88 (0.82, 0.95) | 0.70 (0.66, 0.75) | 0.70 (0.56, 0.88) |
| Medium, increase | 0.50 (0.37, 0.64) | 0.97 (0.91, 1.04) | 0.73 (0.69, 0.77) | 0.80 (0.64, 1.01) |
| Medium, stable | 0.51 (0.37, 0.65) | Ref               | 0.77 (0.74, 0.80) | Ref               |

**Physical activity**

| High, decrease | 0.45 (0.31, 0.58) | 0.74 (0.61, 0.90) | 0.74 (0.61, 0.84) | 1.17 (0.62, 2.18) |
|----------------|-------------------|-------------------|-------------------|-------------------|
| High, increase | 0.43 (0.31, 0.57) | 1.08 (0.93, 1.25) | 0.72 (0.65, 0.79) | 1.06 (0.73, 1.54) |
| High, stable   | 0.38 (0.26, 0.51) | 0.92 (0.81, 1.06) | 0.77 (0.65, 0.86) | 1.38 (0.75, 2.56) |
| Low, decrease  | 0.49 (0.35, 0.63) | 0.66 (0.52, 0.82) | 0.56 (0.47, 0.64) | 0.51 (0.35, 0.73) |
| Low, increase  | 0.41 (0.29, 0.55) | 0.75 (0.70, 0.80) | 0.59 (0.50, 0.68) | 0.59 (0.40, 0.86) |
| Low, stable    | 0.37 (0.25, 0.51) | 0.75 (0.70, 0.81) | 0.61 (0.59, 0.64) | 0.63 (0.55, 0.73) |
| Medium, decrease | 0.50 (0.36, 0.64) | 0.84 (0.71, 0.99) | 0.64 (0.57, 0.71) | 0.73 (0.53, 1.00) |
| Medium, increase | 0.46 (0.33, 0.59) | 1.08 (1.00, 1.16) | 0.75 (0.70, 0.78) | 1.19 (0.94, 1.50) |
| Medium, stable | 0.46 (0.33, 0.60) | Ref               | 0.71 (0.68, 0.75) | Ref               |

**Alcohol intake**

| High, decrease | 0.53 (0.38, 0.66) | 0.62 (0.52, 0.74) | 0.70 (0.61, 0.78) | 0.71 (0.46, 1.10) |
|----------------|-------------------|-------------------|-------------------|-------------------|
| High, increase | 0.51 (0.37, 0.65) | 0.80 (0.70, 0.91) | 0.75 (0.66, 0.82) | 0.88 (0.56, 1.38) |
| High, stable   | 0.45 (0.31, 0.59) | 0.72 (0.64, 0.81) | 0.76 (0.70, 0.81) | 0.93 (0.68, 1.27) |
| Low, decrease  | 0.59 (0.44, 0.72) | 0.65 (0.52, 0.81) | 0.75 (0.50, 0.90) | 0.88 (0.30, 2.60) |
| Low, increase  | 0.53 (0.39, 0.67) | 1.18 (0.95, 1.45) | 0.75 (0.63, 0.85) | 0.92 (0.50, 1.67) |
| Low, stable    | 0.51 (0.37, 0.65) | 0.95 (0.89, 1.00) | 0.74 (0.71, 0.77) | 0.86 (0.72, 1.02) |
| Medium, decrease | 0.56 (0.41, 0.70) | 0.77 (0.68, 0.87) | 0.71 (0.64, 0.77) | 0.74 (0.52, 1.04) |
| Medium, increase | 0.48 (0.34, 0.62) | 1.02 (0.92, 1.14) | 0.73 (0.67, 0.79) | 0.81 (0.58, 1.13) |
| Medium, stable | 0.47 (0.33, 0.62) | Ref               | 0.77 (0.73, 0.80) | Ref               |

*Logistic model adjusted for baseline age (continuous), race (White, Black, Asian, and other), family histories of cancer (yes, no), myocardial infarction (yes, no), and type 2 diabetes (yes, no), multivitamin use (yes, no), menopausal status (yes, no, women only), postmenopausal hormone use (yes, no, women only), cohort, education (registered nurse, bachelor degree, master degree and higher, women only), social economic status (annual family income [quartiles] for women and work status [disabled, retired, part-time, full-time] for men), aspirin use (yes, no), use of antihypertensive medications (yes, no), use of cholesterol lowering medications (yes, no), and the other four risk factors at baseline as continuous variables.
#Predicted odds of longevity was obtained using ‘estimate’ statement under ‘proc logistic’ command in SAS, assuming mean values of continuous covariates and median values of categorical covariates.

