

A simulation study of the impact of population-wide lifestyle modifications on life expectancy in the Chinese population

Supplementary Material.

Members of the China Kadoorie Biobank collaborative group:

International Steering Committee: Junshi Chen, Zhengming Chen (PI), Robert Clarke, Rory Collins, Liming Li (PI), Jun Lv, Richard Peto, Robin Walters.

National Co-ordinating Centre, Beijing: Xiao Han, Can Hou, Qingmei Xia, Chao Liu, Jun Lv, Pei Pei, Dianjianyi Sun, Canqing Yu, Lang Pan.

International Co-ordinating Centre, Oxford: Daniel Avery, Maxim Barnard, Derrick Bennett, Lazaros Belbasis, Ruth Boxall, Ka Hung Chan, Yiping Chen, Zhengming Chen, Charlotte Clarke, Johnathan Clarke; Robert Clarke, Huaidong Du, Ahmed Edris Mohamed, Hannah Fry, Simon Gilbert, Pek Kei Im, Andri Iona, Maria Kakkoura, Christiana Kartsonaki, Hubert Lam, Kuang Lin, James Liu, Mohsen Mazidi, Iona Millwood, Sam Morris, Qunhua Nie, Alfred Pozarickij, Maryanm Rahmati, Paul Ryder, Saredo Said, Dan Schmidt, Becky Stevens, Iain Turnbull, Robin Walters, Baihan Wang, Lin Wang, Neil Wright, Ling Yang, Xiaoming Yang, Pang Yao.

10 Regional Co-ordinating Centres: Qingdao CDC: Zengchang Pang, Ruqin Gao, Shanpeng Li, Haiping Duan, Shaojie Wang, Yongmei Liu, Ranran Du, Yajing Zang, Liang Cheng, Xiaocao Tian, Hua Zhang, Yaoming Zhai, Feng Ning, Xiaohui Sun, Feifei Li. **Licang CDC:** Silu Lv, Junzheng Wang, Wei Hou. **Heilongjiang Provincial CDC:** Wei Sun, Shichun Yan, Xiaoming Cui. **Nangang CDC:** Chi Wang, Zhenyuan Wu, Yanjie Li, Quan Kang. **Hainan Provincial CDC:** Huiming Luo, Tingting Ou. **Meilan CDC:** Xiangyang Zheng, Zhendong Guo, Shukuan Wu, Yilei Li, Huimei Li. **Jiangsu Provincial CDC:** Ming Wu, Yonglin Zhou, Jinyi Zhou, Ran Tao, Jie Yang, Jian Su. **Suzhou CDC:** Fang Liu, Jun Zhang, Yihe Hu, Yan Lu, Liangcai Ma, Aiyu Tang, Shuo Zhang, Jianrong Jin, Jingchao Liu. **Guangxi Provincial CDC:** Mei Lin, Zhenzhen Lu. **Liuzhou CDC:** Lifang Zhou, Changping Xie, Jian Lan, Tingping Zhu, Yun Liu, Liuping Wei, Liyuan Zhou, Ningyu Chen, Yulu Qin, Sisi Wang. **Sichuan Provincial CDC:** Xianping Wu, Ningmei Zhang, Xiaofang Chen, Xiaoyu Chang. **Pengzhou CDC:** Mingqiang Yuan, Xia Wu, Xiaofang Chen, Wei Jiang, Jiaqiu Liu, Qiang Sun. **Gansu Provincial CDC:** Faqing Chen, Xiaolan Ren, Caixia Dong. **Maiji CDC:** Hui Zhang, Enke Mao, Xiaoping Wang, Tao Wang, Xi zhang.

Henan Provincial CDC: Kai Kang, Shixian Feng, Huizi Tian, Lei Fan. **Huixian CDC:** XiaoLin Li, Huarong Sun, Pan He, Xukui Zhang. **Zhejiang Provincial CDC:** Min Yu, Ruying Hu, Hao Wang. **Tongxiang CDC:** Xiaoyi Zhang, Yuan Cao, Kaixu Xie, Lingli Chen, Dun Shen. **Hunan Provincial CDC:** Xiaojun Li, Donghui Jin, Li Yin, Huilin Liu, Zhongxi Fu. **Liuyang CDC:** Xin Xu, Hao Zhang, Jianwei Chen, Yuan Peng, Libo Zhang, Chan Qu.

Assessment of candidate predictors

The CKB collected information using a laptop-based questionnaire with built-in functions to avoid missing items and logic errors. Weight, height, blood pressure, and resting heart rate were measured with well-calibrated instruments. Information on smoking frequency, type, and amount of tobacco smoked per day was ascertained from all participants with a history of tobacco use. Those who had ceased use were also asked how long they had stopped smoking and why they had quit. Participants were asked about their typical drinking frequency, as well as those who drank at least once a week, the type of alcoholic beverage they drank habitually, and the amount of alcohol drunk on a typical drinking day in the previous year. For physical activity, the usual type and duration of activities in occupational, commuting, domestic, and leisure-time-related domains in the previous year were collected. To calculate the daily total physical activity level, we multiplied the metabolic equivalent of task (MET) value of each type of activity by the hours spent on that activity per day then summed the MET-hours for all activities.⁸ We used a validated qualitative food frequency questionnaire (FFQ) to assess habitual dietary intakes in the previous year.⁹ In the second resurvey, we also asked about the amount consumed by each food group on a typical consumption day.

Predictors were assessed similarly in CNHS. The questions about smoking were basically the same as those in the CKB questionnaire, with the exception that only cigarettes were used to calculate the daily smoking amount. A quantitative FFQ was used to collect information on the frequency and amount of various foods and alcoholic drinks consumed in the previous year. Physical activity was investigated using an adapted version of the International Physical Activity Questionnaire (IPAQ)-long form, and total physical activity level was calculated in the same way as in the CKB.

Calculation of the usual intake amount for each baseline frequency group

In the CKB, we used the resurvey data to estimate the usual intake amount of each baseline frequency group for fresh fruits, fresh vegetables, red meat, and fish/seafood, using the following steps with fresh fruits as an example:¹¹

1. We first calculated the mean intake amount on a typical consuming day for participants in each group, namely “never/rarely”, “monthly”, “1-3 days/wk”, “4-6 days/wk”, “daily”, using data collected during the 2nd resurvey (n=25,041). The corresponding values for each group were 0 g/d for the “never/rarely” group, 137.0 g/d for the “monthly” group, 142.5 g/d for the “1-3 days/wk” group, 140.3 g/d for the “4-6 days/wk” group, and 168.6 g/d for the “daily” group.
2. Assuming that the intake amount on a typical consuming day did not vary much from the baseline to the 2nd resurvey (i.e., people may change how frequently they consume fruits but the daily amount of fresh fruits consumption stays roughly the same), the mean daily grams derived in step 1 was used as the proxy for baseline mean daily grams.
3. The “baseline intake amount” for each group (column B in Supplementary Table 8) were estimated by multiplying the mean daily grams intake by the number of consuming days per month and dividing by 30. The number of consuming days per month assigned to each group was 0 for the “never/rarely” group, 2.5 days for the “monthly” group, 8.6 days for the “1-3 days/wk” group, 21.5 days for the “4-6 days/wk” group, and 30 days for the “daily” group. The estimates of the baseline intake amount were 0g, 11.4g, 40.8g, 100.5g, and 168.6g, respectively, for “never/rarely”, “monthly”, “1-3 days/wk”, “4-6 days/wk”, and “daily” group.
4. The consumption frequency recorded at the 1st resurvey (n=19,786) was used to assess the variation in fresh fruits consumption from the baseline. The “usual intake amount” for each group (column U in Supplementary Table 8) was calculated by multiplying the “baseline intake amount” for each group by the percentage of participants in each of the group at the 1st resurvey. For example, for the “never/rarely” group at baseline, 17.9% remained at this group at the 1st resurvey, 55.7% switched to the “monthly” group, 18.0% switched to the “1-3 days/wk” group, 4.2% switched to the “4-6 days/wk” group, and 4.2% switched to the “daily” group. The “usual intake amount” for the baseline “never/rarely” group was therefore equal to $17.9\% \times 0 + 55.7\% \times 11.4 + 18.0\% \times 40.8 + 4.2\% \times 100.5 + 4.2\% \times 168.6 = 25$ g/d. Corresponding values for other baseline groups, namely “monthly”, “1-3 days/wk”, “4-6 days/wk”, and “daily” group were 32.6g/d, 56.2g/d, 79.3g/d, and 123.4g/d, respectively.
5. The “usual intake amount” was then assigned to each individual participant in the

corresponding baseline group as the proxy of the intake amount.

