






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

Ming Ding <sup>1,2</sup> Garrett M Fitzmaurice,<sup>3,4,5</sup> Mariel Arvizu <sup>1</sup> Walter C Willett,<sup>1,6,7</sup> JoAnn E Manson,<sup>6,7,8</sup> Kathryn M Rexrode,<sup>9</sup> Frank B Hu,<sup>1,6,7</sup> Jorge E Chavarro <sup>1,6,7</sup>

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjmed-2021-000098>).

For numbered affiliations see end of article.  
Correspondence to: Dr Ming Ding, Department of Emergency Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; [ming\\_ding@med.unc.edu](mailto:ming_ding@med.unc.edu)

Cite this as: *BMJMED* 2022;1:e000098. doi:10.1136/bmjmed-2021-000098

Received: 18 December 2021  
Accepted: 12 July 2022

## 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 low 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  $\geq 85$  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 had physical activity levels and alcohol consumption that met public health recommendations.

## Introduction

Life expectancy has increased considerably worldwide.<sup>1</sup> 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.<sup>1</sup> Prolonged life expectancy has seen a growth in age related diseases, which impair quality of life and impose a substantial medical burden on society.<sup>2</sup> 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.<sup>3-7</sup> Firstly, studies on lifestyle and

### 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

mortality usually use baseline or cumulative data,<sup>3-7</sup> 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,<sup>8</sup> are under used. These methods could be used creatively to efficiently capture patterns of modifiable risk factors.<sup>8</sup> Thirdly, although research has been conducted on mortality, few studies have taken into account morbidity status, such as cardiovascular 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 121 700 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 51 529 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, callisthenics, aerobics, aerobic dance, use of rowing machine, lap swimming, playing tennis, and playing squash or racquet ball.<sup>9</sup> Weekly energy expenditure in metabolic equivalent task hours (MET hours/week) was calculated.<sup>10</sup> 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.<sup>11</sup> 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.

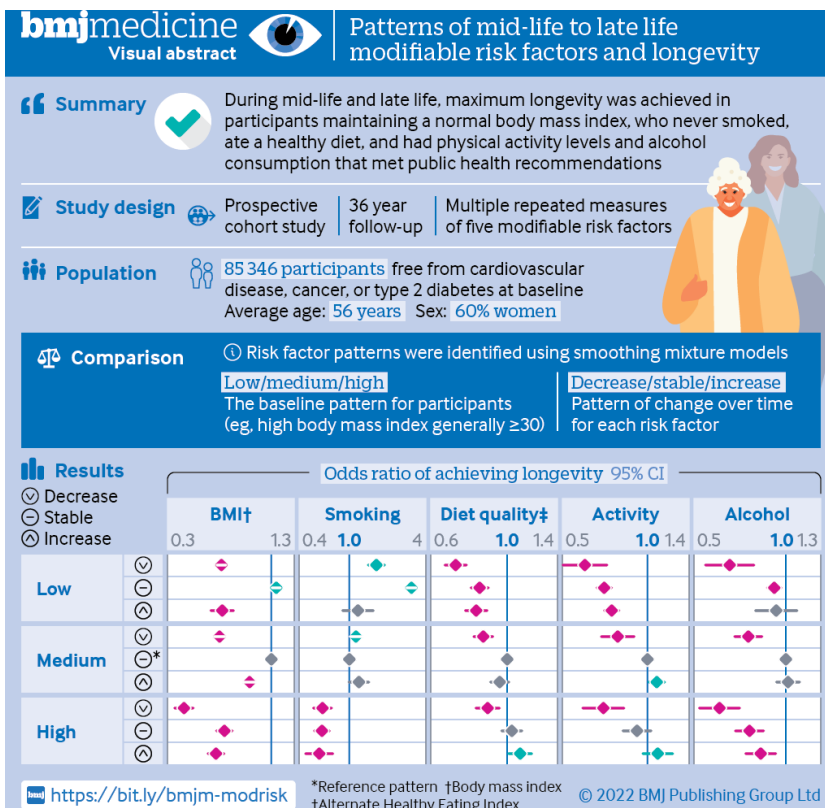


Figure 1 | Visual abstract

### 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.<sup>12</sup> 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  $\geq 85$  years, according to the life expectancy of the US population and the vital status of our data.<sup>1</sup> 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 3900-4590 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, anti-hypertensive agents (including  $\beta$  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.<sup>8</sup> 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/>

