



OPEN ACCESS

CrossMark  
click for updates

# Fruit and vegetable consumption in adolescence and early adulthood and risk of breast cancer: population based cohort study

Maryam S Farvid,<sup>1,2</sup> Wendy Y Chen,<sup>3,4</sup> Karin B Michels,<sup>3,5,6</sup> Eunyoung Cho,<sup>3,7</sup> Walter C Willett,<sup>1,3,5</sup> A Heather Eliassen<sup>3,5</sup>

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

<sup>2</sup>Harvard/Massachusetts General Hospital Center on Genomics, Vulnerable Populations, and Health Disparities, Mongan Institute for Health Policy, Massachusetts General Hospital, Boston, MA, USA

<sup>3</sup>Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA

<sup>4</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA

<sup>5</sup>Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

<sup>6</sup>Obstetrics and Gynecology Epidemiology Center, Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA

<sup>7</sup>Department of Dermatology, The Warren Alpert Medical School of Brown University, Providence, RI, USA

Correspondence to: M S Farvid mfarvid@hsph.harvard.edu

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2016;353:i2343 <http://dx.doi.org/10.1136/bmj.i2343>

Accepted: 12 April 2016

## ABSTRACT

### OBJECTIVE

To evaluate the association between fruit and vegetable intake during adolescence and early adulthood and risk of breast cancer.

### DESIGN

Prospective cohort study.

### SETTING

Health professionals in the United States.

### PARTICIPANTS

90 476 premenopausal women aged 27-44 from the Nurses' Health Study II who completed a questionnaire on diet in 1991 as well as 44 223 of those women who completed a questionnaire about their diet during adolescence in 1998.

### MAIN OUTCOME MEASURE

Incident cases of invasive breast cancer, identified through self report and confirmed by pathology report.

### RESULTS

There were 3235 cases of invasive breast cancer during follow-up to 2013. Of these, 1347 cases were among women who completed a questionnaire about their diet during adolescence (ages 13-18). Total fruit consumption during adolescence was associated with a lower risk of breast cancer. The hazard ratio was 0.75 (95% confidence interval 0.62 to 0.90;  $P=0.01$  for trend) for the highest (median intake 2.9 servings/day) versus the lowest (median intake 0.5 serving/day) fifth of intake. The association for fruit intake during adolescence was independent of adult fruit intake. There was no association between risk and total fruit intake in early adulthood and total vegetable intake in either adolescence or early adulthood. Higher early adulthood intake of fruits and vegetables rich in  $\alpha$

carotene was associated with lower risk of premenopausal breast cancer. The hazard ratio was 0.82 (0.70 to 0.96) for the highest fifth (median intake 0.5 serving/day) versus the lowest fifth (median intake 0.03 serving/day) intake. The association with adolescent fruit intake was stronger for both estrogen and progesterone receptor negative cancers than estrogen and progesterone receptor positive cancers ( $P=0.02$  for heterogeneity). For individual fruits and vegetables, greater consumption of apple, banana, and grapes during adolescence and oranges and kale during early adulthood was significantly associated with a reduced risk of breast cancer. Fruit juice intake in adolescence or early adulthood was not associated with risk.

### CONCLUSION

There is an association between higher fruit intake and lower risk of breast cancer. Food choices during adolescence might be particularly important.

### Introduction

Fruits and vegetables are important sources of fiber, vitamins, and other biologically active substances that can favorably affect the pathogenesis of breast cancer through several biological mechanisms.<sup>1-3</sup> Results from epidemiological studies assessing fruit and vegetable intake and risk of breast cancer, however, have been inconsistent.<sup>4-16</sup> A recent pooled analysis of 20 prospective studies showed an inverse association between vegetable intake and estrogen receptor negative tumors, but not with estrogen receptor positive tumors or breast cancer overall.<sup>17</sup> All these results, however, were from dietary assessments during mid-life or later; the role of fruit and vegetable intake in earlier life is unclear. Breast tissue might be particularly susceptible to carcinogenic exposures during childhood and early adult life, as seen in women exposed to ionizing radiation.<sup>18-20</sup>

We used unique information on diet in early life in the Nurses' Health Study II (NHSII) to conduct a detailed analysis of associations between fruit and vegetable intake during adolescence and early adulthood and risk of breast cancer. This included consideration of tumor hormone receptor and menopausal status at diagnosis and the relation of specific fruits and vegetables to risk.

### Methods

#### Study population

The NHSII is an ongoing prospective cohort study that began in 1989 with enrollment of 116 430 female registered nurses aged 25-42. Our analyses included 97 813

## WHAT IS ALREADY KNOWN ON THIS TOPIC

Fruit and vegetable consumption has been hypothesized to protect against breast cancer but the overall evidence has not been supportive

Most studies have assessed intakes during midlife and later, which could be after the period when breast tissue is more vulnerable to carcinogenic influences

## WHAT THIS STUDY ADDS

Fruit intake during adolescence seem beneficial in terms of preventing breast cancer

Higher intake of fruit and vegetables rich in  $\alpha$  carotene during early adulthood was specifically associated with lower risk

The associations with breast cancer differed significantly among individual fruits and vegetables: greater consumption of apple, orange, banana, grapes, and kale was significantly associated with a reduced risk

women who returned the 1991 food frequency questionnaire (FFQ). We excluded women with implausible total energy intake (<2508 kJ or 14 630 kJ (<600 or >3500 kcal)/day) or who left all items on fruit and/or vegetables blank in the questionnaire and those who were postmenopausal in 1991, had missing information on age, or had previously had cancer (except non-melanoma skin cancer) before the recruitment questionnaire. The primary analyses included 90 476 women. The follow-up rate was over 96% of total potential person years from 1991 to 2013.

In 1997, participants were asked about their willingness to complete a supplemental food frequency questionnaire about diet during high school (age range 13-18) (HS-FFQ). Among 47 355 of women who returned the HS-FFQ in 1998, we excluded those who had had cancer (except non-melanoma skin cancer) before 1998, reported implausible daily caloric intake (<2508 kJ or 20 900 kJ (<600 or  $\geq$ 5000 kcal)), or who left all items on fruit and/or vegetables blank. Thus we evaluated adolescent fruit and vegetable intake and risk of breast cancer in 44 223 women. The follow-up rate was over 98% of total potential person years from 1998 through 2013.