6. Using the same approach as described above, the “usual intake amount” for each baseline group of fresh vegetables, red meat, and seafood were presented in Supplementary Table 9.

Supplementary Table 1. Characteristics of the study participants in the China Kadoorie Biobank (CKB) and the China Nutrition and Health Surveillance (CNHS).

	CKB population				CNHS population	
	Men		Women		Men	Women
	Derivation	Validation	Derivation	Validation		
No. of participants	140,135	70,068	201,678	100,840	31,515	36,049
Age, years	52.4 (10.9)	52.4 (10.9)	51.0 (10.5)	50.9 (10.5)	55.5 (12.5)	54.5 (12.3)
Urban area, %	43.3	43.7	44.6	44.6	51.3	49.6
Highest education, %						
No formal school	8.9	8.8	25.4	25.2	4.5	18.9
Primary school	33.4	33.2	31.5	31.3	28.5	33.2
Middle school	32.4	32.5	25.3	25.5	37.3	27.3
High school	17.5	17.4	13.4	13.6	17.8	11.1
College/university	7.8	8.1	4.4	4.4	12.0	9.5
Marital status, %						
Married	92.9	93.0	89.0	88.9	94.6	93.3
Widowed	4.1	4.0	9.3	9.3	1.7	5.4
Separated/divorced	1.6	1.6	1.5	1.5	1.3	0.5
Never married	1.4	1.5	0.2	0.2	2.4	0.8
Smoking, % [*]						
Never	25.6	25.7	96.8	96.8	35.5	96.8
Former	6.7	6.8	0.4	0.4	12.9	0.8
Current (number of cigarettes or equivalents per day)						
<20	30.3	29.9	2.4	2.4	23.8	1.8
≥20	37.4	37.6	0.5	0.5	27.8	0.6
Alcohol intake, % [†]						
Less than daily	70.7	70.6	98.2	98.2	83.2	98.6
Current daily (g of pure alcohol per day)						
1-29	4.7	4.7	0.6	0.6	4.9	1.0
30-59	6.8	6.8	0.2	0.3	4.6	0.2
≥60	17.8	17.8	1.0	1.0	7.3	0.2
Physical activity, MET-h/d	22.1 (15.3)	21.9 (15.2)	20.4 (12.8)	20.4 (12.7)	16.3 (20.3)	13.7 (16.4)
Food consumption, g/d [‡]						
Fresh fruits	56.9 (31.3)	56.9 (31.3)	63.9 (34.9)	63.8 (34.8)	79.7 (95.1)	92.6 (102.1)
Red meat	44.3 (17.1)	44.3 (17.1)	41.3 (17.3)	41.3 (17.3)	72.5 (76.6)	53.7 (63.9)
Fish/seafood	21.8 (14.1)	21.9 (14.2)	20.6 (14.2)	20.6 (14.2)	21.1 (33.8)	17.3 (31.2)
Body mass index (kg/m ²), %						
<18.5	4.5	4.4	4.3	4.3	3.1	3.0
18.5-23.9	53.8	54.0	50.4	50.6	42.7	46.4
24.0-27.9	32.8	32.7	33.6	33.4	38.4	35.2
≥28	8.9	9.0	11.8	11.7	15.9	15.4
Systolic blood pressure, mmHg	132.8 (20.0)	132.8 (20.0)	129.9 (22.0)	129.9 (22.1)	137.3 (19.9)	136.2 (22.2)
Resting heart rate, beats/minute	77.7 (12.1)	77.8 (12.1)	79.6 (11.6)	79.7 (11.6)	75.1 (11.3)	76.9 (10.6)
Self-rated good health, %						

	CKB population				CNHS population	
	Men		Women		Men	Women
	Derivation	Validation	Derivation	Validation		
Excellent/good	49.1	49.4	43.2	43.6	44.9	41.5
Fair	41.9	41.7	45.4	45.1	46.5	47.8
Poor	8.9	8.9	11.4	11.3	8.6	10.7
Prevalent disease, %						
Stroke	2.3	2.4	1.3	1.3	3.4	3.0
Cancer	0.5	0.5	0.5	0.5	1.6	2.1
COPD	8.8	8.6	6.1	6.2	5.6	3.8
Diabetes	5.6	5.5	6.2	6.1	4.1	4.4

MET-h/d indicates metabolic equivalent task hours per day; COPD, chronic obstructive pulmonary disease.

The baseline characteristics of the CKB population were shown. Data are presented as percentages or means (standard deviations), if applicable.

*Former smoking status was defined as having quit smoking for reasons other than illness.

Participants who had quit smoking due to illness were defined as current smoking status.

†Less than daily group included both never-regular drinkers and current weekly drinkers. Former alcohol drinkers were included in the heavy drinking category (≥ 60 g of pure alcohol per day).

‡The food consumption amount of the CKB population was estimated as described in Supplementary Table 8.

Supplementary Table 2. Definition, type, and levels/range of variables in the prediction model.

Variable	Definition	Type	Levels / range
Age	Age at baseline	Continuous	30-79 years
Highest education	Highest level of received school education	Categorical	5 levels
Marital status	Current marital status	Categorical	4 levels
Smoking	Smoking status	Categorical	4 levels
Alcohol intake	Drinking status	Categorical	4 levels
Physical activity level	Daily total physical activity level during the past 12 months	Continuous	0-59.3 MET-h/d
Fresh fruits	Daily consumption of fresh fruits	Continuous	25-123.4 g/day
Red meat	Daily consumption of red meat	Continuous	11.7-65.2 g/day
Fish/seafood	Daily consumption of fish or seafood	Continuous	3-50.5 g/day
Body mass index	Height (kg)/weight (m) ²	Categorical	4 levels
Systolic blood pressure	Onsite-measured systolic blood pressure	Continuous	94-194.5 mmHg
Resting heart rate	Onsite-measured resting heart rate	Continuous	56-112 beats/min
Self-rated health	Self-rated health status	Categorical	3 levels
Stroke at baseline	Stroke status based on self-report	Categorical	2 levels
Cancer at baseline	Cancer status based on self-report	Categorical	2 levels
COPD at baseline	COPD status based on self-report and pulmonary function test at enrollment	Categorical	2 levels
Diabetes at baseline	Diabetes status based on self-report	Categorical	2 levels

COPD indicates chronic obstructive pulmonary disease. MET-h/d, metabolic equivalent task hours per day.

Supplementary Table 3. Calculation of 5-year all-cause mortality risk.

$$\begin{aligned} \text{Score} = & \beta_{\text{Age_cont}} \times \text{Age_cont} + \beta_{\text{AgeSquared_cont}} \times \text{AgeSquared_cont} + \beta_{\text{Edu_primary_cat}} \\ & \times \text{Edu_primary_cat} + \beta_{\text{Edu_middle_cat}} \times \text{Edu_middle_cat} + \beta_{\text{Edu_high_cat}} \times \text{Edu_high_cat} \\ & + \beta_{\text{Edu_college_cat}} \times \text{Edu_college_cat} + \beta_{\text{Mar_widowed_cat}(M)} \times \text{Mar_widowed_cat}(M) \\ & + \beta_{\text{Mar_separated_cat}(M)} \times \text{Mar_separated_cat}(M) + \beta_{\text{Mar_never_cat}(M)} \times \text{Mar_never_cat}(M) \\ & + \beta_{\text{Smk_former_cat}} \times \text{Smk_former_cat} + \beta_{\text{Smk_dailylt20_cat}} \times \text{Smk_dailylt20_cat} \\ & + \beta_{\text{Smk_dailyge20_cat}} \times \text{Smk_dailyge20_cat} + \beta_{\text{Alc_daily1to29_cat}} \times \text{Alc_daily1to29_cat} \\ & + \beta_{\text{Alc_daily30to59_cat}} \times \text{Alc_daily30to59_cat} + \beta_{\text{Alc_dailyge60_cat}} \times \text{Alc_dailyge60_cat} \\ & + \beta_{\text{Ln_pa_cont}} \times \text{Ln_pa_cont} + \beta_{\text{Fruits_cont}} \times \text{Fruits_cont} + \beta_{\text{Meat_cont}} \times \text{Meat_cont} \\ & + \beta_{\text{Fish_cont}} \times \text{Fish_cont} + \beta_{\text{BMI_low_cat}} \times \text{BMI_low_cat} + \beta_{\text{BMI_overw_cat}} \times \text{BMI_overw_cat} \\ & + \beta_{\text{BMI_obese_cat}} \times \text{BMI_obese_cat} + \beta_{\text{Ln_sbp_cont}} \times \text{Ln_sbp_cont} + \beta_{\text{Heart_rate_cont}} \\ & \times \text{Heart_rate_cont} + \beta_{\text{SRH_fair_cat}} \times \text{SRH_fair_cat} + \beta_{\text{SRH_poor_cat}} \times \text{SRH_poor_cat} \\ & + \beta_{\text{Stroke_cat}} \times \text{Stroke_cat} + \beta_{\text{Cancer_cat}} \times \text{Cancer_cat} + \beta_{\text{COPD_cat}} \times \text{COPD_cat} \\ & + \beta_{\text{Diabetes_cat}} \times \text{Diabetes_cat} \end{aligned}$$

where (M) denotes that the predictor variable was only included in the male model.