mingding-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.<sup>13</sup> 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,<sup>14</sup> we standardised physical activity and dietary factors (alcohol intake and Alternative Healthy Eating Index score) during each follow-up cycle,<sup>15</sup> 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), or 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 the number of years lived from baseline until death or to the end of follow-up, whichever came first. We assumed an exponential distribution of survival time and applied transformation  $100 \times (e^{\beta} - 1)$  to the regression coefficient  $\beta$ , which can be interpreted as the per cent increase in life expectancy comparing two groups. We pooled the data of both cohorts to obtain overall effects. All statistical tests were two sided at a type I error rate of 0.05. The smoothing mixture model was performed with R 3.5.0, and logistic regression and accelerated failure time model were applied with SAS version 9.2 for UNIX (SAS Institute, Cary, NC).

#### Patient and public involvement

No patients were involved in setting the research question or the outcome measures, nor were they involved in developing plans for design or implementation of the study. No patients were asked to advise on interpretation or writing up of results. There are no plans to disseminate the results of the research to study participants or the relevant patient community.

#### Results

Our analysis involved 85 346 participants, of whom 51 442 were women from the Nurses' Health Study and 33 904 were men from the Health Professionals Follow-up Study. Mean age at baseline was 56 years (standard deviation 5 years). In the Nurses' Health Study, 29 016 participants achieved longevity (including 10 470 participants who died after age 85 years and 18 546 who were alive at the end of follow-up), and 22 426 participants died before age 85 years. In the Health Professionals Follow-up Study, 17 026 participants achieved longevity (including 10 834 participants who died after age 85 years and 6 192 who were alive at the end of follow-up), and 16 878 participants died before age 85. Over 36 years of follow-up, 46 042 participants achieved longevity and 25 322 participants achieved healthy longevity (no diagnosis of cardiovascular disease, type 2 diabetes, or cancer). Online supplemental table S3 shows the baseline characteristics of

participants by trajectories of modifiable risk factors. Trajectory patterns of a risk factor were strongly correlated with the factor at baseline, highlighting the importance of identifying trajectories of change in the factor which removes the effect of the baseline risk factor. Participants with healthier patterns were more likely to take multivitamins, and trajectory patterns of a risk factor were moderately correlated with other factors.

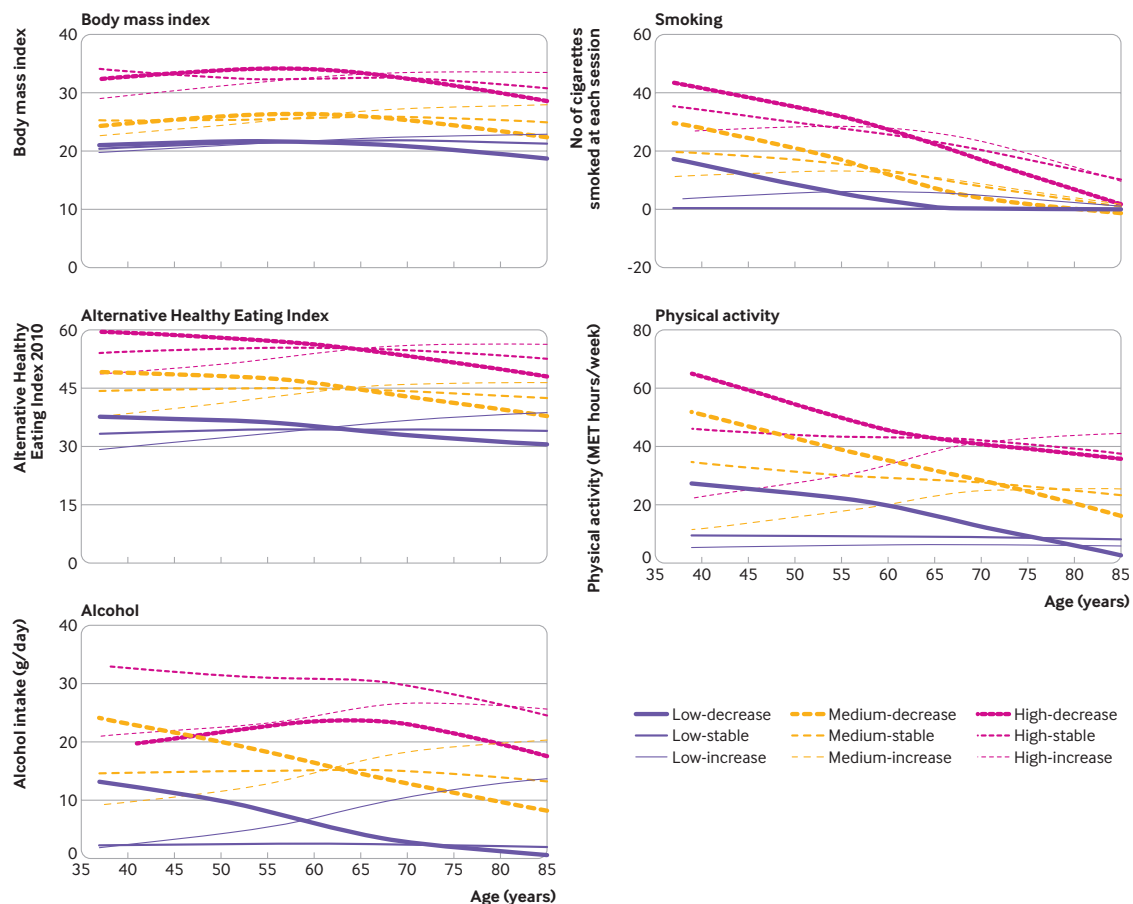
In each cohort, we identified trajectories of risk factors as high, medium, and low patterns (online supplemental figures S3 and S4) and trajectories of change in risk factors as increase, no change, and decrease patterns (online supplemental figures S5 and S6). We classified nine joint patterns: high-stable, 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  $\geq 30$ ).