#### Dietary assessment

In 1991 and every four years thereafter, dietary intake and alcohol consumption were measured with a validated semiquantitative FFQ with about 130 items about usual dietary intake and alcohol consumption during the past year (available at <http://www.nurseshealth-study.org/participants/questionnaires>). Validity of the questionnaire to assess long term intake was evaluated by comparison with weighed diet records collected over a period of six years; the average correlation for dietary nutrient intakes, corrected for variation in diet records, was 0.83.<sup>21</sup> In addition, higher fruit and/or vegetable intake assessed by this questionnaire has been associated with lower risks of diabetes<sup>22</sup> and coronary heart disease,<sup>23</sup> indirectly supporting the validity of the questionnaire. We asked women about the frequency of consumption of fruits and vegetables in nine categories ranging from “never or less than once per month” to “6 or more times per day.” Total fruit, fruit juice and vegetable items listed on the questionnaire are presented in appendix table A. Vegetables were categorized into five subgroups including green leafy vegetables, yellow/orange vegetables, tomatoes, cruciferous vegetables, and other vegetables. Fruits were grouped into seven subgroups including citrus fruits, berries, pomes (fruits with a core of several small seeds), drupes (fruits with a stone), tropical fruits, grapes, and melon. We also grouped fruits and vegetables by their content of vitamin C ( $\geq$ 40 mg/100 g),  $\alpha$  carotene ( $\geq$ 3000  $\mu$ g  $\alpha$  carotene/100 g),  $\beta$  carotene ( $\geq$ 3000  $\mu$ g/100 g), or lutein ( $\geq$ 10 mg/100 g) (appendix table A).<sup>23-27</sup>

We evaluated diet in adolescence with a FFQ that included 124 foods typically consumed between 1960 and 1980 when participants were in high school (HS-FFQ, see appendix table B). We categorized intakes of fruits and vegetables similar to early adulthood intakes. The validity of the HS-FFQ was evaluated

among 80 young women who completed three 24 hour recalls and two FFQ during high school with the HS-FFQ administered 10 years later; the mean of corrected correlation coefficients for energy adjusted nutrient intakes was 0.58 for the HS-FFQ and the earlier FFQ's and 0.45 for the HS-FFQ and the earlier 24 hour recalls.<sup>28</sup> Furthermore, evidence of validity and reproducibility also came from the comparison of their dietary reports with the information provided four years later. In 2002, HS-FFQ was readministered to 333 women who randomly selected from those initial participants. The mean Pearson correlation was 0.65 (0.50-0.77) for 38 nutrient intakes, and the mean Spearman rank correlation was 0.60 (0.37-0.77) for food intakes.<sup>29</sup> In addition, validity was assessed by comparing adolescent diet reported by 272 NHS II participants with their diet reported by their mothers; the mean correlation was 0.40 for nutrients.<sup>29</sup>

#### Identification of cases of breast cancer

In follow-up questionnaires administered every two years, we asked women about diagnosis of breast cancer and the date of diagnosis. Deaths were reported by family members and the postal service in response to the follow-up questionnaires or identified through the National Death Index. When a case of breast cancer was identified, we asked the participant (or next of kin for those who had died) for permission to obtain relevant hospital records and pathology reports. Because accuracy was high for self reporting (99%), we included in the analysis cases in women with unavailable medical records (n=547). Estrogen receptor and progesterone receptor status of tumors was extracted from medical records.

#### Assessment of other variables

We obtained data on potential risk factors for breast cancer from the biennial NHSII questionnaires including age, height, weight, history of benign breast disease, family history of breast cancer, smoking, race, menopausal status, menopausal hormone use, oral contraceptive use, and physical activity, updated to the most recent information, if available. Participants provided informed consent by returning their questionnaires. Body mass index (BMI) at age 18 and alcohol consumption in adolescence was obtained from the 1989 questionnaire. Women were considered as premenopausal if they still had periods or had had a hysterectomy with at least one ovary remaining and were aged under 46 for smokers or 48 for non-smokers. Women were considered postmenopausal if they reported natural menopause or had undergone bilateral oophorectomy. We defined women of unknown menopausal status or who had hysterectomy without bilateral oophorectomy as postmenopausal if they were aged  $\geq$ 54 for smokers or  $\geq$ 56 for non-smokers.<sup>30</sup>

#### Statistical analysis

Food intakes reported in 1991, when participants were aged 27-44, were considered as early adulthood dietary intake. To evaluate diet in early adulthood and breast

cancer, participants contributed person years from the date of return of the 1991 questionnaire until the date of any diagnosis of cancer except non-melanoma skin cancer, death, or end of follow-up period (1 June 2013), whichever occurred first. For adolescent fruit, vegetable, and fruit juice intake, participants contributed person years similarly except that follow-up began with return of the adolescent diet questionnaire in 1998. Participants were divided into fifths according to their dietary intake. We used Cox proportional hazards regression to estimate hazard ratios and 95% confidence intervals for each category, using the lowest fifth of intake as the reference category. To control for confounding by age or calendar time, or any possible two way interactions between these two time scales, the regression models included age in months as the time scale, stratified by calendar year of the current questionnaire cycle. Multivariable models also simultaneously adjusted for race, history of breast cancer in mother or sisters, history of benign breast disease, smoking, height, physical activity, BMI at age 18, weight change since age 18, age at menarche, parity and age at first birth, oral contraceptive use, menopausal status, menopausal hormone use, age at menopause, and intakes of alcohol and energy.