Absolute risk of mortality in 5 years:

$$F(5) = 1 - S_0(5)^{\exp(\text{Score})}$$

$S_0(5)$ is the baseline survival probability at 5 years, which is 0.984 in the male model and 0.992 in the female model.

Supplementary Table 4. Description and beta coefficients of the predictor variables in the prediction models for men and women.

Parameter	Description	Beta coefficients	
		Men	Women
Age_cont	Age at baseline (years)	0.07466	0.07276
AgeSquared_cont	Squared age at baseline	0.00031	0.00063
Edu_no_cat	No formal education	Reference	Reference
Edu_primary_cat	Primary school education	-0.13903	-0.10375
Edu_middle_cat	Middle school education	-0.23563	-0.21047
Edu_high_cat	High school education	-0.39169	-0.29393
Edu_college_cat	College or university education	-0.54788	-0.45000
Mar_married_cat	Married	Reference	--
Mar_widowed_cat	Widowed	0.16495	--
Mar_separated_cat	Separated or divorced	0.55843	--
Mar_never_cat	Never married	0.59950	--
Smk_never_cat	Never smoked	Reference	Reference
Smk_former_cat	Former smoking	0.08252	0.08210
Smk_dailylt20_cat	Current smoking, <20 cigarettes or equivalents per day	0.22444	0.28894
Smk_dailyge20_cat	Current smoking, ≥20 cigarettes or equivalents per day	0.27836	0.47761
Alc_nodaily_cat	Non-daily drinker	Reference	Reference
Alc_daily1to29_cat	Daily drinker, 1-29g pure alcohol per day	0.01344	-0.05296
Alc_daily30to59_cat	Daily drinker, 30-59g pure alcohol per day	0.04206	0.10347
Alc_dailyge60_cat	Daily drinker, ≥60g pure alcohol per day	0.23151	0.12238
Ln_pa_cont	Log transformation of total physical activity level ()	-0.08969	-0.18258
Fruits_cont	Daily fresh fruits intake (10g/d)	-0.12426	-0.20809
Meat_cont	Daily red meat intake (10g/d)	-0.27162	-0.34874
Fish_cont	Daily fish/seafood intake (10g/d)	-0.35292	-0.61505
BMI_low_cat	Underweight, BMI <18.5 kg/m ²	0.43572	0.43791
BMI_normal_cat	Normal weight, BMI 18.5-23.9 kg/m ²	Reference	Reference
BMI_overw_cat	Overweight, BMI 24.0-27.9 kg/m ²	-0.15787	-0.09857
BMI_obese_cat	Obese, BMI ≥28 kg/m ²	-0.14856	-0.07868
Ln_sbp_cont	Log transformation of systolic blood pressure (MET-h/d)	1.26924	1.32205
Heart_rate_cont	Resting heart rate (beats/minute)	0.00875	0.00565
SRH_good_cat	Excellent or good self-rated health	Reference	Reference
SRH_fair_cat	Fair self-rated health	0.18767	0.15813
SRH_poor_cat	Poor self-rated health	0.57303	0.49777
Stroke_cat	Has stroke at baseline	0.47988	0.47740
Cancer_cat	Has cancer at baseline	0.75013	0.77251
COPD_cat	Has COPD at baseline	0.27701	0.33181
Diabetes_cat	Has diabetes at baseline	0.39295	0.57525

BMI indicates body mass index; COPD, chronic obstructive pulmonary disease; MET-h/d, metabolic equivalent task hours per day.

Continuous predictor variables were centered as that in Table 1.

Supplementary Table 5. HRs (95% CIs) of all-cause mortality for all predictor variables in the whole China Kadoorie Biobank population for men and women.

	Men (n=210,203)			Women (n=302,518)		
	Deaths	Deaths/PYs (/1,000)	HRs (95% CIs)	Deaths	Deaths/PYs (/1,000)	HRs (95% CIs)
Age (years)						
Linear	--	--	1.08 (1.08-1.08)	--	--	1.08 (1.07-1.08)
Squared	--	--	1.00 (1.00-1.00)	--	--	1.00 (1.00-1.00)
Highest education						
No formal school	5,797	28.9	1.00 (Reference)	11,302	12.6	1.00 (Reference)
Primary school	14,154	18.0	0.86 (0.83-0.89)	8,125	7.2	0.91 (0.88-0.94)
Middle school	7,105	8.9	0.80 (0.77-0.83)	3,163	3.4	0.82 (0.79-0.86)
High school	3,307	7.7	0.68 (0.65-0.72)	1,510	3.1	0.77 (0.72-0.81)
College/university	1,593	8.3	0.57 (0.54-0.61)	493	3.1	0.69 (0.63-0.77)
Marital status						
Married	27,340	12.1	1.00 (Reference)	--	--	--
Widowed	3,277	38.1	1.19 (1.15-1.24)	--	--	--
Separated/divorced	594	16.0	1.71 (1.58-1.86)	--	--	--
Never married	745	22.6	1.81 (1.68-1.95)	--	--	--
Smoking*						
Never	7,007	11.3	1.00 (Reference)	22,364	6.4	1.00 (Reference)
Former	2,333	14.6	1.06 (1.01-1.11)	281	21.2	1.16 (1.03-1.30)
Current (number of cigarettes or equivalents per day)						
<20	11,214	15.6	1.25 (1.21-1.29)	1,621	20.3	1.34 (1.26-1.41)
≥20	11,402	12.5	1.33 (1.29-1.38)	327	20.5	1.60 (1.43-1.79)
Alcohol intake†						
Less than daily	20,610	12.0	1.00 (Reference)	23,832	6.7	1.00 (Reference)
Current daily (g of pure alcohol per day)						
1-29	1,657	14.7	1.00 (0.95-1.05)	204	9.8	0.96 (0.83-1.10)
30-59	2,129	12.8	1.06 (1.01-1.11)	89	10.8	1.04 (0.84-1.29)
≥60	7,560	18.0	1.23 (1.20-1.27)	468	13.3	1.10 (1.00-1.21)
Ln(physical activity level [MET-h/day])	--	--	0.91 (0.91-0.92)	--	--	0.83 (0.82-0.84)
Food consumption (per 10g/d)						
Fresh fruits	--	--	0.99 (0.98-0.99)	--	--	0.98 (0.98-0.99)
Red meat	--	--	0.97 (0.96-0.98)	--	--	0.97 (0.96-0.98)
Fish/seafood	--	--	0.96 (0.95-0.98)	--	--	0.94 (0.92-0.95)
Body mass index (kg/m ²)						
<18.5	3,082	31.7	1.52 (1.47-1.59)	2,169	14.8	1.52 (1.45-1.60)
18.5-23.9	17,793	13.7	1.00 (Reference)	11,400	6.3	1.00 (Reference)
24.0-27.9	8,668	10.8	0.85 (0.83-0.88)	7,609	6.3	0.90 (0.87-0.92)
≥28	2,413	11.1	0.86 (0.82-0.90)	3,415	8.1	0.93 (0.90-0.97)
Ln(systolic blood pressure [mmHg])	--	--	3.52 (3.26-3.79)	--	--	3.75 (3.45-4.07)
Resting heart rate (beats/minute)	--	--	1.01 (1.01-1.01)	--	--	1.01 (1.00-1.01)