&We included interaction terms of age (<65, ≥65 years) and joint patterns of lifestyle factors (9 categories) into the model (18 categories of interaction terms). P for interaction was obtained using likelihood ratio test comparing models with and without interaction terms.

Table S8. Percentage (%) of missing of risk factors at each assessment in the Nurses’ Health Study (NHS) and the Health Professionals Follow-up Study (HPFS).

| Time at assessment of risk factors | 84 | 86 | 88 | 90 | 92 | 94 | 96 | 98 | 00 | 02 | 04 | 06 | 08 | 10 | 12 | 14 | 16 |
|-----------------------------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| **BMI**                           |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| **NHS**                           | 6  | 10 | 14 | 3  | 6  | 1  | 2  | 2  | 3  | 7  | 6  | 8  | 8  | 9  | 11 | 12 | 14 |
| **HPFS**                          | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| **Smoking**                       |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| **NHS**                           | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| **HPFS**                          | 0  | 4  | 2  | 1  | 2  | 2  | 2  | 3  | 2  | 3  | 4  | 6  | 8  | 3  | 3  |    |    |
| **Diet quality**                  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| **NHS**                           | 0  | 14 | 12 | 9  | 10 | 10 | 9  | 11 |    |    |    |    |    |    |    |    |    |
| **HPFS**                          | 0  | 23 | 20 | 19 | 17 | 18 |    |    |    |    |    |    |    |    |    |    |
| **Physical activity**             |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| **NHS**                           | 9  | 4  | 2  | 1  | 1  | 1  | 0  | 0  | 0  | 0  | 10 | 9  |    |    |    |    |
| **HPFS**                          | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |    |    |    |    |    |    |    |
| **Alcohol intake**                |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| **NHS**                           | 0  | 14 | 12 | 9  | 10 | 10 | 9  | 11 |    |    |    |    |    |    |    |
| **HPFS**                          | 0  | 23 | 20 | 19 | 17 | 18 | 24 |    |    |    |    |    |    |    |    |    |    |
Table S9. Total number of assessments of risk factors during follow-up period in the Nurses’ Health Study (NHS) (n=51442).