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).<sup>16</sup>

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<sup>17</sup>; 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.



**Figure 2 |** Joint patterns of modifiable risk factors and change in factor from baseline by pooling data from the Nurses' Health Study and the Health Professionals Follow-up Study. For smoking, because the number of cigarettes smoked by participants decreased over time, increase refers to participants with the least decrease in the number of cigarettes smoked, and stable refers to participants with a medium decrease in the number of cigarettes smoked. Patterns of risk factors and patterns of change in risk factors were identified with smoothing mixture models, participants were classified according to joint group membership, and mean value of the risk factor with age within each category was plotted with the command function `loess.smooth` in R

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 (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.

**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			Change in risk factor from baseline		
	Medium	Low	High	No change	Increase	Decrease
<b>Body mass index</b>						
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
Model 1 or 3	1.00	1.05 (1.02 to 1.08)	0.52 (0.50 to 0.55)	1.00	0.43 (0.41 to 0.45)	0.56 (0.54 to 0.58)
Model 2 or 4	1.00	0.98 (0.94 to 1.01)	0.60 (0.57 to 0.64)	1.00	0.72 (0.68 to 0.76)	0.56 (0.54 to 0.59)
<b>Smoking</b>						
No of participants/ Total No of participants	3105/8817	40 637/67 990	1095/4388	39 925/67 287	527/1647	2336/6772
Model 1 or 3	1.00	2.93 (2.80, 3.07)	0.60 (0.56 to 0.66)	1.00	0.49 (0.43 to 0.55)	1.31 (1.20 to 1.44)
Model 2 or 4	1.00	2.75 (2.61 to 2.90)	0.55 (0.50 to 0.60)	1.00	0.63 (0.55 to 0.72)	1.26 (1.14 to 1.40)
<b>Diet quality (AHEI)</b>						
No of participants/ Total No of participants	20 439/35 470	12 395/25 111	11 808/18 972	20 018/32 453	11 114/18 895	9951/17 717
Model 1 or 3	1.00	0.72 (0.69 to 0.74)	1.22 (1.18 to 1.26)	1.00	1.01 (0.96 to 1.05)	0.66 (0.63 to 0.69)
Model 2 or 4	1.00	0.85 (0.81 to 0.88)	1.05 (1.01 to 1.09)	1.00	1.03 (0.98 to 1.08)	0.81 (0.77 to 0.85)
<b>Physical activity</b>						
No of participants/ Total No of participants	14 600/23 867	28 071/54 329	2629/4577	16 862/30 344	23 055/40 387	1259/2570
Model 1 or 3	1.00	0.68 (0.66 to 0.70)	0.86 (0.80 to 0.91)	1.00	1.01 (0.97 to 1.06)	0.56 (0.50 to 0.62)
Model 2 or 4	1.00	0.71 (0.69 to 0.74)	0.88 (0.82 to 0.95)	1.00	1.10 (1.05 to 1.15)	0.72 (0.64 to 0.80)
<b>Alcohol intake</b>						
No of participants/ Total No of participants	10 418/18 068	31 402/55 486	2872/6076	35 516/58 498	3680/6602	1973/4106
Model 1 or 3	1.00	0.94 (0.91 to 0.98)	0.66 (0.62 to 0.70)	1.00	0.78 (0.74 to 0.83)	0.64 (0.58 to 0.69)
Model 2 or 4	1.00	0.91 (0.87 to 0.94)	0.77 (0.72 to 0.82)	1.00	1.01 (0.94 to 1.09)	0.79 (0.71 to 0.87)