	Men (n=210,203)			Women (n=302,518)		
	Deaths	Deaths/PYs (/1,000)	HRs (95% CIs)	Deaths	Deaths/PYs (/1,000)	HRs (95% CIs)
Self-rated health						
Excellent/good	11,711	9.6	1.00 (Reference)	7,353	4.6	1.00 (Reference)
Fair	14,626	14.6	1.18 (1.15-1.21)	11,564	7.1	1.17 (1.13-1.20)
Poor	5,619	28.5	1.74 (1.68-1.80)	5,676	14.4	1.64 (1.58-1.70)
Stroke at baseline						
No	29,702	12.6	1.00 (Reference)	23,351	6.6	1.00 (Reference)
Yes	2,254	48.5	1.62 (1.54-1.69)	1,242	29.9	1.61 (1.52-1.71)
Cancer at baseline						
No	31,499	13.1	1.00 (Reference)	24,196	6.8	1.00 (Reference)
Yes	457	53.4	2.18 (1.98-2.39)	397	23.1	2.13 (1.92-2.35)
COPD at baseline						
No	25,825	11.7	1.00 (Reference)	21,071	6.2	1.00 (Reference)
Yes	6,131	31.5	1.33 (1.29-1.37)	3,522	16.4	1.39 (1.34-1.44)
Diabetes at baseline						
No	28,536	12.5	1.00 (Reference)	20,526	6.1	1.00 (Reference)
Yes	3,420	27.7	1.52 (1.46-1.58)	4,067	19.8	1.78 (1.72-1.85)

PYs indicates person-years; HR, hazard ratio; CI, confidence interval; MET-h/d, metabolic equivalent task hours per day; COPD, chronic obstructive pulmonary disease; -, not applicable.

All predictor variables were included simultaneously in the model. Continuous predictor variables were centered as follows: age at 50 years, physical activity level at 20 MET-h/d, daily fresh fruits intake at 80g, daily red meat intake at 50g, daily fish/seafood intake at 20g, systolic blood pressure at 120 mmHg, and resting heart rate at 80 beats/minute.

*Former smoking status was defined as having stopped smoking for reasons other than illness.

Participants who had stopped smoking due to illness were classified as having a current smoking status.

†Less than daily group included both never-regular drinkers and current weekly drinkers. Former alcohol drinkers were included in the heavy drinking category (≥ 60 g of pure alcohol per day).

Supplementary Table 6. HRs (95% CIs) of all-cause mortality for predictor variables in multiple models for men and women.

	Men (n=140,135)			Women (n=201,678)		
	Model 1 Age + lifestyle factors	Model 2 Model 1 + Sociodemographic factors	Full model Model 2 + health indicators	Model 1 Age + lifestyle factors	Model 2 Model 1 + Sociodemographic factors	Full model Model 2 + health indicators
Age (years)						
Linear	1.09 (1.09, 1.09)	1.09 (1.09, 1.09)	1.08 (1.08, 1.08)	1.10 (1.09, 1.10)	1.09 (1.09, 1.10)	1.08 (1.07, 1.08)
Squared	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)	1.00 (1.00, 1.00)
Smoking						
Never	1.00 (Reference)	1.00 (Reference)	1.00 (Reference)	1.00 (Reference)	1.00 (Reference)	1.00 (Reference)
Former	1.05 (1.00, 1.12)	1.05 (0.99, 1.12)	1.07 (1.01, 1.13)	1.12 (0.97, 1.30)	1.09 (0.94, 1.26)	1.12 (0.97, 1.30)
Current (number of cigarettes or equivalents per day)						
<20	1.31 (1.26, 1.36)	1.28 (1.24, 1.33)	1.24 (1.19, 1.29)	1.34 (1.25, 1.43)	1.31 (1.22, 1.40)	1.31 (1.22, 1.40)
≥20	1.41 (1.35, 1.46)	1.38 (1.33, 1.44)	1.34 (1.29, 1.39)	1.55 (1.35, 1.78)	1.52 (1.32, 1.74)	1.53 (1.34, 1.76)
Alcohol intake						
Less than daily	1.00 (Reference)	1.00 (Reference)	1.00 (Reference)	1.00 (Reference)	1.00 (Reference)	1.00 (Reference)
Current daily (g of pure alcohol per day)						
1-29	1.00 (0.94, 1.06)	0.99 (0.93, 1.06)	1.02 (0.95, 1.08)	0.83 (0.70, 0.98)	0.82 (0.69, 0.98)	0.92 (0.77, 1.09)
30-59	1.06 (1.00, 1.12)	1.04 (0.99, 1.10)	1.05 (0.99, 1.11)	0.97 (0.75, 1.27)	0.96 (0.74, 1.25)	1.01 (0.77, 1.31)
≥60	1.36 (1.32, 1.41)	1.36 (1.31, 1.41)	1.25 (1.21, 1.30)	1.18 (1.05, 1.32)	1.18 (1.05, 1.32)	1.14 (1.02, 1.28)
Ln(physical activity level [MET-h/day])	0.87 (0.86, 0.88)	0.87 (0.86, 0.88)	0.91 (0.90, 0.92)	0.76 (0.75, 0.77)	0.76 (0.75, 0.77)	0.82 (0.81, 0.84)xiyan
Food consumption (per 10g/d)						
Fresh fruits	0.97 (0.96, 0.97)	0.98 (0.97, 0.98)	0.98 (0.98, 0.99)	0.96 (0.96, 0.97)	0.97 (0.96, 0.97)	0.98 (0.97, 0.99)
Red meat	0.95 (0.94, 0.96)	0.96 (0.95, 0.97)	0.97 (0.96, 0.98)	0.97 (0.95, 0.98)	0.97 (0.96, 0.98)	0.97 (0.96, 0.98)

	Men (n=140,135)			Women (n=201,678)		
	Model 1 Age + lifestyle factors	Model 2 Model 1 + Sociodemographic factors	Full model Model 2 + health indicators	Model 1 Age + lifestyle factors	Model 2 Model 1 + Sociodemographic factors	Full model Model 2 + health indicators
Fish/seafood	0.93 (0.91, 0.95)	0.95 (0.93, 0.97)	0.96 (0.94, 0.98)	0.92 (0.90, 0.94)	0.93 (0.91, 0.95)	0.93 (0.92, 0.95)
Highest education						
No formal school	--	1.00 (Reference)	1.00 (Reference)	--	1.00 (Reference)	1.00 (Reference)
Primary school	--	0.86 (0.82, 0.89)	0.86 (0.83, 0.90)	--	0.91 (0.87, 0.95)	0.91 (0.87, 0.95)
Middle school	--	0.79 (0.76, 0.83)	0.79 (0.75, 0.83)	--	0.80 (0.75, 0.85)	0.80 (0.76, 0.85)
High school	--	0.67 (0.63, 0.71)	0.67 (0.63, 0.71)	--	0.73 (0.67, 0.78)	0.74 (0.69, 0.80)
College/university	--	0.57 (0.53, 0.61)	0.59 (0.54, 0.63)	--	0.67 (0.59, 0.75)	0.69 (0.61, 0.78)
Marital status						
Married	--	1.00 (Reference)	1.00 (Reference)	--	--	--
Widowed	--	1.22 (1.16, 1.28)	1.19 (1.14, 1.25)	--	--	--
Separated/divorced	--	1.75 (1.58, 1.93)	1.65 (1.50, 1.83)	--	--	--
Never married	--	2.05 (1.87, 2.24)	1.80 (1.64, 1.97)	--	--	--
Body mass index (kg/m ²)						
<18.5	--	--	1.55 (1.48, 1.62)	--	--	1.53 (1.44, 1.62)
18.5-23.9	--	--	1.00 (Reference)	--	--	1.00 (Reference)
24.0-27.9	--	--	0.85 (0.82, 0.88)	--	--	0.89 (0.86, 0.93)
≥28	--	--	0.87 (0.82, 0.92)	--	--	0.93 (0.89, 0.98)
Ln(systolic blood pressure [mmHg])	--	--	3.41 (3.12, 3.74)	--	--	3.73 (3.37, 4.12)
Resting heart rate (beats/minute)	--	--	1.01 (1.01, 1.01)	--	--	1.01 (1.01, 1.01)
Self-rated health						

	Men (n=140,135)			Women (n=201,678)		
	Model 1 Age + lifestyle factors	Model 2 Model 1 + Sociodemographic factors	Full model Model 2 + health indicators	Model 1 Age + lifestyle factors	Model 2 Model 1 + Sociodemographic factors	Full model Model 2 + health indicators
Excellent/good	--	--	1.00 (Reference)	--	--	1.00 (Reference)
Fair	--	--	1.17 (1.13, 1.21)	--	--	1.17 (1.12, 1.21)
Poor	--	--	1.75 (1.67, 1.82)	--	--	1.64 (1.56, 1.71)
Stroke at baseline						
No	--	--	1.00 (Reference)	--	--	1.00 (Reference)
Yes	--	--	1.59 (1.50, 1.68)	--	--	1.60 (1.49, 1.72)
Cancer at baseline						
No	--	--	1.00 (Reference)	--	--	1.00 (Reference)
Yes	--	--	2.10 (1.87, 2.37)	--	--	2.04 (1.80, 2.31)
COPD at baseline						
No	--	--	1.00 (Reference)	--	--	1.00 (Reference)
Yes	--	--	1.34 (1.29, 1.39)	--	--	1.38 (1.32, 1.45)
Diabetes at baseline						
No	--	--	1.00 (Reference)	--	--	1.00 (Reference)
Yes	--	--	1.51 (1.44, 1.58)	--	--	1.79 (1.71, 1.87)

HR indicates hazard ratio; CI, confidence interval; MET-h/d, metabolic equivalent task hours per day; COPD, chronic obstructive pulmonary disease; --, not applicable.