| Number of assessments | BMI | Percentage of participants in original data (%) | Number of assessments | Smoking | Percentage of participants after censoring participants (%) | Number of assessments | Physical activity | Percentage of participants in original data (%) | Number of assessments | Alcohol intake | Percentage of participants after censoring participants (%) | Number of assessments | Diet quality | Percentage of participants in original data (%) | Number of assessments | Diet quality | Percentage of participants after censoring participants (%) |
|----------------------|-----|-----------------------------------------------|----------------------|---------|-------------------------------------------------------------|----------------------|------------------|-----------------------------------------------|----------------------|--------------|-------------------------------------------------------------|----------------------|-------------|-------------------------------------------------------------|----------------------|-------------|-------------------------------------------------------------|
| 0                    | 0   | 0                                             | 0                    | 0       | 0                                                           | 0                    | 2                             | 0                                                           | 0                    | 1            | 3                                                           | 0                    | 1           | 3                                                           | 0                    | 1           | 3                                                           |
| 1                    | 1   | 1                                             | 1                    | 1       | 3                                                           | 1                    | 1                             | 4                                                           | 2                    | 5            | 9                                                           | 2                    | 5           | 9                                                           |
| 2                    | 1   | 1                                             | 2                    | 1       | 3                                                           | 2                    | 3                             | 7                                                           | 3                    | 6            | 9                                                           | 3                    | 6           | 9                                                           |
| 3                    | 2   | 4                                             | 3                    | 1       | 4                                                           | 3                    | 2                             | 4                                                           | 4                    | 7            | 11                                                          | 4                    | 7           | 11                                                          |
| 4                    | 2   | 4                                             | 4                    | 2       | 4                                                           | 4                    | 2                             | 5                                                           | 5                    | 10           | 12                                                          | 5                    | 10          | 12                                                          |
| 5                    | 2   | 4                                             | 5                    | 2       | 4                                                           | 5                    | 3                             | 6                                                           | 6                    | 13           | 13                                                          | 6                    | 13          | 13                                                          |
| 6                    | 3   | 5                                             | 6                    | 2       | 5                                                           | 6                    | 5                             | 7                                                           | 7                    | 19           | 17                                                          | 7                    | 19          | 17                                                          |
| 7                    | 3   | 5                                             | 7                    | 3       | 5                                                           | 7                    | 8                             | 11                                                          | 8                    | 38           | 22                                                          | 8                    | 38          | 22                                                          |
| 8                    | 4   | 6                                             | 8                    | 3       | 6                                                           | 8                    | 12                            | 14                                                          |                      |               |                                                             |                      |             |                                                             |
| 9                    | 5   | 6                                             | 9                    | 4       | 6                                                           | 9                    | 16                            | 16                                                          |                      |               |                                                             |                      |             |                                                             |
| 10                   | 6   | 7                                             | 10                   | 4       | 6                                                           | 10                   | 11                            | 8                                                           |                      |               |                                                             |                      |             |                                                             |
| 11                   | 7   | 8                                             | 11                   | 6       | 6                                                           | 11                   | 34                            | 14                                                          |                      |               |                                                             |                      |             |                                                             |
| 12                   | 8   | 9                                             | 12                   | 6       | 8                                                           |                      |                                 |                                                             |                      |               |                                                             |                      |             |                                                             |
| 13                   | 9   | 9                                             | 13                   | 7       | 9                                                           |                      |                                 |                                                             |                      |               |                                                             |                      |             |                                                             |
| 14                   | 10  | 8                                             | 14                   | 8       | 8                                                           |                      |                                 |                                                             |                      |               |                                                             |                      |             |                                                             |
| 15                   | 12  | 7                                             | 15                   | 9       | 7                                                           |                      |                                 |                                                             |                      |               |                                                             |                      |             |                                                             |
| 16                   | 13  | 6                                             | 16                   | 10      | 6                                                           |                      |                                 |                                                             |                      |               |                                                             |                      |             |                                                             |
| 17                   | 13  | 5                                             | 17                   | 30      | 10                                                          |                      |                                 |                                                             |                      |               |                                                             |                      |             |                                                             |

*We censored risk factors reported after diagnosis of CVD, type 2 diabetes, or cancer, and risk factors reported after age 85.*
Table S10. Total number of assessments of risk factors during follow-up period in the Health Professionals Follow-up Study (HPFS) (n=33904).

