International Journal of Epidemiology 2002;**31**:847–854

NUTRITION

A prospective study of variety of healthy foods and mortality in women

Karin B Michels[a,b] and Alicja Wolk[b]

Background To assess the overall influence of diet on health and disease in epidemiological studies, the habitual diet of the study participants has to be captured as a pattern rather than individual foods or nutrients. The simplest way to describe dietary preferences is to separate foods considered beneficial to health from foods considered to promote disease, and separate individuals on the basis of their regular consumption of these foods.

Methods We used data from 59 038 women participating in the prospective Mammography Screening Cohort in Sweden to investigate the influence of variety of healthy and less healthy foods on all-cause and cause-specific mortality.

Results Women who followed a healthy diet defined as consumption of a high variety of fruits, vegetables, whole grain breads, cereals, fish, and low fat dairy products had a significantly lower mortality than women who consumed few of these foods (3710 deaths total). Women who reported regularly consuming 16–17 healthy foods had a 42% lower all-cause mortality (95% CI: 32–50%) compared to women reporting consumption of 0–8 healthy foods with any regularity (*P* for trend <0.0001). For each additional healthy food consumed the risk of death was about 5% lower (95% CI: 4–6%). Cardiovascular mortality was particularly low among women who reported consuming a high variety of healthy foods. A less healthy diet defined as consumption of a high variety of red meats, refined carbohydrates and sugars, and foods high in saturated or trans fats was not directly associated with a higher overall mortality. However, women who reported consuming many less healthy foods were significantly more likely to die from cancer than those who consumed few less healthy foods.

Conclusions A healthy diet can affect longevity. It appears more important to increase the number of healthy foods regularly consumed than to reduce the number of less healthy foods regularly consumed.

Keywords Dietary pattern, nutrition, mortality, cohort study

Accepted 8 February 2002

Epidemiological studies investigating the association of diet and health have mostly related individual foods or nutrients to health outcomes. This approach is confounded by other foods or nutrients not studied as no food or nutrient is consumed in isolation. Dietary patterns that represent preferences for commonly consumed foods are considerably more interesting as they consider diet in its entirety. Principal components analysis[1–4] and cluster analysis[5–9] have been employed to derive dietary patterns. Similarly, a variety of dietary indices have been proposed to assess overall diet quality.[10] A simple way to define dietary patterns is to separate 'good' from 'bad' foods to describe a 'healthy' diet and a 'less healthy' diet based on current knowledge and current dietary guidelines. Kant *et al.* have recently reported the use of a Recommended Foods Score (RFS) and found it to be highly predictive of mortality.[11] We have developed a similar RFS and have complemented it with a Not Recommended Foods Score (NRFS). We have related these dietary patterns prospectively to overall mortality during nearly 10 years of follow-up using dietary data collected from 59 038 women in the Swedish Mammography Screening Cohort.

[a] Obstetrics & Gynecology Epidemiology Center, Brigham & Women's Hospital, Harvard Medical School, and Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA.

[b] Department of Medical Epidemiology, Karolinska Institutet, Stockholm, Sweden.

Correspondence: Dr Karin B Michels, Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA. E-mail: kmichels@rics.bwh.harvard.edu

Moreover, we have studied the associations with cause-specific mortality from cancer, coronary heart disease (CHD), and stroke.

Methods

Swedish Mammography Screening Cohort

A population-based mammography screening programme was introduced in two counties (Västmanland and Uppsala) in central Sweden from 1987 to 1990.[12,13] In Västmanland county all women born between 1917 and 1948 were invited by mail to participate in a mammogram screening between March 1987 and March 1989 (n = 41 786). Enclosed with this invitation was a 6-page questionnaire on demographic and lifestyle factors and a 60-item food frequency questionnaire (FFQ); 31 735 women (76%) returned completed questionnaires. In Uppsala county all women born between 1914 and 1948 were invited to the screening and were mailed the same questionnaire between January 1988 and December 1990 (n = 48 517); 34 916 women (72%) returned completed questionnaires. In total, questionnaires completed before mammography were available from 66 651 women.

Women who did not fall within the age range 40–76 years at mammography (n = 165), women with missing (n = 707) or incorrect national identification numbers (n = 415), women with missing return date of the questionnaire (n = 608), those who moved out of the study area at an unknown date (n = 79), and those who had died during follow-up but for whom date of death was missing (n = 16) were excluded from the study population. We also excluded women with self-reported energy intake estimates below or above three standard deviations of the mean $\log_e$-transformed calories (below 417 or above 3729 kcal) from this analysis (n = 793). Women with a previous cancer diagnosis other than non-melanoma skin cancer at baseline were identified by linkage to the Swedish Cancer Registry and excluded from the cohort (n = 2399). A further 2431 women with missing or unreasonable self-reported values for height (<100 cm or >270 cm), weight (<30 kg or >300 kg), body mass index (<15 or >50 kg/m^2), and age at first birth (⩽15 or ⩾60 years of age) were excluded from this analysis, leaving a study cohort of 59 038 women.