For adolescent fruit, vegetable, and fruit juice intake and risk of breast cancer, multivariable models adjusted additionally for adolescent alcohol intake and adolescent energy intake (instead of early adulthood energy intake). The median value for each fifth was used for tests for trend, modeled as a continuous variable. We also used a likelihood ratio test to determine a potential non-linear association by comparing the model with only the linear term to the model with both the linear and the cubic spline terms. We replaced missing covariate data with carried forward method for continuous variables and missing indicator method for categorical variables. To examine if the observed associations were independent of a generally healthy dietary pattern, we further adjusted for a modified alternate healthy eating index<sup>31</sup> that excluded the scores for fruits and vegetables to avoid redundancy with our primary variables. We further examined whether the associations between fruit and vegetable intake and risk of breast cancer depended on the other dietary factors, including intake of fiber, vitamin C,  $\beta$  carotene, animal fat, and total red meat, and we additionally adjusted for these variables as well. To better represent adolescent intakes and show that these associations were independent of adulthood dietary intakes, we adjusted for dietary intake during adult life (cumulative average dietary intakes).

Because dietary intake has been assumed to influence risk of breast cancer over an extended period of time, for a sensitivity analyses, we calculated the average of adolescent and early adult (1991) fruit, vegetable, and fruit juice intake among those with information for both periods. We also calculated premenopausal cumulative average of dietary intakes by using the 1991, 1995, 1999, 2003, 2007, and 2011 dietary

intake data, stopping updating when a woman reached menopause. To examine potential effect modification by BMI at age 18 on associations between fruit and vegetable consumption and risk, we included a cross-product term of the ordinal score for BMI at age 18 and fruit or vegetable intakes in the multivariable model. Tests for interactions were obtained from a likelihood ratio test. Furthermore, to evaluate whether associations varied by age, we conducted tests for interaction and performed analyses separately among women aged 27-35 and 36-44 at baseline. To evaluate the difference between premenopausal and postmenopausal hazard ratios, we evaluated the P value for heterogeneity with the Q statistic.<sup>32,33</sup> To examine differential associations of fruit and vegetable intake with risk of breast cancer by hormone receptor status (both estrogen and progesterone receptor positive, both estrogen and progesterone receptor negative, and estrogen receptor positive and progesterone receptor negative<sup>34</sup>) we used Cox proportional cause specific hazards regression model with a duplication method for competing risk data.<sup>35</sup> All P values were two sided. SAS version 9.3 (SAS Institute, Cary NC) was used for all analyses.

#### Patient involvement

No patients were involved in setting the research question or the outcome measures, nor were they involved in the design and implementation of the study. There are no plans to involve patients in dissemination.

#### Results

During 22 years of follow-up, we documented 3235 cases of invasive breast cancer for the early adulthood dietary analyses. Of these, 1347 cases were in women for whom we had adolescent dietary information. A higher fruit intake in early adulthood and adolescence was associated with lower prevalence of smoking, lower intakes of alcohol and animal fat, and higher fiber intake (tables 1 and 2). Women with higher intakes of vegetables during early adulthood were more likely to drink alcohol and to have higher fiber intake and were less likely to eat animal fat, to smoke, and to be nulliparous (table 1).

#### Fruit and vegetable intake during adolescence and risk of breast cancer

A higher total fruit intake during adolescence was associated with a lower risk of breast cancer (hazard ratio (for highest v lowest fifth of intake) 0.75, 95% confidence interval 0.62 to 0.90;  $P_{\text{trend}}=0.01$ ). Although women in third and fifth category seemed to have a similar lower risk compared with those with the lowest intake, a test for non-linearity was not significant ( $P=0.11$ ). The association was significant for postmenopausal (0.69, 0.52 to 0.90;  $P_{\text{trend}}=0.02$ ) but not premenopausal breast cancer (0.80, 0.60 to 1.05;  $P_{\text{trend}}=0.17$ ). There was, however, no significant difference between premenopausal and postmenopausal hazard ratios ( $P=0.45$  for differences in hazard ratios) (table 3). These associations were materially unchanged with adjustment

for intake of vitamin C,  $\beta$  carotene, animal fat, red meat, or modified alternate healthy eating index (appendix table C). Adjustment for adult fruit intake did not affect the associations (hazard ratio for highest v lowest 0.76 (0.63 to 0.93;  $P_{\text{trend}}=0.04$ ) among all women; 0.87 (0.64 to 1.19;  $P_{\text{trend}}=0.55$ ) among premenopausal women; and 0.67 (0.50 to 0.89;  $P_{\text{trend}}=0.02$ ) among postmenopausal women). The estimate was attenuated but remained significant after adjustment for adolescent total dietary fiber intake (0.80 (0.64 to 0.99;  $P_{\text{trend}}=0.14$ ) among all women; appendix table C).

The association between adolescent vegetable consumption and risk was non-linear ( $P=0.03$ ); the relation seemed to follow a dose-response inverse association up to 2.7 servings/day. Non-significant associations between adolescent vegetable intake and breast cancer were observed (table 4). Higher total fruit and vegetable intake during adolescence was associated with lower risk of breast cancer (hazard ratio (for highest v lowest fifth of intake) 0.82, 0.68 to 0.98;  $P_{\text{trend}}=0.07$ ) (table 5). Adolescent fruit juice intake was not associated with risk (table 6). We did not observe any significant

**Table 1 | Age and age standardized characteristics in 1991 according to intake of fruit and vegetables during early adulthood among women enrolled in Nurses' Health Study II. Figures are means (SD) unless otherwise specified**