| BMI | Number of assessments | Percent of participants in original data (%) | Percent of participants after censoring participants (%) | Number of assessments | Percent of participants in original data (%) | Percent of participants after censoring participants (%) | Number of assessments | Percent of participants in original data (%) | Percent of participants after censoring participants (%) | Number of assessments | Percent of participants in original data (%) | Percent of participants after censoring participants (%) |
|-----|-----------------------|---------------------------------------------|---------------------------------------------------------|-----------------------|---------------------------------------------|---------------------------------------------------------|-----------------------|---------------------------------------------|---------------------------------------------------------|-----------------------|---------------------------------------------|---------------------------------------------------------|
|     |                       |                                             |                                                         |                       |                                             |                                                         |                       |                                             |                                                         |                       |                                             |                                                         |
| 1   | 2                     | 7                                           | 1                                                       | 2                     | 7                                           | 1                                                       | 6                     | 15                                         | 1                                                       | 2                     | 12                                         | 23                                                       |
| 2   | 3                     | 8                                           | 2                                                       | 3                     | 8                                           | 2                                                       | 8                     | 17                                         | 2                                                       | 11                    | 19                                         | 12                                                       |
| 3   | 4                     | 9                                           | 3                                                       | 4                     | 9                                           | 3                                                       | 9                     | 16                                         | 3                                                       | 11                    | 16                                         | 2                                                         |
| 4   | 4                     | 8                                           | 4                                                       | 5                     | 8                                           | 4                                                       | 11                    | 15                                         | 4                                                       | 12                    | 14                                         | 3                                                         |
| 5   | 4                     | 8                                           | 5                                                       | 4                     | 8                                           | 5                                                       | 12                    | 12                                         | 5                                                       | 14                    | 11                                         | 4                                                         |
| 6   | 5                     | 8                                           | 6                                                       | 5                     | 8                                           | 6                                                       | 10                    | 7                                          | 6                                                       | 15                    | 9                                          | 5                                                         |
| 7   | 5                     | 8                                           | 7                                                       | 5                     | 8                                           | 7                                                       | 45                    | 17                                         | 7                                                       | 24                    | 9                                          | 6                                                         |
| 8   | 6                     | 7                                           | 8                                                       | 6                     | 7                                           |                                                         |                       |                                             |                                                         |                       |                                             |                                                         |
| 9   | 6                     | 7                                           | 9                                                       | 6                     | 6                                           |                                                         |                       |                                             |                                                         |                       |                                             |                                                         |
| 10  | 6                     | 6                                           | 10                                                      | 6                     | 6                                           |                                                         |                       |                                             |                                                         |                       |                                             |                                                         |
| 11  | 6                     | 5                                           | 11                                                      | 7                     | 5                                           |                                                         |                       |                                             |                                                         |                       |                                             |                                                         |
| 12  | 7                     | 4                                           | 12                                                      | 7                     | 4                                           |                                                         |                       |                                             |                                                         |                       |                                             |                                                         |
| 13  | 7                     | 4                                           | 13                                                      | 8                     | 4                                           |                                                         |                       |                                             |                                                         |                       |                                             |                                                         |
| 14  | 8                     | 3                                           | 14                                                      | 8                     | 3                                           |                                                         |                       |                                             |                                                         |                       |                                             |                                                         |
| 15  | 7                     | 3                                           | 15                                                      | 9                     | 3                                           |                                                         |                       |                                             |                                                         |                       |                                             |                                                         |
| 16  | 19                    | 5                                           | 16                                                      | 14                    | 4                                           |                                                         |                       |                                             |                                                         |                       |                                             |                                                         |

*We censored risk factors reported after diagnosis of CVD, type 2 diabetes, or cancer, and risk factors reported after age 85.
Figure S13. Plots of joint patterns of risk factors after missing data imputation by pooling the Nurses’ Health Study and the Health Professionals Follow-up Study.
We identified patterns of risk factor and patterns of change in risk factor using smoothing mixture models, classified participants according to joint group membership, and plotted trajectories of risk factor within each category.
Table S11. Associations of patterns of risk factors and patterns of change in risk factors after missing data imputation with odds ratios (OR) of achieving longevity by pooling the Nurses’ Health Study and the Health Professionals Follow-up Study.

| Risk factors | Medium | Low | High | Change in risk factors | No change | Increase | Decrease |
|--------------|--------|-----|------|------------------------|------------|----------|----------|
| BMI          |        |     |      |                        |            |          |          |
| Model 1      | 1.00   | 0.82 (0.80, 0.84) | 0.53 (0.51, 0.55) | Model 3     | 1.00       | 0.61 (0.59, 0.63) | 0.55 (0.53, 0.57) |
| Model 2      | 1.00   | 0.86 (0.83, 0.89) | 0.58 (0.56, 0.61) | Model 4     | 1.00       | 0.83 (0.80, 0.86) | 0.58 (0.55, 0.60) |
| Smoking      |        |     |      |                        |            |          |          |
| Model 1      | 1.00   | 3.00 (2.87, 3.15) | 0.49 (0.44, 0.53) | Model 3     | 1.00       | 0.40 (0.35, 0.46) | 1.14 (1.04, 1.25) |
| Model 2      | 1.00   | 2.89 (2.74, 3.05) | 0.50 (0.45, 0.55) | Model 4     | 1.00       | 0.54 (0.46, 0.62) | 1.11 (1.00, 1.23) |
| Diet quality (AHEI) |       |     |      |                        |            |          |          |
| Model 1      | 1.00   | 0.74 (0.71, 0.76) | 1.21 (1.16, 1.25) | Model 3     | 1.00       | 1.19 (1.14, 1.24) | 0.87 (0.83, 0.90) |
| Model 2      | 1.00   | 0.90 (0.86, 0.93) | 1.05 (1.01, 1.10) | Model 4     | 1.00       | 1.10 (1.05, 1.15) | 0.91 (0.87, 0.95) |
| Physical activity |     |     |      |                        |            |          |          |
| Model 1      | 1.00   | 0.60 (0.58, 0.62) | 1.14 (1.10, 1.19) | Model 3     | 1.00       | 0.97 (0.93, 1.01) | 0.75 (0.69, 0.82) |
| Model 2      | 1.00   | 0.66 (0.63, 0.69) | 1.22 (1.17, 1.28) | Model 4     | 1.00       | 1.01 (0.97, 1.06) | 0.83 (0.76, 0.91) |
| Alcohol intake |     |     |      |                        |            |          |          |
| Model 1      | 1.00   | 0.69 (0.67, 0.71) | 0.95 (0.91, 0.98) | Model 3     | 1.00       | 0.95 (0.90, 0.99) | 0.59 (0.55, 0.64) |
| Model 2      | 1.00   | 0.89 (0.85, 0.93) | 0.95 (0.91, 1.00) | Model 4     | 1.00       | 0.94 (0.89, 1.00) | 0.74 (0.68, 0.81) |