The definition and classification of predictor variables were the same as in Supplementary Table 5.

Supplementary Table 7. Life expectancy at 30 years and the gained life years under different scenarios for smoking and alcohol intake in men and women.

	Men		Women	
	Life expectancy (95% CI), 30y	Gained life years (95% CI)	Life expectancy (95% CI), 30y	Gained life years (95% CI)
Base scenario*	45.46 (44.85-46.27)	--	47.00 (46.23-47.86)	--
Smoking				
Ideal scenario	46.13 (45.51-47.00)	0.68 (0.44-0.91)	47.05 (46.30-47.90)	0.05 (0.00-0.09)
Practical scenario	45.86 (45.27-46.70)	0.40 (0.26-0.54)	--	--
Alternative scenario [†]	46.47 (45.83-47.24)	1.02 (0.88-1.17)	47.11 (46.34-47.95)	0.10 (0.07-0.14)
Healthy China 2030 [‡]	45.60 (45.03-46.42)	0.14 (0.09-0.19)	--	--
Alcohol intake				
Ideal scenario	45.61 (44.98-46.44)	0.15 (0.08-0.23)	47.01 (46.23-47.87)	0.01 (0.00-0.02)
Alternative scenario [§]	45.63 (45.02-46.44)	0.17 (0.13-0.20)	47.00 (46.23-47.86)	0.00 (0.00-0.01)
Physical activity				
Ideal scenario	46.18 (45.58-46.95)	0.72 (0.62-0.84)	48.47 (47.78-49.23)	1.47 (1.30-1.67)
Practical scenario	45.99 (45.38-46.77)	0.53 (0.46-0.62)	48.09 (47.38-48.88)	1.09 (0.96-1.25)
Healthy China 2030	45.63 (45.02-46.43)	0.17 (0.15-0.20)	47.39 (46.63-48.22)	0.38 (0.34-0.43)

CI indicates confidence interval.

*Exposure patterns for all lifestyle factors remained unchanged.

[†]The whole population is assumed to never smoke.

[‡]The goal is to reduce smoking prevalence in the whole population to 20%, and we assumed that this would be achieved only by reducing smoking prevalence in men.

[§]All people with excessive alcohol use in the whole population are assumed to be non-daily drinkers.

^{||}The goal is to increase the proportion of individuals who engage in regular physical activity to 40%. Regular physical activity was defined as engaging in physical activity of moderate intensity or higher ≥ 3 times per week, each time lasting 30 minutes or more.

Supplementary Table 8. Calculation of the usual amount (g/day) for each baseline category of fresh fruits consumption based on 19,786 participants who attended the first resurvey in 2008.

		1 st resurvey (mean 2.6 years later)								
Baseline		F					Consumption days per month	Mean daily grams*	Baseline consumption (g/day) [†] B	Usual consumption (g/day) [‡] U
n		Never/rarely 1	Monthly 2	1-3 days/wk 3	4-6 days/wk 4	Daily 5				
1	Never/rarely	17.9% (244)	55.7% (761)	18.0% (246)	4.2% (58)	4.2% (57)	0	--	B ₁ = 0	U ₁ = 25.0
2	Monthly	7.4% (496)	52.1% (3,504)	29.0% (1,952)	6.8% (460)	4.7% (316)	2.5	137.0	B ₂ = 11.4	U ₂ = 32.6
3	1-3 days/wk	4.0% (250)	31.7% (1,981)	37.4% (2,339)	11.7% (729)	15.2% (950)	8.6	142.5	B ₃ = 40.8	U ₃ = 56.2
4	4-6 days/wk	1.6% (29)	20.1% (368)	35.0% (641)	15.0% (275)	28.2% (516)	21.5	140.3	B ₄ = 100.5	U ₄ = 79.3
5	Daily	1.2% (45)	6.8% (246)	18.3% (662)	13.2% (476)	60.5% (2,185)	30	168.6	B ₅ = 168.6	U ₅ = 123.4

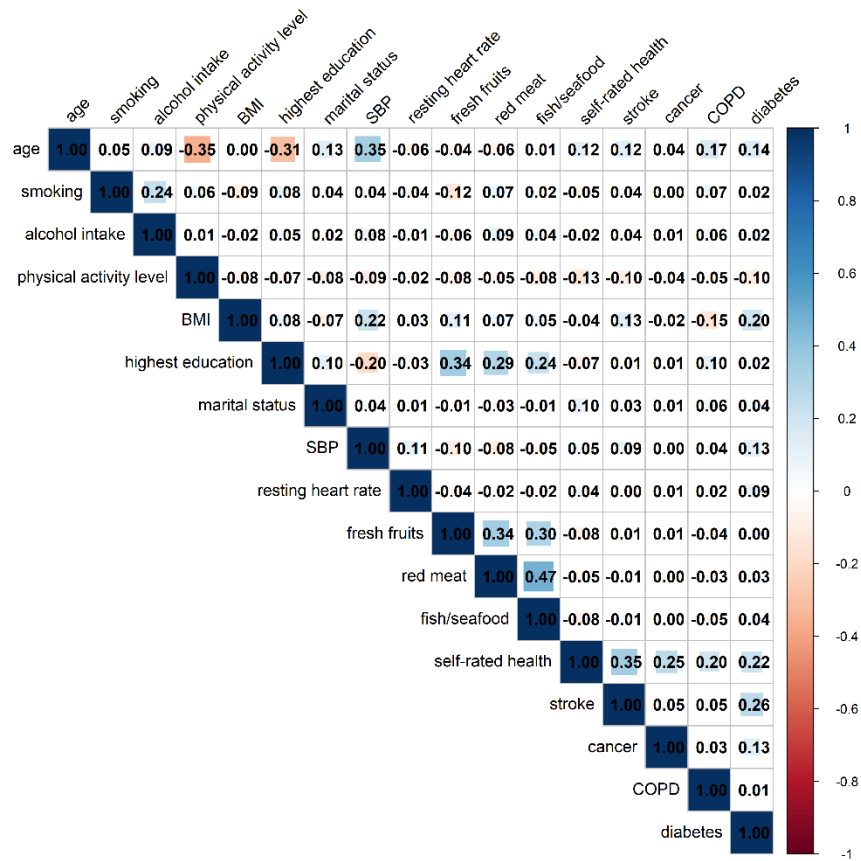
*The mean daily grams were calculated using data from the 2nd resurvey (n=25,041) and used as a proxy for baseline mean daily grams, assuming that the daily amount of fresh fruits consumption did not change significantly from the baseline to the 2nd resurvey (i.e., people may change how frequently they consume fruits but the daily amount of fresh fruits consumption stays roughly the same).

[†]The baseline intake amount for each group was estimated by multiplying the mean daily grams intake by the number of consuming days per month and dividing by 30.

[‡]By accounting for the variation in consumption frequency from the baseline, the usual intake amount for each group was estimated using the following formula: $U_n = \sum_{i=1}^5 F_{(n,i)} \times B_i$. F denotes the percentage in each cell, B the baseline grams per day for each baseline category, and U the usual grams per day for each baseline category.

Supplementary Table 9. Usual intake amount for each baseline group of fresh vegetables, red meat, and seafood.