The study was approved by the Ethical Committee at Uppsala University Hospital and by the Karolinska Institute's Regional Ethical Committee.

Dietary assessment

Diet was assessed using a self-administered semi-quantitative FFQ[14] including 60 foods commonly consumed in Sweden. The FFQ included questions on the regular consumption of butter, margarine, cheese, milk, yoghurt, bread, potatoes, vegetables (root, cabbage, tomatoes, lettuce/cucumber, spinach/kale), fruit (apples/pears, citrus fruit, banana), juice, oats, muesli, pancakes/Belgian waffles, rice, pasta, beans/peas, meat, bacon, sausage, liver, poultry, fish, eggs, chips, cookies, ice cream, jam, soft drinks, candy, chocolate, added sugar, coffee, tea, beer, wine, and hard liquor. We asked how often, on average, over the past 6 months, the participants had consumed these foods. Eight categories for frequency of consumption were prespecified: never/seldom, 1–3 times/month, once a week, 2–3 times/week, 4–6 times/week, once a day, 2–3 times/day, ⩾4 times/day. While the FFQ only assessed average consumption during the 6 months prior to baseline, dietary preferences tend to be fairly stable and dietary patterns are unlikely to have changed in this cohort of women who were healthy at baseline.

The validity of nutrient estimates based on the self-reported food frequencies was evaluated in a subsample of 129 women from this cohort. During four 7-day periods, 3–4 months apart, each of these 129 participants weighed and recorded all foods consumed. The validity of energy-adjusted fat intake from the FFQ (assessed by Pearson correlation coefficients between the FFQ and weighed food record-derived estimates) was as follows: total fat r = 0.5, monounsaturated fat r = 0.5, and total polyunsaturated fat r = 0.4. After completion of diet recording, a subcutaneous adipose tissue sample was taken from each participant by needle aspiration. The validity of polyunsaturated fat estimated from the food questionnaire (as per cent of total fatty acids) in comparison to adipose tissue composition was r = 0.5. Compared with the food records, the FFQ-based intakes were underestimated on average by 12% for saturated and monounsaturated fat and by 18% for polyunsaturated fat.

Food scores

Dietary patterns were defined as RFS and NRFS by simply separating 'good' from 'bad' foods based on their nutrient content and largely following dietary guidelines[15] and results from large epidemiological studies.[16] The RFS included the following foods: apples/pears, citrus fruit, bananas, lettuce/cucumber, spinach/kale, tomatoes, cabbage, root vegetables (carrots, beets, etc.), beans/peas, milk with 0.5% or 1.5% fat, yoghurt with 1.5% fat, whole grain bread, crisp bread (high fibre content, no fat), oats, salmon/herring/tuna, other fish (excluding shell fish). The NRFS included meat, meat stew, minced meat, bacon, sausages, blood pudding/sausages, cold cuts, pate, liver/kidney, fried potatoes, French fries, chips, cheese (high in saturated fat), butter (high in saturated fat), margarine (high content of trans fatty acids), white bread (high glycaemic index), pancakes/Belgian waffles, cookies (high glycaemic index, high trans fatty acids), ice cream, candy, sugar. For each food constituting a food score that was reported to be consumed at least 1–3 times per month (more often than never/seldom) a score of one was assigned; for the RFS this added up to a maximum score of 17 (one score was assigned for consumption of milk with 0.5% fat or 1.5% fat) and for the NRFS to a maximum score of 21. Consumption of the remaining ten foods (yoghurt with 3.0% fat, potatoes, rice, pasta, poultry, eggs, shell fish, sweet soups, jam, chocolate), seven non-alcoholic beverages (milk with 3.0% fat, juice [in Sweden generally high sugar content], sodas, lemonade, coffee, tea) and five alcoholic beverages (beer with different alcohol content, wine, liquor) assessed using the FFQ did not count towards any food score.

Our RFS differed from the one presented by Kant *et al.*[11] in three food items: we did not include juices due to the high sugar content of most juices consumed in Sweden, potatoes due to their high glycaemic index, and chicken because we do not consider poultry a health-promoting food by itself but rather because it may be substituted for red meat and red meat consumption is captured by our NRFS. Juices, potatoes and poultry were not considered in our RFS or in our NRFS. Furthermore, potatoes are one of the most frequently consumed foods in Sweden and thus would probably not have added much between-person variation in the food scores.

Identification of deaths and follow-up of the cohort

Deaths in the cohort and dates of death were ascertained through the Swedish Death Register. The Swedish Death Register captures the death of all Swedish citizens and the mortality follow-up of our cohort is close to 100%. Causes of death were ascertained through linkage to the Swedish Register of Death Causes.