Model 1 is univariate analysis.

Model 2 adjusted for baseline age (continuous), race (White, Black, Asian, and other), family histories of cancer (yes, no), myocardial infarction (yes, no), and type 2 diabetes (yes, no), multivitamin use (yes, no), menopausal status (yes, no, women only), postmenopausal hormone use (yes, no, women only), cohort, education (registered nurse, bachelor degree, master degree and higher, women only), social economic status (annual family income [quartiles] for women and work status [disabled, retired, part-time, full-time] for men), and the other four risk factors at baseline as continuous variables.

Model 3 is model 1 additionally adjusting for the risk factor at baseline (continuous) and the risk factor patterns (categorical).
Model 4 is model 2 additionally adjusting for the risk factor at baseline (continuous) and the risk factor patterns (categorical).

Figure S14. Kaplan–Meier plot of survival during follow up in the Nurses’ Health Study and the Health Professionals Follow-up Study.
Figure S15. Plots of joint patterns of diet quality, physical activity, alcohol intake using unstandardized data by pooling the Nurses’ Health Study and the Health Professionals Follow-up Study.

We identified patterns of risk factor and patterns of change in risk factor using smoothing mixture models, classified participants according to joint group membership, and plotted trajectories of risk factor within each category.
Table S12. Associations of patterns of risk factor and patterns of change in risk factor using unstandardized data with odds ratios (OR) of achieving longevity by pooling the Nurses’ Health Study and the Health Professionals Follow-up Study.

| Patterns of risk factor | Medium | Low   | High  | Patterns of change in risk factor | No change | Increase | Decrease |
|-------------------------|--------|-------|-------|----------------------------------|-----------|----------|----------|
| Diet quality (AHEI)     |        |       |       | Diet quality (AHEI)              |           |          |          |
| Cases/Participants      | 18346/30941 | 11619/21937 | 10588/17285 | Cases/Participants              | 19432/32253 | 9678/17909 | 11443/19908 |
| Model 1                 | 1.00   | 0.77 (0.75, 0.80) | 1.09 (1.04, 1.13) | Model 3                         | 1.00      | 0.85 (0.82, 0.89) | 0.78 (0.75, 0.81) |
| Model 2                 | 1.00   | 0.76 (0.73, 0.79) | 1.11 (1.06, 1.16) | Model 4                         | 1.00      | 0.99 (0.94, 1.03) | 0.81 (0.77, 0.85) |
| Physical activity       |        |       |       | Physical activity                |           |          |          |
| Cases/Participants      | 12878/20467 | 25323/45630 | 2352/4065   | Cases/Participants              | 28694/49364 | 7120/11862 | 2592/5084  |
| Model 1                 | 1.00   | 0.73 (0.71, 0.76) | 0.81 (0.76, 0.87) | Model 3                         | 1.00      | 0.96 (0.91, 1.01) | 0.55 (0.51, 0.59) |
| Model 2                 | 1.00   | 0.72 (0.70, 0.75) | 0.86 (0.80, 0.93) | Model 4                         | 1.00      | 1.06 (1.00, 1.12) | 0.68 (0.63, 0.73) |
| Alcohol intake          |        |       |       | Alcohol intake                   |           |          |          |
| Cases/Participants      | 9896/16869 | 29757/51438 | 2736/5614   | Cases/Participants              | 32657/52517 | 3609/6342 | 1929/3928  |
| Model 1                 | 1.00   | 0.97 (0.93, 1.00) | 0.67 (0.63, 0.71) | Model 3                         | 1.00      | 0.75 (0.71, 0.80) | 0.62 (0.57, 0.68) |
| Model 2                 | 1.00   | 0.86 (0.83, 0.90) | 0.77 (0.72, 0.83) | Model 4                         | 1.00      | 1.08 (1.01, 1.16) | 0.74 (0.67, 0.82) |

Model 1 is univariate analysis.