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), use of cholesterol lowering drugs (yes, no), and the other four risk factors at baseline as continuous variables.

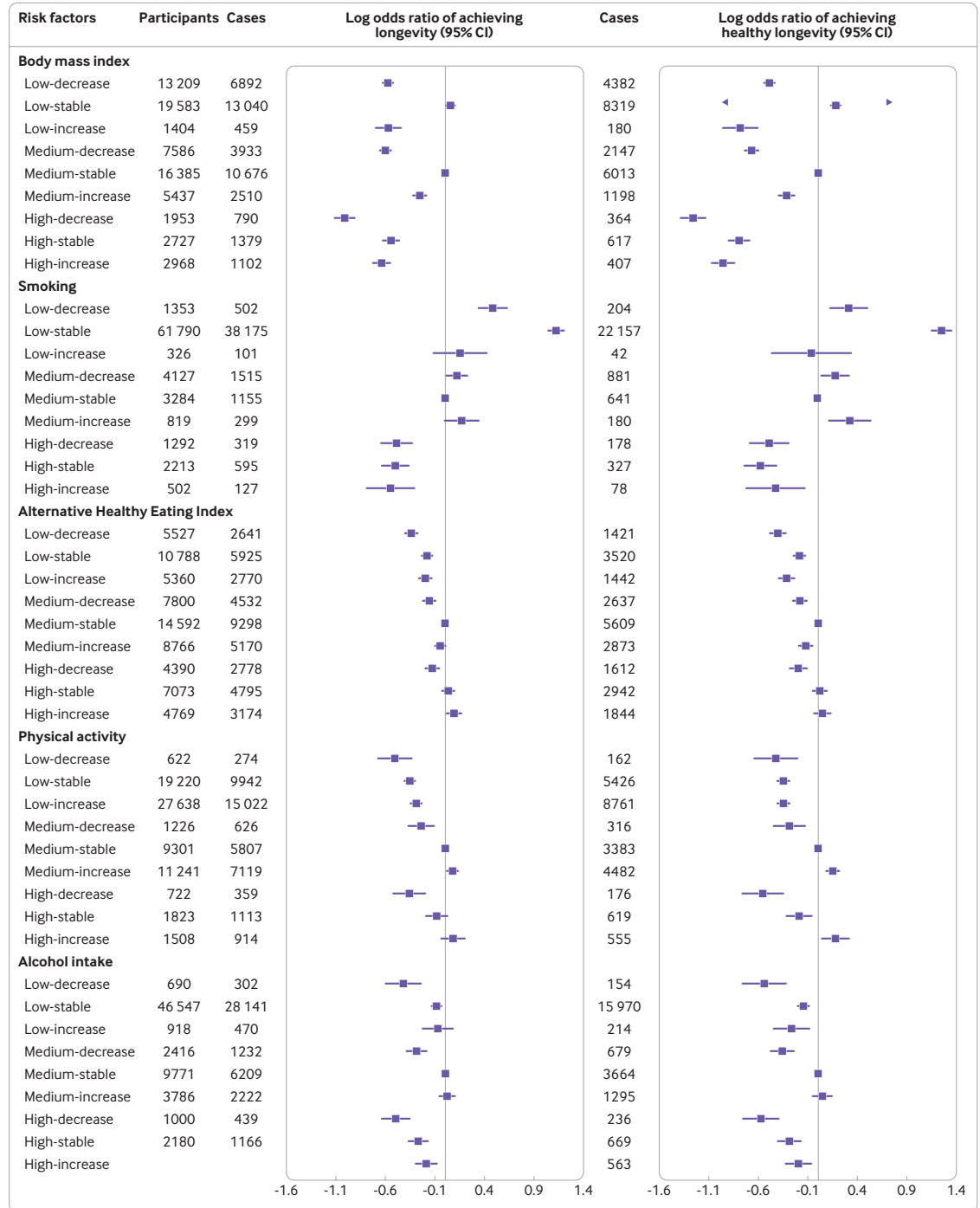
Model 3 applies to changes in risk factors and is adjusted for all variables included in model 1, as well as group membership of trajectories of the risk factor (categorical).