	Fifth of total fruit intake			Fifth of total vegetable intake		
	1 (n=18 004)	3 (n=18 060)	5 (n=18 132)	1 (n=18 048)	3 (n=18 106)	5 (n=18 144)
Age (years)	36 (5)	36 (5)	37 (5)	36 (5)	36 (5)	37 (4)
Intake in early adulthood:						
Total fruit intake (serving/day)	0.2 (0.1)	0.9 (0.1)	2.4 (0.9)	0.6 (0.5)	1.0 (0.7)	1.7 (1.1)
Total vegetable intake (serving/day)	2.2 (1.4)	3.0 (1.5)	4.4 (2.3)	1.2 (0.4)	2.8 (0.3)	6.0 (1.9)
Total energy intake (kJ)	6383 (2111)	7399 (2103)	8786 (2261)	6078 (1977)	7449 (2027)	8987 (2278)
Alcohol consumption (g/day)	3.7 (7.6)	2.9 (5.5)	2.9 (5.3)	2.5 (5.8)	3.1 (5.8)	3.6 (6.5)
Total fiber intake (g/day)	14.6 (4.2)	18.0 (4.4)	22.7 (5.8)	14.4 (4.3)	17.9 (4.3)	23.0 (5.8)
Animal fat intake (% energy)	19.1 (5.0)	17.6 (4.3)	15.5 (4.3)	18.4 (4.9)	17.7 (4.4)	16.2 (4.6)
Total red meat intake (serving/day)	0.8 (0.6)	0.8 (0.5)	0.7 (0.6)	0.7 (0.5)	0.8 (0.5)	0.8 (0.6)
Adult BMI	24.7 (5.7)	24.6 (5.3)	24.5 (5.0)	24.6 (5.5)	24.4 (5.1)	24.8 (5.3)
BMI at age 18	21.1 (3.4)	21.3 (3.2)	21.5 (3.4)	21.2 (3.3)	21.2 (3.2)	21.6 (3.5)
Age at first birth (years)	25 (4)	26 (4)	26 (4)	26 (4)	26 (4)	26 (4)
Age at menarche (years)	12.5 (1.4)	12.4 (1.4)	12.4 (1.5)	12.5 (1.4)	12.5 (1.4)	12.3 (1.5)
No (%):						
Current smokers	3555 (20)	1891 (10)	1347 (7)	2470 (14)	2086 (12)	2109 (12)
Current oral contraceptive use	2106 (12)	2016 (11)	1734 (10)	2451 (14)	1977 (11)	1622 (9)
History of benign breast disease	1698 (9)	1755 (10)	1641 (9)	1573 (9)	1716 (9)	1770 (10)
Family history of breast cancer in mother or sisters	2639 (15)	2735 (15)	2862 (16)	2527 (14)	2709 (15)	2848 (16)
Nulliparous	5391 (30)	4445 (25)	4727 (27)	5760 (32)	4396 (25)	4687 (26)

**Table 2 | Age standardized characteristics in 1998 according to intake of fruit and vegetables during adolescence among women enrolled in Nurses' Health Study II. Figures are means (SD) unless otherwise specified**

	Fifths of total fruit intake			Fifths of total vegetable intake		
	1 (n=8850)	3 (n=8830)	5 (n=8840)	1 (n=8869)	3 (n=8829)	5 (n=8869)
Age (years)	43 (5)	43 (5)	43 (5)	43 (5)	43 (5)	43 (5)
Intake in adolescence*:						
Total fruit intake (serving/day)	0.4 (0.2)	1.4 (0.1)	3.2 (1.1)	0.9 (0.7)	1.5 (0.8)	2.4 (1.4)
Total vegetables intake (serving/day)	1.9 (1.0)	2.6 (1.2)	3.9 (1.8)	1.1 (0.3)	2.4 (0.2)	5.0 (1.5)
Total energy intake (kJ)	9781 (3127)	11 407 (2972)	13 397 (3223)	9631 (2993)	11 370 (2955)	13 468 (3298)
Alcohol consumption (g/day)	1.2 (4.2)	1.0 (3.1)	0.9 (3.1)	1.1 (3.9)	1.0 (3.2)	1.0 (3.5)
Total fiber intake (g/day)	17.3 (4.1)	20.5 (4.1)	25.5 (5.8)	17.0 (3.8)	20.5 (3.8)	26.1 (6.0)
Animal fat intake (% energy)	27.7 (6.1)	26.0 (5.3)	23.7 (5.2)	27.4 (6.0)	26.0 (5.4)	24.1 (5.5)
Total red meat intake (serving/day)	1.5 (0.7)	1.5 (0.7)	1.5 (0.7)	1.4 (0.7)	1.5 (0.7)	1.6 (0.8)
BMI at age 18	21.1 (3.2)	21.2 (3.3)	21.2 (3.2)	21.0 (3.2)	21.2 (3.2)	21.4 (3.4)
Age at first birth (years)	26 (5)	27 (5)	27 (5)	26 (5)	26 (5)	27 (5)
Age at menarche (years)	12.4 (1.4)	12.4 (1.4)	12.4 (1.5)	12.5 (1.4)	12.4 (1.4)	12.3 (1.5)
No (%):						
Current smokers	1078 (12)	759 (9)	639 (7)	864 (10)	848 (10)	732 (8)
Current oral contraceptive use	778 (9)	737 (8)	741 (8)	735 (8)	773 (9)	761 (9)
History of benign breast disease	1426 (16)	1343 (15)	1399 (16)	1356 (15)	1431 (16)	1394 (16)
Family history of breast cancer in mother or sisters	1296 (15)	1420 (16)	1493 (17)	1363 (15)	1438 (16)	1518 (17)
Nulliparous	1836 (21)	1736 (20)	1815 (21)	1645 (19)	1722 (20)	2063 (23)

\*Participants recalled their dietary intake during adolescence in 1998.



associations between subgroups of adolescent vegetable intake and risk of cancer, but the power was limited and the confidence intervals were wide (appendix table D). For adolescent diet, the highest intakes of pomes, tropical fruits, grapes, and melon were each associated with lower risk (appendix table E). When we evaluated the associations with intakes of individual fruits and vegetables (fig 1) during adolescence as continuous variables in multivariate models, we observed significant inverse associations for each two servings/week grapes, bananas, and apples and a positive association for other fruit juice, and mixed vegetables.

### Fruit and vegetable intake during early adulthood and risk of breast cancer

Total fruit intake during early adulthood was not significantly associated with a lower risk of breast cancer (hazard ratio (for highest v lowest fifth of intake) 0.96, 95% confidence interval 0.85 to 1.09;  $P_{\text{trend}}=0.46$ ) (table 7). Similar estimates were noted after additional adjustment for red meat intake or modified alternate healthy eating index (appendix table F). Non-significant associations between early adulthood fruit intake and breast cancer were observed among both premenopausal and postmenopausal women at diagnosis (table 7). Risk of

**Table 3 | Hazard ratios and 95% confidence intervals for breast cancer in two models\* according to fifths of intake of fruit during adolescence among women in Nurses' Health Study II**