Model 2 adjusted for baseline age (continuous), race (White, Black, Asian, and other), family histories of cancer (yes, no), myocardial infarction (yes, no), and type 2 diabetes (yes, no), multivitamin use (yes, no), menopausal status (yes, no, women only), postmenopausal hormone use (yes, no, women only), cohort, education (registered nurse, bachelor degree, master degree and higher, women only), social economic status (annual family income [quartiles] for women and work status [disabled, retired, part-time, full-time] for men), and the other four risk factors at baseline as continuous variables.

Model 3 is model 1 additionally adjusting for the risk factor at baseline (continuous) and the risk factor patterns (categorical).

Model 4 is model 2 additionally adjusting for the risk factor at baseline (continuous) and the risk factor patterns (categorical).
Table S13. Comparison between participants who were excluded due to unknown longevity status (i.e. who were alive and did not reach 85) and the main study population.

|                    | NHS (1984)          | HPFS (1986)          |
|--------------------|---------------------|----------------------|
|                    | Main population     | Participants with unknown longevity status | Main population | Participants with unknown longevity status |
|                    | (n=51442)           | (n=26210)            | (n=33904)       | (n=15330)                                    |
| Age                | 54.21 (5.41)        | 42.86 (3.14)         | 59.52 (8.04)    | 44.36 (3.19)                                 |
| BMI (kg/m²)        | 25.23 (4.72)        | 24.37 (4.46)         | 25.01 (5.27)    | 24.73 (4.52)                                 |
| Alcohol intake (g) | 7.37 (11.92)        | 6.16 (9.91)          | 12.04 (16.33)   | 9.86 (13.30)                                 |
| Physical activity  | 13.99 (20.42)       | 14.51 (22.09)        | 19.53 (26.59)   | 23.65 (32.28)                                |
| AHEI               | 44.22 (10.32)       | 41.41 (9.83)         | 47.77 (10.92)   | 45.43 (10.62)                                |
| Current smoker, %  | 26                  | 21                   | 11              | 8                                             |
| Caucasian, %       | 98                  | 98                   | 91              | 91                                           |
| Menopausal status, % | 69                | 8                    | NA              | NA                                           |
| Postmenopausal hormone use, % | 18          | 6                    | NA              | NA                                           |
| Family history of cancer, % | 42            | 36                   | 9               | 7                                            |
| Family history of cardiovascular disease, % | 36         | 31                   | 35              | 29                                           |
| Family history of diabetes, % | 23           | 21                   | 15              | 11                                           |
| Multivitamin use, % | 38                  | 36                   | 33              | 29                                           |
Table S14. Associations of patterns of risk factor and patterns of change in risk factor with odds ratios (OR) of mortality further including participants that were excluded due to unknown longevity status by pooling the Nurses’ Health Study and the Health Professionals Follow-up Study.