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).

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

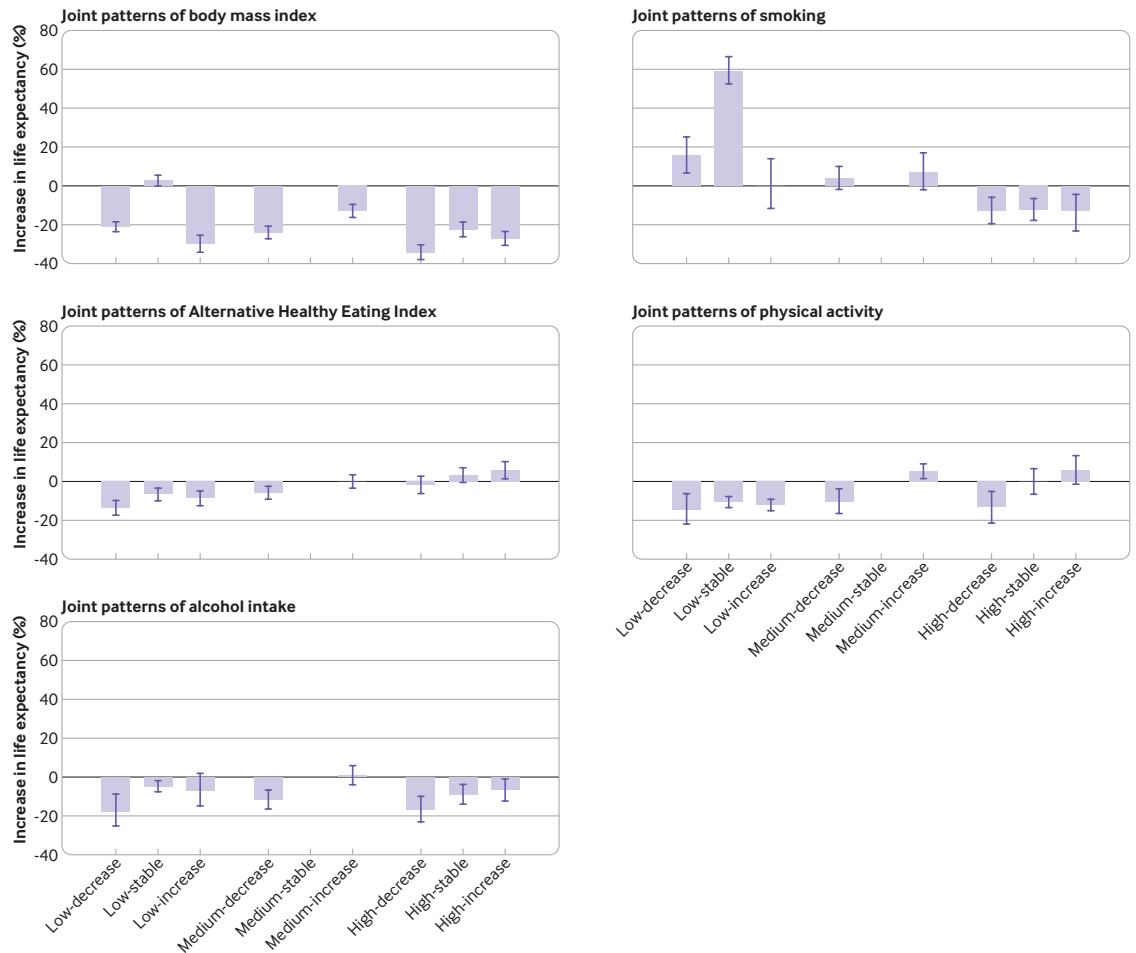
a low-stable pattern (mainly never smokers) achieved the highest longevity, followed by a low-decrease pattern (light smokers who stopped smoking during follow-up), a low-increase pattern (light smokers with a long duration of smoking), and participants in the medium patterns. Participants with a high-increase 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