Statistical analysis

Women were divided into five categories based on their RFS and NRFS, respectively. Categories were chosen to divide the study population into approximately similar size categories while maintaining approximately equally spaced categories (at the higher end of the distribution). Intake scores clustered among women reporting higher food scores. Therefore, women reporting consumption of 0–8 recommended foods were combined in the lowest category, the second category included 9, 10, and 11 recommended foods, the third 12 and 13 recommended foods, the fourth 14 and 15 recommended foods, and the highest category comprised 16 and 17 recommended foods. The respective numbers for NRFS were 0–8, 9–11, 12–14, 15–17, and 18–21.

The RFS and NRFS were both included in each model to represent dietary intake most completely and to obtain results for RFS adjusted for NRFS and vice versa. Death rates were calculated by dividing the number of deaths by person-years of follow-up. Hazard ratios (HR) and 95% CI were estimated using Cox proportional hazards models.[17] Follow-up was censored at date of death or end of the follow-up period defined as the last match with the Swedish Death Register (31 December 1998). Equivalent analyses were performed for cause-specific mortality due to cancer, CHD and stroke, respectively. Secondary analyses were conducted excluding deaths during the first 3 and the first 5 years of follow-up, respectively.

The second, third, fourth, and fifth category of intake were each related to the first (lowest). A test for trend in mortality rates employed the median value for each category. The association between food scores and mortality was also assessed using food scores as continuous variables. Total energy intake (as estimated by the FFQ) was included in the model to provide some degree of control for between-person variation in total energy intake (due to real between-person differences as well as due to reporting bias).

Hazard ratio estimates were adjusted for age (in 5-year categories), height (continuous), body mass index (BMI) (continuous), parity (0, 1 or 2, ⩾3), age at first birth (⩽25, 26–30, ⩾31), education (less than high school, high school, university), marital status (single, married, living with partner, divorced, widowed), total energy intake (quintiles), and alcohol consumption (quintiles). Data on supplement intake, smoking, and physical activity were not collected in this cohort.

For cause-specific mortality analyses, women who died of causes other than the one under study were censored from the analysis at the time of death.

Results

Among our study cohort of 59 038 women, 3710 died during an average of 9.9 years of follow-up and a total of 585 636 person-years of observation. The distribution of the number of deaths, age at baseline, height, BMI, number of children, age at first birth, education, and marital status by categories of RFS is given in Table 1. Married women and women with two children

Table 1 Distribution of risk factors for mortality by categories of recommended foods score among 59 038 women in the Swedish Mammography Cohort Study

	Recommended Foods Score				
	Category 1	Category 2	Category 3	Category 4	Category 5
Recommended Foods Score Median (Range)	6 (0–8)	10 (9–11)	13 (12–13)	14 (14–15)	16 (16–17)
No. of women	3769	11 402	17 024	18 957	7886
No. of deaths	500	906	1036	947	321
Age[a]	58.2	54.9	53.5	52.9	52.0
Height (m)[a]	1.63	1.64	1.64	1.64	1.65
Body mass index (kg/m^2)[a]	25.3	24.8	24.7	24.6	24.8
No. of children (%)					
Nulliparous	14.9	12.6	10.4	9.5	8.5
1	19.2	18.6	17.6	15.7	13.9
2	30.9	36.5	40.0	41.1	43.8
3+	35.0	32.3	31.9	33.7	33.8
Age at first birth[a]	23.7	23.8	24.1	24.3	24.4
⩾10 years education (%)	5.7	9.5	11.5	12.8	14.2
Marital status (%)					
Single	8.8	7.2	6.1	5.4	4.6
Married	57.4	64.7	70.0	73.4	76.9
Divorced	11.6	10.5	8.8	8.1	7.2
Widowed	16.2	10.8	8.4	7.2	5.6
Living with partner	6.1	6.8	6.8	6.0	5.7

[a] Mean values.

had the highest healthy foods scores. A lower variety of recommended foods was associated with older age, lower education, and non-marital status (single, divorced, widowed). Similarly, a lower variety of not recommended foods was reported by women of older age, lower education, with less children, and of non-marital status (data not shown). The daily mean intake of energy and selected nutrients by categories of RFS and NRFS are presented in Table 2. Energy intake and alcohol consumption increased with RFS. Dietary fibre, vitamin C, vitamin E, folate, and beta-carotene consumption expressed as density per 1000 calories increased with increasing RFS, whereas percentage of energy from fat (in particular saturated fat) was marginally inversely associated and percentage of energy from carbohydrates and from protein was marginally positively related with RFS thereby confirming the higher quality of diet. A higher NRFS score was associated with higher energy, fat (in particular saturated and monounsaturated fatty acids) and alcohol consumption, and marginally inversely with carbohydrate and protein intake. The intake of fibre, vitamin C, vitamin E and folate was marginally higher among women with high NRFS, but beta-carotene intake was inversely related.