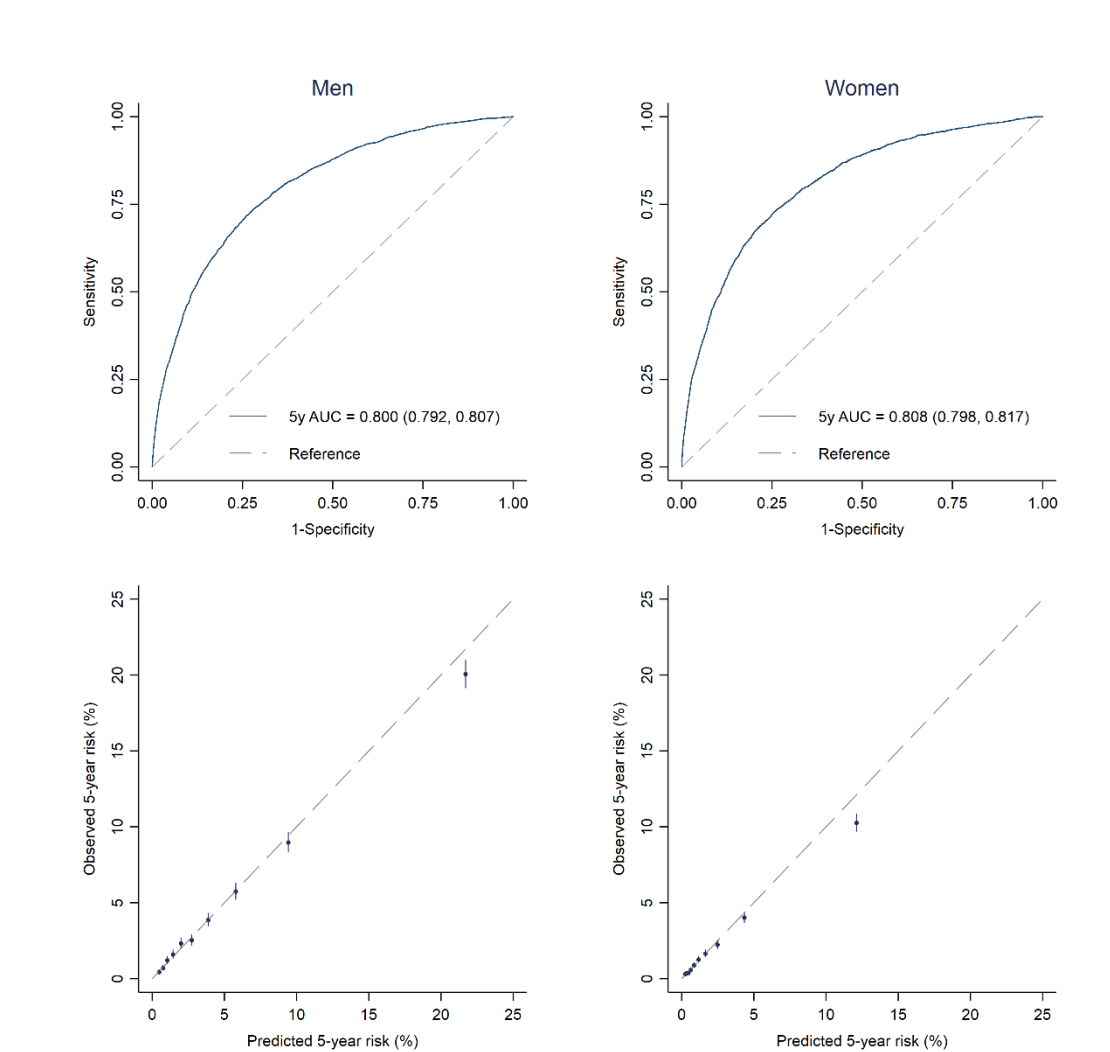
	Usual intake amount (g/day)		
	Fresh vegetables	Red meat	Seafood
Never/rarely	194.3	11.7	3.0
Monthly	204.8	17.9	19.7
1-3 days/wk	215.3	34.5	31.3
4-6 days/wk	210.6	47.0	44.4
Daily	234.1	65.2	50.5



Supplementary Figure 1. Correlation matrix between the predictor variables (n=512,721).

BMI indicates body mass index; SBP, systolic blood pressure; COPD, chronic obstructive pulmonary disease.

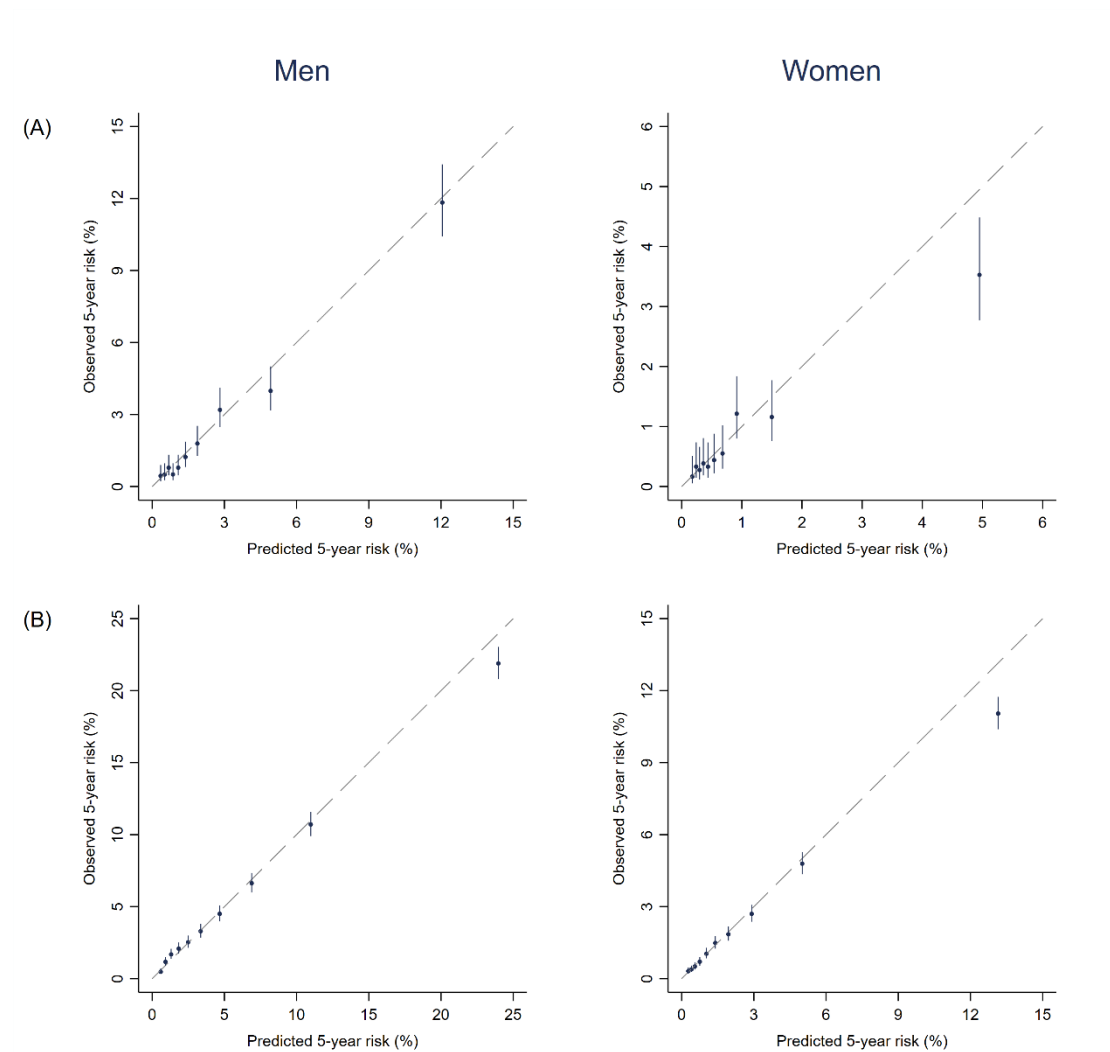
We evaluated correlations between different types of variables using appropriate coefficients: Pearson correlation for continuous variables, point-biserial correlation between continuous and categorical variables, polychoric correlation for ordinal variables, Cramer's V value for nominal variables, and tetrachoric correlation for binary variables. Negatively correlated variables were shown in blue, while those positively correlated were colored in red. The size of the box and the shade of the color were defined by the correlation strength. Source data are provided as a Source Data file.