| Risk factors | Medium | Low | High | Change in risk factors | No change | Increase | Decrease |
|--------------|--------|-----|------|------------------------|-----------|----------|----------|
| BMI          |        |     |      |                        |           |          |          |
| Cases/Participants | 22904/49215 | 27025/57979 | 7162/15198 | Cases/Participants | 20558/47718 | 12797/33165 | 18365/33377 |
| Model 1      | 1.00   | 1.03 (1.00, 1.05) | 1.02 (0.98, 1.06) | Model 3 | 1.00   | 0.58 (0.56, 0.60) | 1.96 (1.90, 2.02) |
| Model 2      | 1.00   | 1.00 (0.96, 1.03) | 1.75 (1.67, 1.84) | Model 4 | 1.00   | 0.99 (0.95, 1.04) | 1.46 (1.40, 1.51) |
| Smoking      |        |     |      |                        |           |          |          |
| Cases/Participants | 6810/11682 | 44749/103633 | 5338/6891 | Cases/Participants | 45508/103297 | 1625/2497 | 4975/9767 |
| Model 1      | 1.00   | 0.48 (0.46, 0.50) | 2.57 (2.40, 2.75) | Model 3 | 1.00   | 0.85 (0.77, 0.93) | 0.72 (0.66, 0.77) |
| Model 2      | 1.00   | 0.36 (0.34, 0.38) | 1.94 (1.78, 2.12) | Model 4 | 1.00   | 1.72 (1.52, 1.95) | 0.70 (0.64, 0.78) |
| AHEI         |        |     |      |                        |           |          |          |
| Cases/Participants | 24478/53366 | 17631/37512 | 13415/28677 | Cases/Participants | 23075/50907 | 9992/27788 | 13114/28040 |
| Model 1      | 1.00   | 1.05 (1.02, 1.08) | 1.02 (0.99, 1.05) | Model 3 | 1.00   | 0.81 (0.78, 0.84) | 0.88 (0.85, 0.92) |
| Model 2      | 1.00   | 1.16 (1.12, 1.20) | 0.94 (0.90, 0.98) | Model 4 | 1.00   | 0.95 (0.91, 1.00) | 1.18 (1.13, 1.24) |
| Physical activity |        |     |      |                        |           |          |          |
| Cases/Participants | 14685/36412 | 41241/80802 | 2611/6630 | Cases/Participants | 41088/82399 | 7812/22674 | 2244/5689 |
| Model 1      | 1.00   | 1.57 (1.53, 1.61) | 0.95 (0.90, 1.00) | Model 3 | 1.00   | 0.59 (0.57, 0.62) | 0.78 (0.73, 0.84) |
| Model 2      | 1.00   | 1.28 (1.24, 1.32) | 1.11 (1.04, 1.19) | Model 4 | 1.00   | 0.83 (0.80, 0.87) | 1.46 (1.34, 1.59) |
| Alcohol intake |        |     |      |                        |           |          |          |
| Cases/Participants | 12489/27992 | 38183/82791 | 4903/8992 | Cases/Participants | 39277/89581 | 3932/11527 | 3057/5985 |
| Model 1      | 1.00   | 1.13 (1.10, 1.16) | 1.47 (1.40, 1.54) | Model 3 | 1.00   | 0.64 (0.61, 0.68) | 0.85 (0.79, 0.92) |
| Model 2      | 1.00   | 1.06 (1.02, 1.10) | 1.26 (1.18, 1.34) | Model 4 | 1.00   | 1.08 (1.01, 1.16) | 1.18 (1.07, 1.30) |

Model 1 is univariate analysis.
Model 2 adjusted for baseline age (continuous), race (White, Black, Asian, and other), family histories of cancer (yes, no), myocardial infarction (yes, no), and type 2 diabetes (yes, no), multivitamin use (yes, no), menopausal status (yes, no, women only), postmenopausal hormone use (yes, no, women only), education (registered nurse, bachelor degree, master degree and higher, women only), social economic status (annual family income [quartiles] for women and work status [disabled, retired, part-time, full-time] for men), and the other four risk factors at baseline as continuous variables.
Model 3 is model 1 additionally adjusting for the risk factor at baseline (continuous) and the risk factor patterns (categorical).
Model 4 is model 2 additionally adjusting for the risk factor at baseline (continuous) and the risk factor patterns (categorical).
Table S15. Associations of patterns of risk factor and patterns of change in risk factor with odds ratios (OR) of achieving longevity excluding participants who were below 51 years in the Nurses’ Health Study and 49 years in the Health Professionals Follow-up Study.