	Fifth of intake					$P_{\text{trend}}$	Per serving/day
	1	2	3	4	5		
<b>All cases</b>							
Median intake (serving/day)	0.5	0.9	1.4	1.9	2.9	—	—
No of cases/person years	315/134 534	263/134 375	250/134 590	272/134 414	247/134 576	—	—
Model 1	1	0.83 (0.71 to 0.98)	0.79 (0.67 to 0.93)	0.86 (0.73 to 1.01)	0.79 (0.66 to 0.93)	0.02	0.93 (0.87 to 0.99)
Model 2	1	0.82 (0.69 to 0.96)	0.76 (0.64 to 0.90)	0.82 (0.69 to 0.97)	0.75 (0.62 to 0.90)	0.01	0.91 (0.85 to 0.98)
<b>Premenopausal cases</b>							
Median intake (serving/day)	0.5	0.9	1.4	1.9	2.9	—	—
No of cases/person years	135/69 408	105/69 528	123/69 429	104/69 505	109/69 518	—	—
Model 1	1	0.77 (0.60 to 1.00)	0.87 (0.68 to 1.12)	0.75 (0.58 to 0.97)	0.81 (0.62 to 1.04)	0.15	0.93 (0.84 to 1.03)
Model 2	1	0.78 (0.60 to 1.01)	0.87 (0.68 to 1.13)	0.75 (0.57 to 0.99)	0.80 (0.60 to 1.05)	0.17	0.93 (0.83 to 1.03)
<b>Postmenopausal cases</b>							
Median intake (serving/day)	0.5	0.9	1.3	1.9	2.9	—	—
No of cases/person years	149/53 265	134/53 368	108/53 454	138/53 178	111/53 219	—	—
Model 1	1	0.89 (0.71 to 1.13)	0.72 (0.56 to 0.92)	0.92 (0.73 to 1.16)	0.74 (0.57 to 0.95)	0.05	0.91 (0.82 to 1.00)
Model 2	1	0.85 (0.67 to 1.08)	0.67 (0.52 to 0.87)	0.84 (0.66 to 1.07)	0.69 (0.52 to 0.90)	0.02	0.88 (0.79 to 0.98)

\*Model 1: stratified by age in months at state of follow-up and calendar year of current questionnaire cycle. Model 2: as for model 1 and simultaneously adjusted for smoking (never, past, current 1-14/day, current 15-24/day, current  $\geq 25$ /day), race (white/non-white), parity and age at first birth (nulliparous, parity  $\leq 2$  and age at first birth  $< 25$ , parity  $\leq 2$  and age at first birth  $\geq 30$ , parity 3-4 and age at first birth  $< 25$ , parity 3-4 and age at first birth  $25 < 30$ , parity 3-4 and age at first birth  $\geq 30$ , parity  $\geq 5$  and age at first birth  $< 25$ , parity  $\geq 5$  and age at first birth  $\geq 25$ ), height ( $< 157$ ,  $157 < 165$ ,  $165 < 173$ ,  $\geq 173$  cm), BMI at age 18 ( $< 18.5$ ,  $18.5 < 22.5$ ,  $22.5 < 25$ ,  $25.0 < 30$ ,  $\geq 30.0$ ), weight change since age 18 (continuous, missing indicator), age at menarche ( $< 12$ , 12, 13,  $\geq 14$ ), family history of breast cancer (yes, no), history of benign breast disease (yes, no), oral contraceptive use (never, past, current), adolescent alcohol intake (non-drinker,  $< 5$ ,  $\geq 5$  g/day), adult alcohol intake (non-drinker,  $< 5$ ,  $5 < 15$ ,  $\geq 15$  g/day), adolescent energy intake (fifth). In postmenopausal women, we additionally adjusted for hormone use (postmenopausal never users, postmenopausal past users, postmenopausal current users), age at menopause ( $< 45$ , 45-46, 47-48, 49-50, 51-52,  $\geq 53$ ). Among all women, we additionally adjusted for hormone use and menopausal status (premenopausal, postmenopausal never users, postmenopausal past users, postmenopausal current users, unknown menopausal status) and age at menopause (premenopausal, unknown menopause,  $< 45$ , 45-46 years, 47-48, 49-50, 51-52,  $\geq 53$ ).

**Table 4 | Hazard ratios and 95% confidence intervals for breast cancer in two models\* according to fifths of intake of vegetables during adolescence among women in Nurses' Health Study II**

	Fifth of intake					$P_{\text{trend}}$	Per serving/day
	1	2	3	4	5		
<b>All cases</b>							
Median intake (serving/day)	1.2	1.8	2.4	3.2	4.5	—	—
No of cases/person years	310/134 501	255/134 513	248/134 563	264/134 493	270/134 419	—	—
Model 1	1	0.83 (0.70 to 0.98)	0.81 (0.68 to 0.95)	0.85 (0.72 to 1.00)	0.86 (0.73 to 1.01)	0.20	0.97 (0.93 to 1.02)
Model 2	1	0.82 (0.69 to 0.97)	0.79 (0.66 to 0.93)	0.83 (0.70 to 0.98)	0.85 (0.71 to 1.01)	0.21	0.97 (0.92 to 1.02)
<b>Premenopausal cases</b>							
Median intake (serving/day)	1.2	1.8	2.4	3.2	4.5	—	—
No of cases/person years	138/69 460	110/69 410	111/69 487	99/69 487	118/69 544	—	—
Model 1	1	0.78 (0.61 to 1.01)	0.80 (0.62 to 1.03)	0.69 (0.53 to 0.90)	0.82 (0.64 to 1.05)	0.14	0.95 (0.88 to 1.02)
Model 2	1	0.80 (0.62 to 1.04)	0.81 (0.62 to 1.05)	0.71 (0.54 to 0.94)	0.84 (0.64 to 1.10)	0.25	0.95 (0.88 to 1.03)
<b>Postmenopausal cases</b>							
Median intake (serving/day)	1.2	1.9	2.5	3.2	4.6	—	—
No of cases/person years	147/53 314	122/53 335	116/53 319	137/53 289	118/53 225	—	—
Model 1	1	0.83 (0.65 to 1.06)	0.78 (0.61 to 1.00)	0.93 (0.73 to 1.17)	0.80 (0.62 to 1.02)	0.21	0.96 (0.90 to 1.03)
Model 2	1	0.79 (0.62 to 1.02)	0.73 (0.57 to 0.94)	0.87 (0.68 to 1.11)	0.75 (0.57 to 0.98)	0.13	0.94 (0.88 to 1.02)