**Figure 3 | Associations between joint patterns of modifiable risk factors and change in risk factors from baseline and odds ratios of achieving longevity and healthy longevity (ie, no diagnosis of cardiovascular disease, type 2 diabetes, or cancer) by pooling data from the Nurses' Health Study and Health Professionals Follow-up Study. Logistic model (longevity) and multinomial logistic model (healthy longevity) 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), use of aspirin (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**

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



**Figure 4 | Associations between joint patterns of modifiable risk factors and change in the factor from baseline and life expectancy by pooling data from the Nurses’ Health Study and Health Professionals Follow-up Study. Accelerated failure time model adjusting 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), use of 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**

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



**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	9350	7040	5756
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 (-13 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.					

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, online supplemental table S11).

Kaplan-Meier curves of survival (online supplemental figure S14) showed constant slopes during most of the follow-up periods. Thus we considered it appropriate to assume a constant hazard and an exponential function of survival time for the accelerated failure time model. We plotted the joint patterns of physical activity, quality of diet, and alcohol intake with unstandardised data (online supplemental figure S15) and found an obvious trend of

higher quality of diet over time, showing the importance of standardising these factors. The associations between patterns of these factors with unstandardised data and longevity were similar to our main findings, showing the robustness of our findings (online supplemental table S12).

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.

## 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 (ie, 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.<sup>18 19</sup> 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.<sup>20-24</sup> 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.<sup>25</sup> 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.<sup>26</sup>

### Comparison with other studies

The association between physical activity with mortality has been extensively studied,<sup>27</sup> 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.<sup>17</sup> 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.<sup>28 29</sup> 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,<sup>30</sup> 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 mortality<sup>31 32</sup> and that improving the quality of your diet was associated with a lower risk of death.<sup>6</sup> Randomised clinical trials have shown that a healthy diet promotes weight loss,<sup>33</sup> substantially lowers blood pressure,<sup>34 35</sup> and efficiently prevents cardiovascular disease.<sup>36</sup> A plausible mechanism is that the beneficial effects might be mediated through improvement in endothelial function, inflammatory markers, and insulin resistance,<sup>37</sup> 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.<sup>16</sup> 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.<sup>38</sup> Prospective studies found that moderate intake of alcohol was

associated with the lowest risk of total mortality and the highest probability of achieving longevity.<sup>39,40</sup> 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.<sup>41</sup> 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.<sup>41</sup> 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.<sup>42</sup> Alcohol intake with family and friends has social and psychological benefits that might also contribute to health and wellbeing.<sup>43</sup> We found that high patterns of alcohol intake were less likely to be associated with longevity, consistent with previous studies,<sup>38–40</sup> and evidence has been convincing that alcohol consumption increases the risk of breast cancer.<sup>44</sup>

Smoking was the leading modifiable risk factor for mortality worldwide in 2015,<sup>45</sup> and 11.5% of global deaths were attributed to smoking.<sup>46</sup> 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).<sup>46</sup> 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 benefitted 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.<sup>8</sup> 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.<sup>8</sup> 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.<sup>47</sup> 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,<sup>48–51</sup> however, and body mass index and physical activity have been validated against standard measurements.<sup>9,52</sup> 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

<sup>1</sup>Department of Nutrition, Harvard T H Chan School of Public Health, Boston, MA, USA

<sup>2</sup>Department of Emergency Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

<sup>3</sup>Laboratory for Psychiatric Biostatistics, McLean Hospital, Harvard Medical School, Boston, MA, USA

<sup>4</sup>Department of Biostatistics, Harvard T H Chan School of Public Health, Boston, MA, USA

<sup>5</sup>Department of Psychiatry, Harvard Medical School, Boston, MA, USA

<sup>6</sup>Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

<sup>7</sup>Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

<sup>8</sup>Division of Preventive Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

<sup>9</sup>Division 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 R03AG060247, P30-DK046200, R33 DA042847, R01 HL34594, R01 HL088521, U01 CA167552, R01 HL35464, U01-HL145386, and UM1 CA186107 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/disclosure-of-interest/](http://www.icmje.org/disclosure-of-interest/) and declare: support from the National Institutes of Health for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