| Risk factors | Medium | Low       | High      | Change in risk factors | No change | Increase | Decrease |
|--------------|--------|-----------|-----------|------------------------|-----------|----------|----------|
| BMI          | 15705/25184 | 18926/30518 | 2798/5445 | Model 1                | 1.00      | 0.98 (0.94, 1.01) | 0.63 (0.59, 0.67) |
|              |         |           |           | Model 2                | 1.00      | 0.98 (0.94, 1.02) | 0.63 (0.59, 0.67) |
| BMI          | 21392/31101 | 2649/4466 | 10253/18634 | Model 3               | 1.00      | 0.72 (0.67, 0.77) | 0.56 (0.54, 0.58) |
|              |         |           |           | Model 4               | 1.00      | 0.86 (0.80, 0.93) | 0.58 (0.56, 0.61) |
| Smoking      | 1.00      | 0.98 (0.94, 1.01) | 0.63 (0.59, 0.67) | Model 3       | 1.00      | 0.62 (0.54, 0.72) | 1.49 (1.33, 1.66) |
|              |         |           |           | Model 4               | 1.00      | 0.65 (0.56, 0.75) | 1.44 (1.27, 1.63) |
| AHEI         | 1.00      | 2.83 (2.68, 2.99) | 0.55 (0.50, 0.60) | Model 3       | 1.00      | 0.62 (0.54, 0.72) | 1.49 (1.33, 1.66) |
|              |         |           |           | Model 4               | 1.00      | 0.65 (0.56, 0.75) | 1.44 (1.27, 1.63) |
| Physical activity | 1.00 | 2.72 (2.56, 2.90) | 0.56 (0.51, 0.62) | Model 3       | 1.00      | 0.65 (0.56, 0.75) | 1.44 (1.27, 1.63) |
| Alcohol intake | 1.00 | 2.72 (2.56, 2.90) | 0.56 (0.51, 0.62) | Model 3       | 1.00      | 0.65 (0.56, 0.75) | 1.44 (1.27, 1.63) |

Model 1 is univariate analysis.
Model 2 adjusted for baseline age (continuous), race (White, Black, Asian, and other), family histories of cancer (yes, no), myocardial infarction (yes, no), and type 2 diabetes (yes, no), multivitamin use (yes, no), menopausal status (yes, no, women only), postmenopausal hormone use (yes, no, women only), cohort, education (registered nurse, bachelor degree, master degree and higher, women only), social economic status (annual family income [quartiles] for women and work status [disabled, retired, part-time, full-time] for men), and the other four risk factors at baseline as continuous variables.
Model 3 is model 1 additionally adjusting for the risk factor at baseline (continuous) and the risk factor patterns (categorical).
Model 4 is model 2 additionally adjusting for the risk factor at baseline (continuous) and the risk factor patterns (categorical).
Reference

1. Rimm EB, Stampfer MJ, Colditz GA, Chute CG, Litin LB, Willett WC. Validity of self-reported waist and hip circumferences in men and women. Epidemiology 1990;1:466-73.
2. Feskanich D, Rimm EB, Giovannucci EL, et al. Reproducibility and validity of food intake measurements from a semi-quantitative food frequency questionnaire. Journal of the American Dietetic Association 1993;93:790-6.
3. Rimm EB, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, Willett WC. Reproducibility and validity of an expanded self-administered semi-quantitative food frequency questionnaire among male health professionals. American journal of epidemiology 1992;135:1114-26; discussion 27-36.
4. Salvini S, Hunter DJ, Sampson L, et al. Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. International journal of epidemiology 1989;18:858-67.
5. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semi-quantitative food frequency questionnaire. American journal of epidemiology 1985;122:51-65.
6. Chiuve SE, Fung TT, Rimm EB, et al. Alternative dietary indices both strongly predict risk of chronic disease. The Journal of nutrition 2012;142:1009-18.
7. Wolf AM, Hunter DJ, Colditz GA, et al. Reproducibility and validity of a self-administered physical activity questionnaire. International journal of epidemiology 1994;23:991-9.
8. Ainsworth BE, Haskell WL, Leon AS, et al. Compendium of physical activities: classification of energy costs of human physical activities. Medicine and science in sports and exercise 1993;25:71-80.
9. Ding M, Chavarro JE, Fitzmaurice GM. Development of a mixture model allowing for smoothing functions of longitudinal trajectories. Stat Methods Med Res 2020:962280220966019.
10. Nagin DS, Tremblay RE. Analyzing developmental trajectories of distinct but related behaviors: a group-based method. Psychol Methods 2001;6:18-34.
11. Berlin KS, Parra GR, Williams NA. An introduction to latent variable mixture modeling (part 2): longitudinal latent class growth analysis and growth mixture models. J Pediatr Psychol 2014;39:188-203.
12. Jones BL, Nagin DS, Roeder K. A SAS Procedure Based on Mixture Models for Estimating Developmental Trajectories. Sociol Methods Res 2001;29:374-93.