Age- and covariate-adjusted hazard ratio (HR) estimates for all-cause mortality by categories of RFS and NRFS are presented in Table 3. Mortality was significantly lower for higher RFS (*P* for trend <0.0001). Women who reported consuming 16–17 items from the RFS had an age-adjusted HR of dying of 0.54 (95% CI: 0.46–0.62) compared to women reporting 0–8 items. Associations changed only marginally after adjusting for age, height, BMI, number of children, age at first birth, education, marital status, energy intake, and alcohol consumption. For each

Table 2 Estimated mean daily intake of energy and selected nutrients by categories of recommended foods score among 59 038 women in the Swedish Mammography Cohort Study

	Recommended Foods Score				
	Category 1	Category 2	Category 3	Category 4	Category 5
Recommended Foods Score Median (Range)	6 (0–8)	10 (9–11)	13 (12–13)	14 (14–15)	16 (16–17)
Energy, kcal	1050	1226	1316	1403	1480
Percentage of energy from fat	32.3	31.4	30.7	30.0	29.2
from saturated fat	13.8	13.4	13.0	12.6	12.1
from monounsaturated fat	11.8	11.4	11.1	10.9	10.7
from polyunsaturated fat	4.5	4.5	4.4	4.4	4.4
from omega-6	3.9	3.9	3.8	3.7	3.7
from omega-3	0.5	0.5	0.5	0.5	0.5
Percentage of energy from carbohydrates	49.8	50.6	51.0	51.5	51.9
Percentage of energy from protein	15.9	15.9	16.2	16.5	16.9
Dietary fibre, g per 1000 kcal	11.5	15.0	16.8	18.4	19.8
Vitamin C, mg per 1000 kcal	41.2	57.9	66.5	74.0	82.0
Vitamin E, mg per 1000 kcal	5.0	6.0	6.5	6.9	7.4
Folate, μg per 1000 kcal	132.2	167.3	187.5	208.0	229.7
Beta-carotene, mg per 1000 kcal	1.6	2.3	2.6	3.0	3.4
Alcohol, g/day[a]	2.2	2.9	3.2	3.2	3.3
	Not Recommended Foods Score				
	Category 1	Category 2	Category 3	Category 4	Category 5
Not Recommended Foods Score Median (Range)	6 (0–8)	10 (9–11)	13 (12–14)	16 (15–17)	19 (18–21)
Energy, kcal	1064	1196	1317	1451	1618
Percentage of energy from fat	26.9	29.2	30.6	31.8	32.6
from saturated fat	11.4	12.4	12.9	13.4	13.7
from monounsaturated fat	9.5	10.5	11.2	11.7	12.1
from polyunsaturated fat	4.0	4.3	4.4	4.6	4.6
from omega-6	3.4	3.6	3.8	4.0	4.0
from omega-3	0.6	0.5	0.5	0.5	0.5
Percentage of energy from carbohydrates	54.3	52.0	51.0	50.1	49.6
Percentage of energy from protein	16.6	16.7	16.3	16.0	15.8
Dietary fibre, g per 1000 kcal	16.2	16.5	17.0	17.4	18.2
Vitamin C, mg per 1000 kcal	67.8	66.2	66.6	68.5	73.2
Vitamin E, mg per 1000 kcal	5.1	5.9	6.5	7.1	7.9
Folate, μg per 1000 kcal	181.3	184.6	189.4	198.5	215.5
Beta-carotene, mg per 1000 kcal	3.3	2.9	2.7	2.5	2.5
Alcohol, g/day[a]	2.4	2.8	3.1	3.3	3.9

[a] Includes never drinkers.

Table 3 Age- and covariate-adjusted hazard ratio (HR) estimates and 95% CI for all-cause mortality by categories of recommended foods score (RFS) and not recommended foods score (NRFS)

	Foods score					**Trend test**	
	Category 1	Category 2	Category 3	Category 4	Category 5	χ^2_1	*P*-value
RFS[a]; Median (Range)	6 (0–8)	10 (9–11)	13 (12–13)	14 (14–15)	16 (16–17)		
NRFS[a]; Median (Range)	6 (0–8)	10 (9–11)	13 (12–14)	16 (15–17)	19 (18–21)		
ALL-CAUSE MORTALITY							
No. of deaths RFS	500	906	1036	947	321		
No. of deaths NRFS	578	896	1236	859	141		
Age-adjusted HR (95% CI)							
RFS	1.00	0.77 (0.69–0.86)	0.67 (0.60–0.75)	0.60 (0.53–0.67)	0.54 (0.46–0.62)	98.5	<0.0001
NRFS	1.00	0.98 (0.88–1.09)	0.95 (0.85–1.05)	0.94 (0.84–1.06)	1.02 (0.84–1.23)	0.7	0.40
Covariate-adjusted HR (95% CI)[b]							
RFS	1.00	0.79 (0.70–0.88)	0.71 (0.63–0.80)	0.64 (0.57–0.72)	0.58 (0.50–0.68)	67.8	<0.0001
NRFS	1.00	1.00 (0.90–1.11)	0.98 (0.88–1.09)	0.98 (0.87–1.11)	1.07 (0.88–1.31)	0.01	0.92
Excluding first 5 years of follow-up[c]							
Age-adjusted HR (95% CI)							
RFS	1.00	0.79 (0.69–0.91)	0.69 (0.60–0.79)	0.58 (0.51–0.68)	0.54 (0.45–0.65)	68.5	<0.0001
NRFS	1.00	0.99 (0.87–1.13)	0.92 (0.81–1.05)	0.95 (0.82–1.09)	0.99 (0.78–1.26)	0.8	0.37
Covariate-adjusted HR (95% CI)[b]							
RFS	1.00	0.80 (0.70–0.92)	0.72 (0.63–0.83)	0.62 (0.53–0.72)	0.58 (0.48–0.70)	49.5	<0.0001
NRFS	1.00	1.01 (0.88–1.15)	0.94 (0.83–1.08)	0.97 (0.84–1.13)	1.01 (0.79–1.30)	0.3	0.59