**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/0000-0001-9947-3446>

Mariel Arvizu <http://orcid.org/0000-0003-1860-7276>

Jorge E Chavarro <http://orcid.org/0000-0002-4436-9630>

### REFERENCES

- 1 GBD 2019 Demographics Collaborators. Global age-sex-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)30977-6
- 2 Dzau VJ, 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/NEJMp1912298
- 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/NEJMs1407211
- 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.115.017341
- 5 Ding M, Li J, Qi L, *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-Prieto M, Bhupathiraju SN, Mattei 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/NEJMoa1613502
- 7 Li Y, Schoufour 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:l6669. doi:10.1136/bmj.l6669
- 8 Ding M, Chavarro JE, Fitzmaurice GM. Development of a mixture model allowing for smoothing functions of longitudinal trajectories. *Stat Methods Med Res* 2021;30:549-62. 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* 1993;25:71-80. doi:10.1249/00005768-199301000-00011
- 11 Chiuve 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;140:1016-9. doi:10.1093/oxfordjournals.aje.a117191
- 13 Nagin DS, Tremblay RE. Analyzing developmental trajectories of distinct but related behaviors: a group-based method. *Psychol Methods* 2001;6:18-34. doi:10.1037/1082-989X.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/s10654-011-9626-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](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/paguide.pdf>
- 18 Rizzuto D, Orsini N, Qiu C, *et al*. Lifestyle, social factors, and survival after age 75: population based study. *BMJ* 2012;345:e5568. doi:10.1136/bmj.e5568
- 19 Wilhelmssen L, Svärdsudd 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.1365-2796.2010.02331.x