\*See table 3.

breast cancer was not significantly associated with total vegetable, total fruit and vegetable, or fruit juice intake during early adulthood (tables 8-10). We did not detect any significant non-linearity for fruits or vegetables intake in early adulthood in relation to risk of breast cancer ( $P=0.85$  for non-linearity for total fruits and  $P=0.56$  for non-linearity for total vegetables). For early adult diet, the highest intakes of yellow/orange vegetables (0.89, 0.79 to 0.99;  $P_{\text{trend}}=0.16$ ) and fruits and vegetables rich in  $\alpha$  carotene (0.89, 0.80 to 1.00;  $P_{\text{trend}}=0.28$ ) were each associated with lower risk but tests for trend were not significant. Stronger associations were observed among women who were premenopausal at diagnosis (0.85 (0.72 to 0.99;  $P_{\text{trend}}=0.11$ ) for yellow/orange vegetables; 0.82 (0.70 to 0.96;  $P_{\text{trend}}=0.06$ ) for fruits and vegetables rich in  $\alpha$  carotene) (appendix table G). We did not observe any significant associations between subgroups of early adulthood fruit intake and risk (appendix H). When evaluating the associations

with intakes of individual fruits and vegetables (fig 1) during early adulthood in multivariate models, we observed significant inverse associations for each two servings/week of oranges and kale and a positive association for orange juice.

#### Average fruit and vegetable intake during adolescence and early adulthood and risk of breast cancer

Fruit intake in adolescence and early adulthood (1991) was moderately correlated ( $r=0.43$ ), as was vegetable intake in adolescence and early adulthood ( $r=0.42$ ). Among women with both early adulthood and adolescent dietary data ( $n=41\,032$ ), there were non-significant associations between average of total fruit intake and overall breast cancer (hazard ratio (for highest v lowest fifth of intake) 0.88, 95% confidence interval 0.73 to 1.06;  $P_{\text{trend}}=0.08$ ), premenopausal breast cancer (0.82, 0.61 to 1.10;  $P_{\text{trend}}=0.03$ ), and postmenopausal breast

**Table 5 | Hazard ratios and 95% confidence intervals for breast cancer in two models\* according to fifths of intake of total fruit and vegetables during adolescence among women in Nurses' Health Study II**

	Fifth of intake					$P_{\text{trend}}$	Per serving/day
	1	2	3	4	5		
<b>All cases</b>							
Median intake (serving/day)	1.9	3.0	3.9	5.0	7.1	—	—
No of cases/person years	307/134 478	264/134 563	249/134 499	266/134 496	261/134 452	—	—
Model 1	1	0.86 (0.73 to 1.02)	0.81 (0.69 to 0.96)	0.86 (0.73 to 1.01)	0.84 (0.71 to 0.99)	0.09	0.97 (0.94 to 1.00)
Model 2	1	0.85 (0.72 to 1.00)	0.79 (0.66 to 0.93)	0.82 (0.69 to 0.98)	0.82 (0.68 to 0.98)	0.07	0.97 (0.94 to 1.00)
<b>Premenopausal cases</b>							
Median intake (serving/day)	1.9	3.0	3.9	5.0	7.0	—	—
No of cases/person years	134/69 418	127/69 475	96/69 472	112/69 480	107/69 543	—	—
Model 1	1	0.94 (0.73 to 1.20)	0.72 (0.55 to 0.93)	0.81 (0.63 to 1.04)	0.77 (0.60 to 1.00)	0.04	0.95 (0.91 to 1.00)
Model 2	1	0.93 (0.72 to 1.19)	0.71 (0.54 to 0.94)	0.80 (0.61 to 1.05)	0.76 (0.57 to 1.01)	0.05	0.95 (0.90 to 1.00)
<b>Postmenopausal cases</b>							
Median intake (serving/day)	1.9	3.0	3.9	5.0	7.1	—	—
No of cases/person years	143/53 312	119/53 366	130/53 333	128/53 278	120/53 195	—	—
Model 1	1	0.83 (0.65 to 1.07)	0.91 (0.71 to 1.15)	0.89 (0.70 to 1.14)	0.83 (0.65 to 1.06)	0.27	0.98 (0.93 to 1.02)
Model 2	1	0.81 (0.63 to 1.04)	0.85 (0.66 to 1.09)	0.83 (0.64 to 1.08)	0.79 (0.60 to 1.04)	0.17	0.97 (0.92 to 1.02)

\*See table 3.

**Table 6 | Hazard ratios and 95% confidence intervals for breast cancer in two models\* according to fifths of intake of fruit juice during adolescence among women in Nurses' Health Study II**

	Fifth of intake					$P_{\text{trend}}$	Per serving/day
	1	2	3	4	5		
<b>All cases</b>							
Median intake (serving/day)	0.1	0.4	0.9	1.5	2.1	—	—
No of cases/person years	285/144 750	269/133 916	237/125 374	267/129 621	288/138 185	—	—
Model 1	1	1.03 (0.87 to 1.22)	0.99 (0.83 to 1.18)	1.05 (0.89 to 1.25)	1.07 (0.91 to 1.26)	0.39	1.03 (0.96 to 1.11)
Model 2	1	1.02 (0.86 to 1.20)	0.96 (0.80 to 1.14)	1.00 (0.84 to 1.19)	1.02 (0.86 to 1.21)	0.85	1.01 (0.93 to 1.09)
<b>Premenopausal cases</b>							
Median intake (serving/day)	0.1	0.4	0.9	1.5	2.2	—	—
No of cases/person years	107/71 645	123/68 496	108/67 377	126/69 940	111/69 591	—	—
Model 1	1	1.22 (0.94 to 1.59)	1.10 (0.84 to 1.44)	1.17 (0.90 to 1.52)	1.06 (0.81 to 1.38)	0.96	1.00 (0.90 to 1.12)
Model 2	1	1.25 (0.96 to 1.62)	1.11 (0.84 to 1.46)	1.17 (0.90 to 1.54)	1.06 (0.80 to 1.41)	1.00	1.00 (0.89 to 1.12)
<b>Postmenopausal cases</b>							
Median intake (serving/day)	0.1	0.3	0.9	1.5	2.1	—	—
No of cases/person years	121/47 907	141/65 884	111/48 809	119/51 122	148/52 498	—	—
Model 1	1	0.84 (0.66 to 1.07)	0.91 (0.71 to 1.18)	0.92 (0.71 to 1.19)	1.08 (0.84 to 1.37)	0.19	1.07 (0.97 to 1.19)
Model 2	1	0.83 (0.65 to 1.06)	0.87 (0.67 to 1.14)	0.86 (0.66 to 1.12)	1.02 (0.79 to 1.32)	0.44	1.04 (0.94 to 1.17)

\*See table 3.