[a] For RFS number of recommended foods increases with increasing category; for NRFS number of not recommended foods increases with increasing category.

[b] Hazard Ratio and 95% CI adjusted for age, height, body mass index, number of children, age at first birth, education, marital status, alcohol consumption, energy intake, and simultaneous adjustment for RFS or NRFS.

[c] 2413 deaths occurred after more than 5 years of follow-up.

additional healthy food consumed the risk of death was about 5% lower (covariate-adjusted HR = 0.95; 95% CI: 0.94–0.96). The association between RFS and mortality was not affected by accounting for NRFS. The NRFS was not related to all-cause mortality (Table 3).

To exclude the possibility that the inverse association between RFS and mortality was the result of a morbidity-related change in diet or altered reporting of diet, the analysis was repeated excluding deaths during the first 3 and 5 years of follow-up. The inverse association between RFS and mortality was not affected by these exclusions. (The results for exclusion of the first 5 years of follow-up are shown in Table 3.)

The relation of RFS and NRFS and mortality was also examined separately for mortality from cancer, CHD, and stroke (Table 4). The two most common fatal cancers were lung cancer (n = 231) and breast cancer (n = 154). Cancer mortality remained considerably reduced with a higher variety of recommended foods (RFS: 16–17 versus 0–8: covariate-adjusted HR = 0.76; 95% CI: 0.60–0.96), but was significantly elevated among women who reported a high variety of less healthy foods (NRFS: 18–21 versus 0–8: covariate-adjusted HR = 1.52; 95% CI: 1.13–2.05). A high RFS was particularly important to reduce CHD and stroke mortality (CHD: RFS: 16–17 versus 0–8: covariate-adjusted HR = 0.47; 95% CI: 0.33–0.68; stroke: RFS: 16–17 versus 0–8: covariate-adjusted HR = 0.40; 95% CI: 0.22–0.73), whereas the NRFS did not appear to play an important role for these causes of death (CHD: NRFS: 18–21 versus 0–8: covariate-adjusted HR = 0.79; 95% CI: 0.47–1.32; stroke: NRFS: 18–21 versus 0–8: covariate-adjusted HR = 0.96; 95% CI: 0.47–1.97).

To investigate the possibility that smoking may have confounded the associations observed we related RFS and NRFS to cancer deaths other than lung cancer. The associations were somewhat weakened but remained supportive of the findings including lung cancer deaths: RFS: 9–11, 12–13, 14–15, 16–17 versus 0–8: covariate-adjusted HR: 0.82, 0.72, 0.73, 0.81, respectively; NRFS: 9–11, 12–14, 15–17, 18–21 versus 0–8: covariate-adjusted HR: 1.22, 1.18, 1.19, 1.37, respectively.

Discussion

In this population-based prospective cohort study of Swedish women we found a strong association between quality of diet and longevity during 9.9 years of follow-up. Women who reported consuming a diet characterized by a great variety of vegetables, fruit, whole grain breads and cereals, fish, and low-fat dairy products—reflected also in a higher density of health-favouring dietary constituents—had a significantly reduced risk of mortality compared to women who consumed few of these foods. Interestingly, overall mortality was not affected by consuming foods commonly perceived as less healthy such as red meats, refined carbohydrates and sugars, and foods rich in saturated fats. In fact, women who reported consuming a larger number of less healthy foods had a similar life expectancy during the 9.9 years of follow-up as women who ate very few of these less healthy foods. It is possible that it is more important for longevity to consume a large variety of health-promoting foods to help prevent disease than to reduce the number of less healthy foods. Such interpretation has to be made cautiously, however, as we did not consider frequency and quantity of

Table 4 Age- and covariate-adjusted hazard ratio (HR) estimates and 95% CI for cause-specific mortality by categories of recommended foods score (RFS) and not recommended foods score (NRFS)