Supplementary Figure 2. Receiver operating characteristic (ROC) curves and calibration plots of the 5-year all-cause mortality prediction model in the validation dataset of China Kadoorie Biobank for men (n=70,068) and women (n=100,840).

The horizontal axis of the ROC curve represents the false positive rate (1-specificity) at various probability cut-offs, and the vertical axis indicates the true positive rate (sensitivity). AUC indicates the area under the ROC curve, with higher values indicating better discrimination between participants at high and low risk of death.

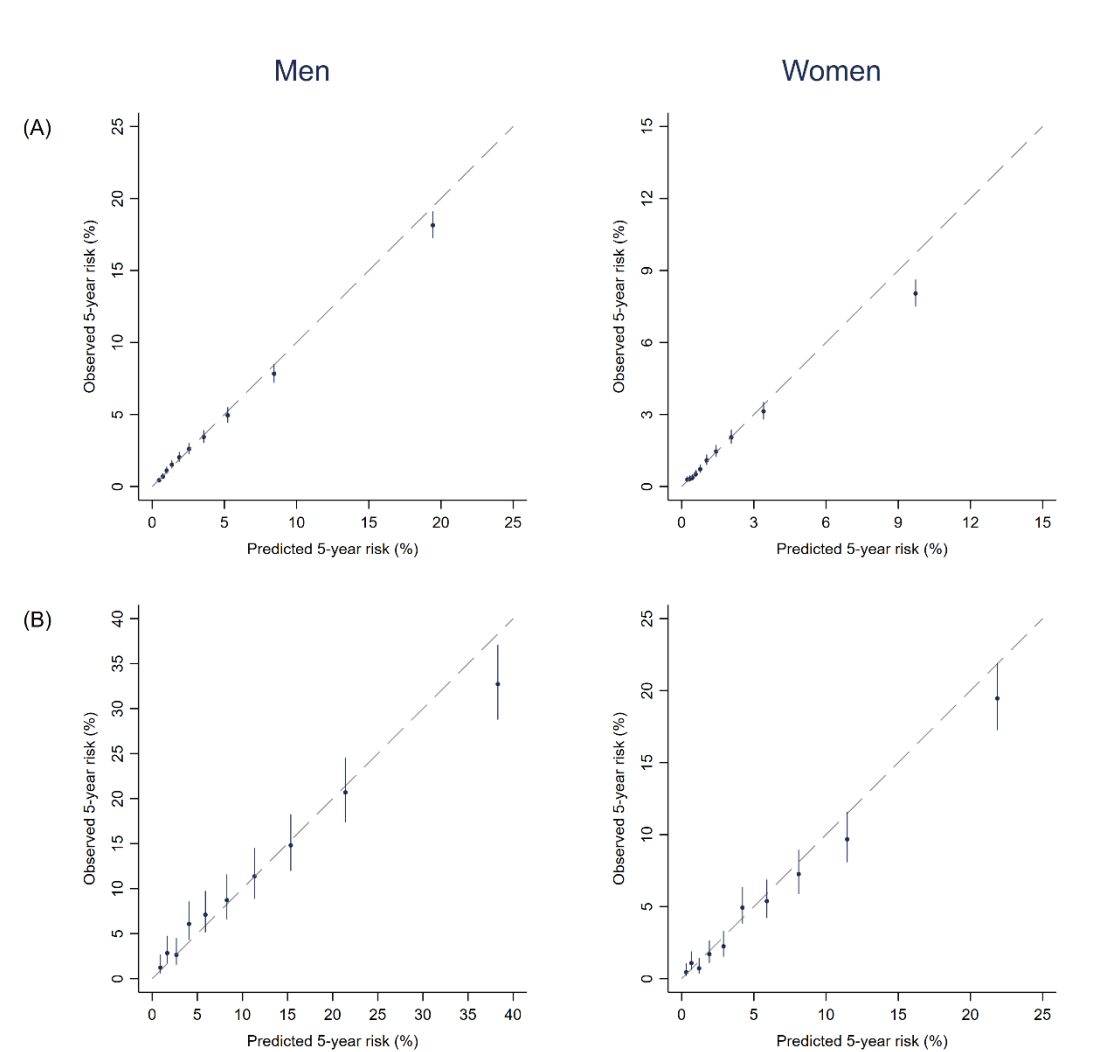
The observed 5-year risk was estimated using Kaplan-Meier and plotted against the model-predicted risk by decile. The error bars are the 95% confidence intervals of the observed risk. Source data are provided as a Source Data file.



Supplementary Figure 3. Calibration plots of the 5-year all-cause mortality prediction model in the validation dataset of China Kadoorie Biobank by education level for men (n=70,068) and women (n=100,840).

(A) Middle school and higher; (B) No education and primary school.

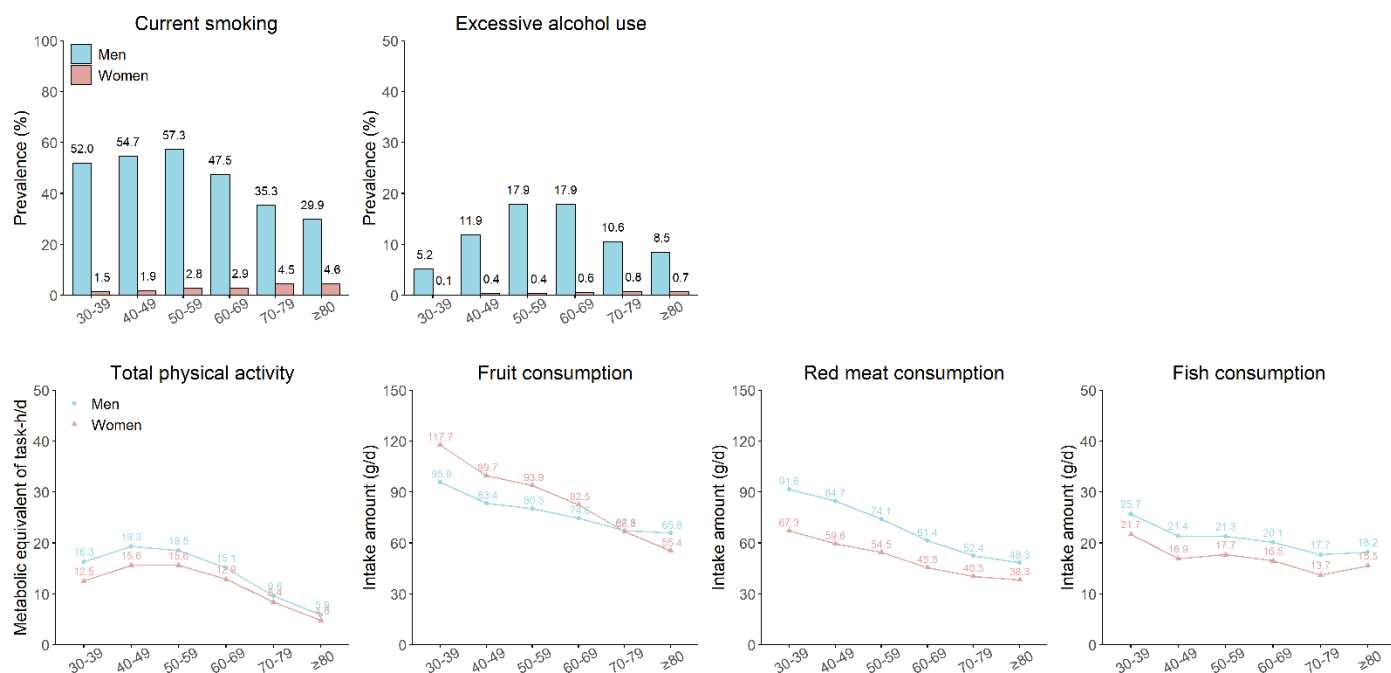
The observed 5-year risk was estimated using Kaplan-Meier and plotted against the model-predicted risk by decile. The error bars are the 95% confidence intervals of the observed risk. Source data are provided as a Source Data file.



Supplementary Figure 4. Calibration plots of the 5-year all-cause mortality prediction model in the validation dataset of China Kadoorie Biobank by marital status for men (n=70,068) and women (n=100,840).