- 20 Berrington de Gonzalez A, Hartge P, Cerhan JR, *et al.* Body-mass index and mortality among 1.46 million white adults. *N Engl J Med* 2010;363:2211–9. doi:10.1056/NEJMoa1000367
- 21 Bhaskaran K, Dos-Santos-Silva I, Leon DA, *et al.* Association of BMI with overall and cause-specific mortality: a population-based cohort study of 3.6 million adults in the UK. *Lancet Diabetes Endocrinol* 2018;6:944–53. doi:10.1016/S2213-8587(18)30288-2
- 22 Global BMI Mortality Collaboration, Di Angelantonio E, Bhupathiraju S, *et al.* Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. *Lancet* 2016;388:776–86. doi:10.1016/S0140-6736(16)30175-1
- 23 GBD 2015 Obesity Collaborators, Afshin A, Forouzanfar MH, *et al.* Health effects of overweight and obesity in 195 countries over 25 years. *N Engl J Med* 2017;377:13–27. doi:10.1056/NEJMoa1614362
- 24 Aune D, Sen A, Prasad M, *et al.* BMI and all cause mortality: systematic review and non-linear dose-response meta-analysis of 230 cohort studies with 3.74 million deaths among 30.3 million participants. *BMJ* 2016;353:i2156. doi:10.1136/bmj.i2156
- 25 Zheng Y, Manson JE, Yuan C, *et al.* Associations of weight gain from early to middle adulthood with major health outcomes later in life. *JAMA* 2017;318:255–69. doi:10.1001/jama.2017.7092
- 26 Gregg EW, Gerzoff RB, Thompson TJ, *et al.* Intentional weight loss and death in overweight and obese U.S. adults 35 years of age and older. *Ann Intern Med* 2003;138:383–9. doi:10.7326/0003-4819-138-5-200303040-00007
- 27 Moore SC, Patel AV, Matthews CE, *et al.* Leisure time physical activity of moderate to vigorous intensity and mortality: a large pooled cohort analysis. *PLoS Med* 2012;9:e1001335. doi:10.1371/journal.pmed.1001335
- 28 Blond K, Brinkløv CF, Ried-Larsen M, *et al.* Association of high amounts of physical activity with mortality risk: a systematic review and meta-analysis. *Br J Sports Med* 2020;54:1195–201. doi:10.1136/bjsports-2018-100393
- 29 Arem H, Moore SC, Patel A, *et al.* Leisure time physical activity and mortality: a detailed pooled analysis of the dose-response relationship. *JAMA Intern Med* 2015;175:959–67. doi:10.1001/jamainternmed.2015.0533
- 30 Mok A, Khaw K-T, Luben R, *et al.* Physical activity trajectories and mortality: population based cohort study. *BMJ* 2019;365:l2323. doi:10.1136/bmj.l2323
- 31 Trichopoulou A, Costacou T, Bamia C, *et al.* Adherence to a Mediterranean diet and survival in a Greek population. *N Engl J Med* 2003;348:2599–608. doi:10.1056/NEJMoa025039
- 32 Knuops KTB, de Groot LCPGM, Kromhout D, *et al.* Mediterranean diet, lifestyle factors, and 10-year mortality in elderly European men and women: the HALE project. *JAMA* 2004;292:1433–9. doi:10.1001/jama.292.12.1433
- 33 Sacks FM, Bray GA, Carey VJ, *et al.* Comparison of weight-loss diets with different compositions of fat, protein, and carbohydrates. *N Engl J Med* 2009;360:859–73. doi:10.1056/NEJMoa0804748
- 34 Sacks FM, Svetkey LP, Vollmer WM, *et al.* Effects on blood pressure of reduced dietary sodium and the dietary approaches to stop hypertension (DASH) diet. DASH-Sodium Collaborative Research Group. *N Engl J Med* 2001;344:3–10. doi:10.1056/NEJM200101043440101
- 35 Appel LJ, Moore TJ, Obarzanek E, *et al.* A clinical trial of the effects of dietary patterns on blood pressure. DASH Collaborative Research Group. *N Engl J Med* 1997;336:1117–24. doi:10.1056/NEJM199704173361601
- 36 Estruch R, Ros E, Salas-Salvadó J, *et al.* Primary prevention of cardiovascular disease with a Mediterranean diet supplemented with extra-virgin olive oil or nuts. *N Engl J Med* 2018;378:e34. doi:10.1056/NEJMoa1800389
- 37 Esposito K, Marfella R, Ciotola M, *et al.* Effect of a mediterranean-style diet on endothelial dysfunction and markers of vascular inflammation in the metabolic syndrome: a randomized trial. *JAMA* 2004;292:1440–6. doi:10.1001/jama.292.12.1440
- 38 Di Castelnuovo A, Costanzo S, Bagnardi V, *et al.* Alcohol dosing and total mortality in men and women: an updated meta-analysis of 34 prospective studies. *Arch Intern Med* 2006;166:2437–45. doi:10.1001/archinte.166.22.2437
- 39 Xi B, Veeranki SP, Zhao M, *et al.* Relationship of alcohol consumption to all-cause, cardiovascular, and cancer-related mortality in U.S. adults. *J Am Coll Cardiol* 2017;70:913–22. doi:10.1016/j.jacc.2017.06.054
- 40 van den Brandt PA, Brandts L. Alcohol consumption in later life and reaching longevity: the Netherlands cohort study. *Age Ageing* 2020;49:395–402. doi:10.1093/ageing/afaa003
- 41 GBD 2016 Alcohol Collaborators. Alcohol use and burden for 195 countries and territories, 1990–2016: a systematic analysis for the global burden of disease study 2016. *Lancet* 2018;392:1015–35. doi:10.1016/S0140-6736(18)31310-2
- 42 Booyse FM, Pan W, Grenett HE, *et al.* Mechanism by which alcohol and wine polyphenols affect coronary heart disease risk. *Ann Epidemiol* 2007;17:S24–31. doi:10.1016/j.annepidem.2007.01.006
- 43 Chang S-C, Glymour M, Cornelis M, *et al.* Social integration and reduced risk of coronary heart disease in women: the role of lifestyle behaviors. *Circ Res* 2017;120:1927–37. doi:10.1161/CIRCRESAHA.116.309443
- 44 Smith-Warner SA, Spiegelman D, Yaun 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;388:1659–724. doi:10.1016/S0140-6736(16)31679-8
- 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;389:1885–906. 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:10.1080/01621459.1999.10474144
- 48 Feskanich D, Rimm EB, Giovannucci EL, *et al.* Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. *J Am Diet Assoc* 1993;93:790–6. doi:10.1016/0002-8223(93)91754-e
- 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/oxfordjournals.aje.a116211
- 50 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. *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;122:51–65. doi:10.1093/oxfordjournals.aje.a114086
- 52 Rimm EB, Stampfer MJ, Colditz GA, *et al.* Validity of self-reported waist and hip circumferences in men and women. *Epidemiology* 1990;1:466–73. doi:10.1097/00001648-199011000-00009

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjmed-2021-000098>).