	Foods score					**Trend test**	
	Category 1	Category 2	Category 3	Category 4	Category 5	χ^2_1	*P*-value
RFS[a]; Median (Range)	6 (0–8)	10 (9–11)	13 (12–13)	14 (14–15)	16 (16–17)		
NRFS[a]; Median (Range)	6 (0–8)	10 (9–11)	13 (12–14)	16 (15–17)	19 (18–21)		
ALL SITES CANCER							
No. deaths RFS	158	325	409	416	170		
No. deaths NRFS	168	342	508	385	75		
Age-adjusted HR (95% CI)							
RFS	1.00	0.78 (0.65–0–95)	0.72 (0.59–0.87)	0.68 (0.56–0.83)	0.71 (0.56–0.89)	12.8	0.0004
NRFS	1.00	1.21 (1.00–1.45)	1.19 (0.99–1.42)	1.19 (0.98–1.45)	1.40 (1.05–1.86)	3.2	0.07
Covariate-adjusted HR (95% CI)[b]							
RFS	1.00	0.80 (0.66–0.98)	0.75 (0.62–0.91)	0.72 (0.59–0.88)	0.76 (0.60–0.96)	7.8	0.005
NRFS	1.00	1.24 (1.03–1.49)	1.24 (1.03–1.50)	1.27 (1.04–1.55)	1.52 (1.13–2.05)	5.7	0.02
CORONARY HEART DISEASE							
No. deaths RFS	120	208	224	179	48		
No. deaths NRFS	137	221	253	150	18		
Age-adjusted HR (95% CI)							
RFS	1.00	0.80 (0.63–1.00)	0.71 (0.56–0.89)	0.57 (0.45–0.73)	0.43 (0.30–0.61)	30.6	<0.0001
NRFS	1.00	1.07 (0.86–1.32)	0.91 (0.73–1.13)	0.85 (0.66–1.09)	0.79 (0.48–1.32)	3.2	0.07
Covariate-adjusted HR (95% CI)[b]							
RFS	1.00	0.81 (0.65–1.03)	0.75 (0.59–0.95)	0.62 (0.48–0.80)	0.47 (0.33–0.68)	27.2	<0.0001
NRFS	1.00	1.08 (0.87–1.34)	0.91 (0.73–1.14)	0.85 (0.65–1.10)	0.79 (0.47–1.32)	2.9	0.09
STROKE							
No. deaths RFS	48	74	112	91	17		
No. deaths NRFS	62	81	103	86	10		
Age-adjusted HR (95% CI)							
RFS	1.00	0.72 (0.50–1.04)	0.87 (0.61–1.24)	0.69 (0.47–1.01)	0.36 (0.20–0.64)	7.2	0.007
NRFS	1.00	0.85 (0.61–1.19)	0.78 (0.56–1.09)	1.01 (0.71–1.44)	0.90 (0.45–1.80)	0.0	0.98
Covariate-adjusted HR (95% CI)[b]							
RFS	1.00	0.77 (0.53–1.12)	0.96 (0.66–1.38)	0.78 (0.53–1.15)	0.40 (0.22–0.73)	7.2	0.007
NRFS	1.00	0.88 (0.63–1.24)	0.83 (0.53–1.71)	1.07 (0.73–1.56)	0.96 (0.47–1.97)	0.0	0.98

[a] For RFS number of recommended foods increases with increasing category; for NRFS number of not recommended foods increases with increasing category.

[b] Hazard Ratio and 95% CI adjusted for age, height, body mass index, number of children, age at first birth, education, marital status, alcohol consumption, energy intake, and simultaneous adjustment for RFS/NRFS.

foods consumed in this study and we therefore cannot exclude the possibility that a few less healthy foods consumed often and in large amounts are associated with an increased mortality. Cancer mortality, however, was affected by both the variety of recommended and non-recommended foods.

Kant *et al.* have recently reported results from a similar study.[11] They used data from the Breast Cancer Detection Demonstration project, a cohort of 42 254 US American women who were followed for 5.6 years. Dietary intake was assessed using a comparable 62-item FFQ and their RFS was similar to ours but additionally included fruit juice, poultry and potatoes. Furthermore, Kant *et al.* based their scoring on a frequency of consumption of at least once per week whereas we simplified this to ever intake with any regularity (more often than 'never/seldom'). In our study with about double the length of follow-up and almost double the number of deaths of the US cohort we found a risk reduction of similar magnitude associated with a high RFS. More importantly, Kant *et al.* restricted their research to the use of an RFS and did not study the influence of a diet including regular consumption of a large number of not recommended foods.