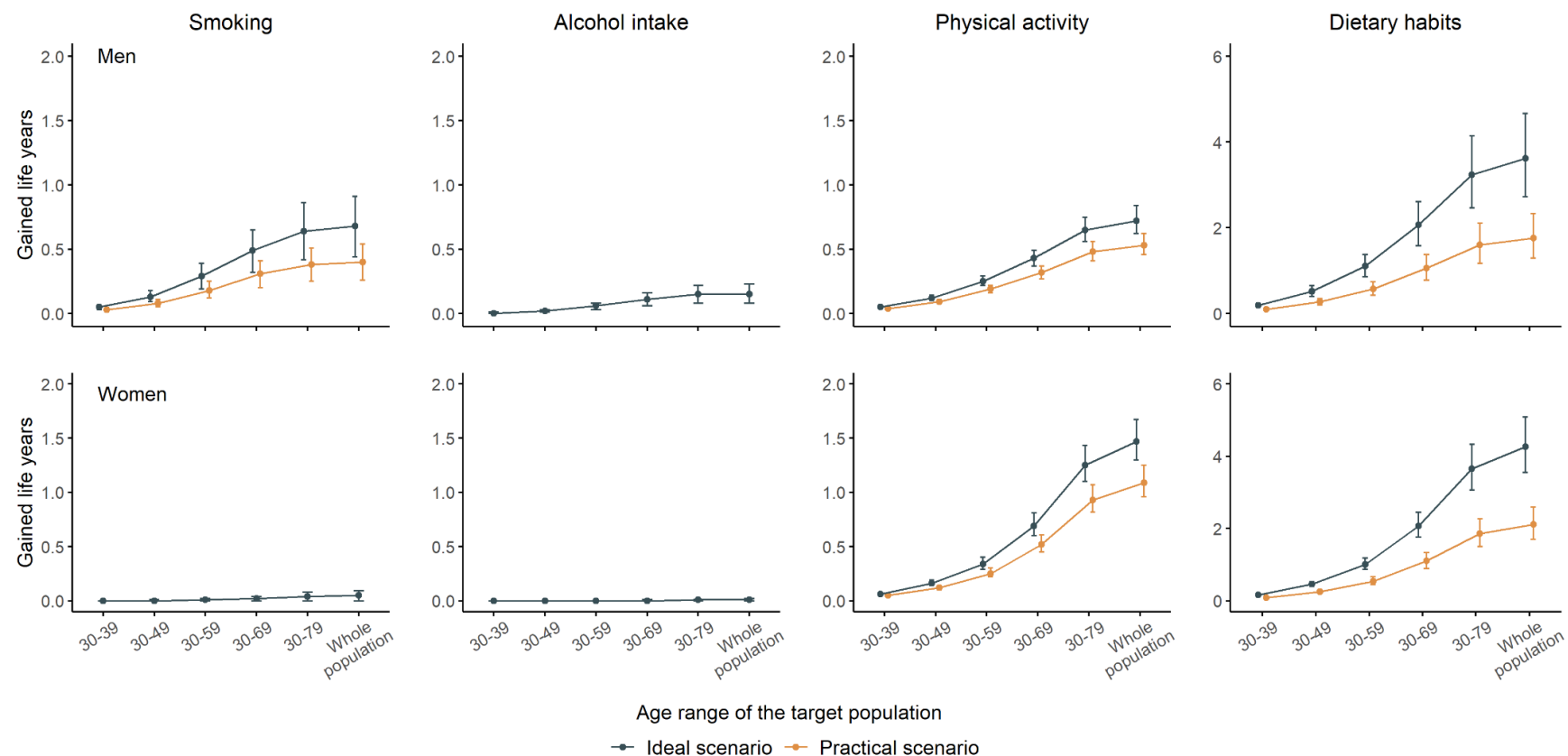
(A) Married status; (B) Non-married status.

The observed 5-year risk was estimated using Kaplan-Meier and plotted against the model-predicted risk by decile. The error bars are the 95% confidence intervals of the observed risk. Source data are provided as a Source Data file.



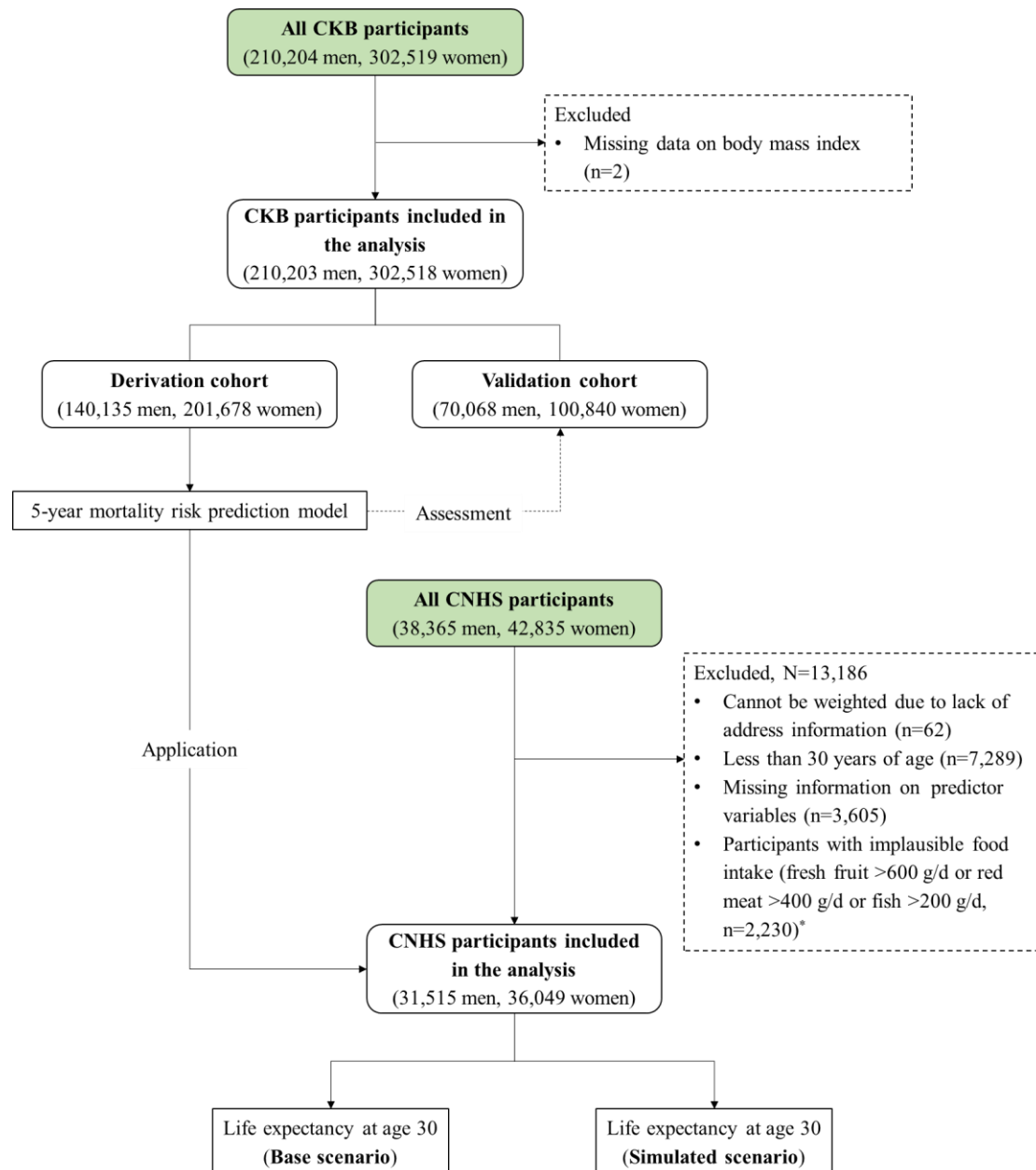
Supplementary Figure 5. Age-specific distribution of lifestyle factors in China Nutrition and Health Surveillance for men (n=31,515, blue) and women (n=36,049, red).

Excessive alcohol use was defined as consuming ≥ 30 g of pure alcohol per day. Source data are provided as a Source Data file.



Supplementary Figure 6. Gained life years at age 30 for men (n=31,515) and women (n=36,049) from adopting individual low-risk lifestyle habits, with the target population having different age ranges.

The definitions of ideal (dark blue) and practical (dark yellow) scenarios for each lifestyle factor are the same as in Table 3. Data are presented as point estimates of gained life years (centers of error bars) and the corresponding 95% confidence limits (error bars). Source data are provided as a Source Data file.



Supplementary Figure 7. Analytic flow chart of the primary analysis.

CKB indicates China Kadoorie Biobank; CNHS, China Nutrition and Health Surveillance.

*The cut-off values were approximately equal to the 99th percentile values.