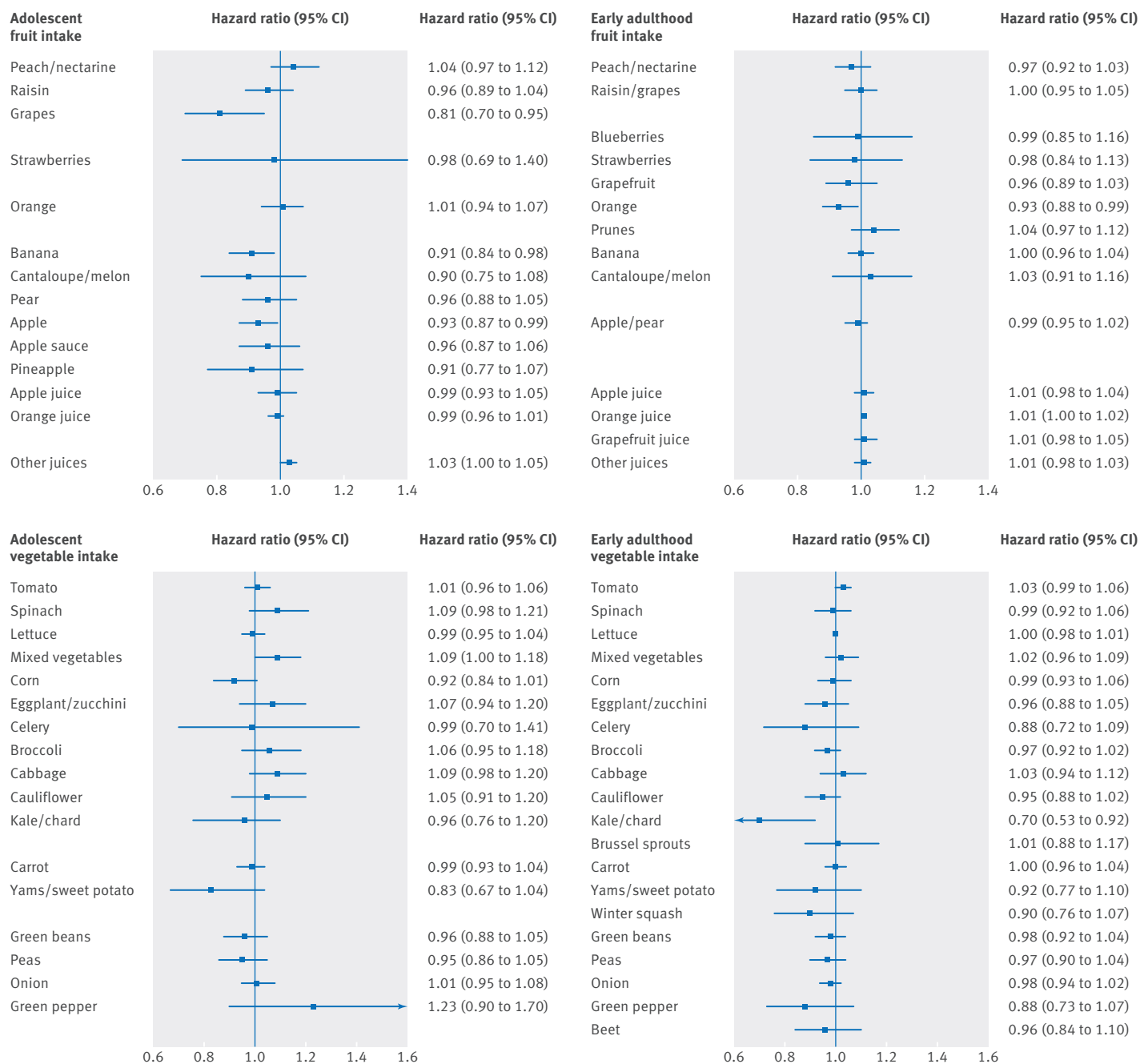


Fig 1 | Multivariate hazard ratios for every two servings/week of specific fruits and vegetables during adolescence and early adulthood and risk of breast cancer

cancer (0.91, 0.69 to 1.19;  $P_{\text{trend}}=0.55$ ). Average of adolescent and early adulthood vegetable intake was associated with lower risk of premenopausal breast cancer (0.74, 0.55 to 0.98;  $P_{\text{trend}}=0.05$ ) but not lower risk of overall breast cancer (0.85, 0.71 to 1.03;  $P_{\text{trend}}=0.18$ ) or postmenopausal breast cancer (0.93, 0.71 to 1.22;  $P_{\text{trend}}=0.63$ ).

#### Cumulative average of premenopausal fruit and vegetable intake and risk of breast cancer

When we examined cumulative average of premenopausal fruit intake, we observed similar risk estimates among all women (hazard ratio (for highest v lowest fifth of intake) 0.91, 95% confidence interval 0.81 to 1.03;

$P_{\text{trend}}=0.05$ ), for premenopausal breast cancer (0.96, 0.82 to 1.14;  $P_{\text{trend}}=0.57$ ), and for postmenopausal breast cancer (0.87, 0.71 to 1.07;  $P_{\text{trend}}=0.12$ ). Cumulative average of premenopausal vegetable intake was not associated with risk of breast cancer (data not shown). Cumulative average of fruit juice intake before menopause, however, was positively associated with risk of premenopausal breast cancer (1.18, 1.01 to 1.39;  $P_{\text{trend}}=0.06$ ).