Some other studies have defined simple dietary patterns and related them to mortality. In a case-control study conducted in Greece, individuals who followed dietary patterns common in the Mediterranean region (high in fruit, vegetables, legumes, cereals, low in meat, milk and dairy products, and high monounsaturated:saturated fat ratio) had a low mortality.[18] In a cohort study including men in Finland, Italy, and the Netherlands[19] a healthy diet indicator was calculated following guidelines for the prevention of chronic diseases defined by the World Health Organization.[20] A healthy diet defined as high consumption of fruit, vegetables, pulses, nuts, seeds, low intake of saturated fatty acids, cholesterol and mono- and disaccharides, and intake of polyunsaturated fatty acids, protein, complex carbohydrates, and dietary fibre within recommended ranges reduced mortality.[19] Osler and colleagues followed 7316 Danish women and men and identified three dietary patterns from a 28-item FFQ.[21] A healthy food index and a prudent diet were inversely related to all-cause and cardiovascular mortality while a Western dietary pattern characterized by frequent intake of meat products, white bread, butter and lard was not associated with mortality. These findings are consistent with ours.

Other more refined dietary pattern analyses have been derived using factor analysis or cluster analysis. Such dietary patterns are data-driven and therefore more specifically tailored to the cohort studied. We aimed to develop an intuitive, common sense classification of foods that may be easy to communicate to the public as dietary recommendations. Our paper adds to existing work in considering the overall dietary pattern as a risk factor for mortality. While many studies have considered specific foods or nutrients in relation to specific diseases, there is little evidence relating overall diet to overall risk of death.

Measurement error is inherent in questionnaire-based dietary assessment. Random within-person variation might misclassify true average intake and could affect ranking of individuals; such misclassification would lead to an underestimation of the association with disease outcome (regression dilution bias). Regression dilution bias, however, is less of a problem with an instrument that is asking people to report an average consumption over time. Due to its structured form and its limited number of food items, however, the FFQ generally leads to an underestimation of intake as reflected in the relatively low total caloric intake in this cohort. As we simplified our classification to ignore frequency of consumption as well as portion size, our results should be largely unaffected by quantitative misreporting and it is unlikely that the underreporting in energy intake may have differentially affected reporting of recommended and non-recommended foods. Rather, our scores give weight to foods reported on the FFQ as consumed with any regularity, thus indicating dietary preferences. The FFQ should be able to roughly separate individuals with a very high RFS (or NRFS) from women with a low RFS (or NRFS). Therefore, comparison of extreme categories of intake should be informative. Indeed, both of our diet scores separated women by nutrient intake. This was not explained by an increase in total caloric intake: the mean energy intake in the highest category was 141% of the lowest category while mean beta-carotene intake was 213%, vitamin C 199%, folate 172%, and fibre 172% of the lowest category. Conversely, the mean caloric intake in the highest NRFS category was 152% of that in the lowest category, but the mean beta-carotene intake was lowest in the highest category and folate and fibre intake were 119% and 112% of the lowest category, respectively (Table 2).

Our study shows the importance of studying diet as a whole rather than investigating the health effects of individual foods or nutrients. Analytical models including single foods or nutrients to predict disease are inherently confounded by the foods and nutrients not accounted for. Thus, a harmful effect of a single food or nutrient may reflect lower intake of healthy foods since caloric intake is fairly constant across humans.

The most important threat to validity when relating a healthy diet to disease or mortality is confounding by other lifestyle factors which are associated with healthy eating as well as longevity. While we adjusted for a number of potential confounders we lacked information on smoking habits, physical activity, and dietary supplement use all of which are likely to confound the association. Smoking rates in Sweden among women in the age group studied are low. Among women aged 55–64 years the prevalence of current smoking is 23.4%, among women 65- to 74-years-old the prevalence is 14.2%.[22] When we resurveyed our study participants in 1997, the prevalence of current smokers was 19%, former smokers 27%, and never smokers 52%. Body mass index and energy intake are considered to be a good proxy for physical activity.[23] Furthermore, dietary supplements have not been found to be associated with mortality.[24] In the study by Kant *et al.* information on these factors was available.[11] While covariate adjustment attenuated relative risk estimates somewhat in their study, the overall results and trends remained unchanged.[11] In our data, the relation of RFS and NRFS with cancer mortality was slightly reduced after excluding lung cancer deaths but associations remained strong. While confounding by smoking might explain the stronger relation between RFS and cardiovascular disease it could not explain the lack of association between NRFS and cardiovascular disease. Nevertheless, residual confounding by a healthy lifestyle has to be considered as a possible explanation for at least part of the observed associations.

Our study on diet and longevity does not permit inferences about biological mechanisms underlying our observations. We can only speculate on essential dietary constituents that might aid the prevention of disease. Public health recommendations, however, do not require complete mechanistic insights and recommendations of preferable foods and food groups may be easier to implement by the public than nutrient-based advice.

In conclusion, we found a strong association between a diet rich in healthy foods and low mortality whereas a diet with a large variety of less healthy foods was only found to affect cancer mortality. For each additional healthy food added to the diet, overall mortality was lowered by 5%. Our study supports the recommendation to consume an abundance of vegetables, fruit, whole grain breads and cereals, fish, and low-fat dairy products.

Acknowledgements

Dr Michels was supported by a grant from the US National Institutes of Health R01 DK 54900–02 to perform this research.

KEY MESSAGES

- A healthy diet can affect longevity.
- Regular consumption of a high variety of healthy foods is associated with longevity and a lower mortality from cardiovascular disease and cancer.
- A diet with a high variety of less healthy foods appears to be less important for longevity but resulted in increased mortality from cancer in our study.

References

[1] Schwerin HS, Stanton JL, Riley AM Jr *et al.* Food eating patterns and health: a re-examination of the Ten-State and HANES I surveys. *Am J Clin Nutr* 1981;**34:**568–80.

[2] Jacobson HN, Stanton JL. Pattern analysis in nutrition. *Clin Nutr* 1986; **5:**249–53.

[3] Slattery ML, Boucher KM, Caan BJ, Potter JD, Ma K-N. Eating patterns and risk of colon cancer. *Am J Epidemiol* 1998;**148:**4–16.

[4] Hu FB, Rimm E, Smith-Warner SA *et al.* Reproducibility and validity of dietary patterns assessed with a food-frequency questionnaire. *Am J Clin Nutr* 1999;**69:**243–49.

[5] Millen BE, Quatromoni PA, Gagnon DR, Cupples LA, Franz MM, D'Agostino RB. Dietary patterns of men and women suggest targets for health promotion: The Framingham Nutrition Studies. *Am J Health Promotion* 1996;**11:**42–53.

[6] Tucker KL, Dallal GE, Rush D. Dietary patterns of elderly Boston-area residents defined by cluster analysis. *J Am Diet Assoc* 1992;**92:**1487–91.

[7] Wirfalt E, Mattisson I, Gullberg B, Berglund G. Food patterns defined by cluster analysis and their utility as dietary exposure variables: a report from the Malmo Diet and Cancer Study. *Public Health Nutr* 2000;**3:**159–73.

[8] Akin JS, Guilkey DK, Popkin BM, Fanelli MT. Cluster analysis of food consumption patterns of older Americans. *J Am Diet Assoc* 1986;**86:** 616–24.

[9] Farchi G, Mariotti S, Menotti A, Seccareccia F, Torsello S, Fidanza F. Diet and 20-year mortality in two rural population groups of middle-aged men in Italy. *Am J Clin Nutr* 1989;**50:**1095–103.

[10] Kant AK. Indices of overall diet quality: a review. *J Am Diet Assoc* 1996; **96:**785–91.

[11] Kant AK, Schatzkin A, Graubard BI, Schairer C. A prospective study of diet quality and mortality in women. *JAMA* 2000;**283:**2109–15.

[12] Wolk A, Bergström R, Hunter D *et al.* A prospective study of association of monounsaturated and other types of fat with risk of breast cancer. *Arch Intern Med* 1998;**158:**41–45.

[13] Michels KB, Holmberg L, Bergkvist L, Ljung H, Bruce Å, Wolk A. Dietary antioxidant vitamins, retinol, and breast cancer incidence in a cohort of Swedish women. *Int J Cancer* 2001;**91:**563–67.

[14] Willett WC, Sampson L, Stampfer MJ *et al.* Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol* 1985;**122:**51–65.

[15] National Research Council, Commission on Life Sciences, Food and Nutrition Board, Committee on Diet and Health. *Diet and Health. Implications for Reducing Chronic Disease Risk*. Washington DC: National Academy Press, 1989.

[16] Willett WC. Diet and health: what should we eat? *Science* 1994;**264:** 532–37.

[17] Cox DR. Regression models and life tables. *J R Statist Soc* 1972;**32:** 187–220.

[18] Trichopoulou A, Kouris-Blazos A, Wahlqvist ML *et al.* Diet and overall survival in elderly people. *BMJ* 1995;**311:**1457–60.

[19] Huijbregts P, Feskens E, Räsänen L *et al.* Dietary pattern and 20 year mortality in elderly men in Finland, Italy, and the Netherlands: longitudinal cohort study. *BMJ* 1997;**315:**13–17.

[20] World Health Organization. *Diet, Nutrition, and the Prevention of Chronic Diseases*. Report of a WHO Study Group. Geneva: WHO, 1990. (WHO Technical Report Series No. 797).

[21] Osler M, Heitmann BL, Gerdes LU, Jorgensen LM, Schroll M. Dietary patterns and mortality in Danish men and women: a prospective observational study. *Br J Nutr* 2001;**85:**219–25.

[22] Statistics Sweden. *Health and Medical Care 1980–89. Living Conditions.* (In Swedish.) Report 76. SCB Tryckeri, Örebro 1992, p. 171.

[23] Willett W, Howe G, Kushi L. Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr* 1997;**65(Suppl.4):**S1220–28.

[24] Kim J, Williamson DF, Byers T, Kaplan JP. Vitamin and mineral supplement use and mortality in a US cohort. *Am J Public Health* 1993; **83:**546–50.