#### Sensitivity analyses

The difference in the associations for tumors positive for both estrogen and progesterone receptors, negative

for both estrogen and progesterone receptors, and estrogen receptor positive and progesterone receptor negative was significant only for adolescent but not early adulthood total fruit intake (table 11). The inverse association with adolescent fruit intake was stronger for both estrogen and progesterone receptor negative cancers (hazard ratio 0.70, 95% confidence interval

0.57 to 0.86, per serving/day) compared with both estrogen and progesterone receptor positive cancers (0.91, 0.83 to 1.00) and estrogen receptor positive and progesterone receptor negative (1.03, 0.82 to 1.29) ( $P_{\text{heterogeneity}}=0.02$ ). Associations between total vegetable intake and risk of breast cancer did not differ significantly by estrogen and progesterone receptor positive

**Table 7 | Hazard ratios and 95% confidence intervals for breast cancer in two models\* according to fifths of intake of fruit during early adulthood among women in the Nurses' Health Study II**

	Fifth of intake					P <sub>trend</sub> †	Per serving/day
	1	2	3	4	5		
<b>All cases</b>							
Median intake (serving/day)	0.2	0.5	0.9	1.3	2.1	—	—
No of cases/person years	641/375 096	659/378 629	657/377 373	642/377 000	636/376 894	—	—
Model 1	1	1.01 (0.91 to 1.13)	1.00 (0.90 to 1.11)	0.97 (0.87 to 1.08)	0.96 (0.86 to 1.07)	0.29	0.97 (0.92 to 1.02)
Model 2	1	0.99 (0.89 to 1.11)	0.98 (0.88 to 1.10)	0.96 (0.85 to 1.07)	0.96 (0.85 to 1.09)	0.46	0.98 (0.92 to 1.04)
<b>Premenopausal cases</b>							
Median intake (serving/day)	0.2	0.5	0.9	1.3	2.1	—	—
No of cases/person years	340/237 032	345/234 894	339/236 114	342/236 389	353/235 939	—	—
Model 1	1	1.00 (0.86 to 1.16)	0.96 (0.82 to 1.12)	0.96 (0.82 to 1.11)	0.98 (0.85 to 1.14)	0.79	0.99 (0.92 to 1.06)
Model 2	1	0.99 (0.85 to 1.15)	0.96 (0.82 to 1.12)	0.96 (0.82 to 1.12)	0.99 (0.84 to 1.17)	0.94	1.00 (0.92 to 1.08)
<b>Postmenopausal cases</b>							
Median intake (serving/day)	0.2	0.5	0.9	1.3	2.1	—	—
No of cases/person years	237/105 846	231/106 794	237/105 759	245/106 175	219/106 040	—	—
Model 1	1	0.96 (0.80 to 1.15)	0.99 (0.83 to 1.19)	1.01 (0.84 to 1.21)	0.91 (0.75 to 1.09)	0.39	0.96 (0.88 to 1.05)
Model 2	1	0.95 (0.79 to 1.15)	0.98 (0.81 to 1.18)	1.00 (0.83 to 1.21)	0.91 (0.74 to 1.11)	0.46	0.97 (0.88 to 1.06)

\*Model 1 stratified by age in months at state of follow-up and calendar year of current questionnaire cycle. Model 2 as for model 1 and simultaneously adjusted for race (white, non-white), family history of breast cancer in mother or sisters (yes, no), history of benign breast disease (yes, no), smoking (never, past, current 1-14/day, current 15-24/day, current  $\geq 25$ /day), height (<157, 157-165, 165-173,  $\geq 173$  cm), BMI at age 18 (<18.5, 18.5-22.5, 22.5-25, 25.0-30,  $\geq 30.0$ ), weight change since age 18 (continuous, missing indicator), age at menarche (<12, 12, 13,  $\geq 14$ ), parity and age at first birth (nulliparous, parity  $\leq 2$  and age at first birth <25, parity  $\leq 2$  and age at first birth 25-30, parity  $\leq 2$  and age at first birth  $\geq 30$ , parity 3-4 and age at first birth <25, parity 3-4 and age at first birth 25-30, parity 3-4 and age at first birth  $\geq 30$ , parity  $\geq 5$  and age at first birth <25, parity  $\geq 5$  and age at first birth  $\geq 25$ ), oral contraceptive use (never, past, current), alcohol intake (non-drinker, <5, 5-15,  $\geq 15$  g/day), and energy (fifth). In postmenopausal women, we additionally adjusted for hormone use (postmenopausal never users, postmenopausal past users, postmenopausal current users), age at menopause (<45, 45-46, 47-48, 49-50, 51-52,  $\geq 53$ ). Among all women, we additionally adjusted for hormone use and menopausal status (premenopausal, postmenopausal never users, postmenopausal past users, postmenopausal current users, unknown menopausal status) and age at menopause (premenopausal, unknown menopause, <45, 45-46, 47-48, 49-50, 51-52,  $\geq 53$ ).

†P<sub>trend</sub> calculated with median intake of each variable in each fifth as continuous variable.

**Table 8 | Hazard ratios and 95% confidence intervals for breast cancer in two models\* according to fifths of intake of vegetables during early adulthood among women in the Nurses' Health Study II**

	Fifth of intake					P <sub>trend</sub> †	Per serving/day
	1	2	3	4	5		
<b>All cases</b>							
Median intake (serving/day)	1.1	1.8	2.5	3.3	4.8	—	—
No of cases/person years	639/376 975	626/377 037	691/376 930	613/377 098	666/376 953	—	—
Model 1	1	0.95 (0.85 to 1.06)	1.04 (0.93 to 1.16)	0.90 (0.81 to 1.01)	0.96 (0.86 to 1.07)	0.35	0.99 (0.96 to 1.01)
Model 2	1	0.92 (0.82 to 1.03)	1.01 (0.90 to 1.13)	0.89 (0.79 to 1.00)	0.97 (0.86 to 1.09)	0.62	0.99 (0.96 to 1.02)
<b>Premenopausal cases</b>							
Median intake (serving/day)	1.1	1.8	2.4	3.2	4.8	—	—
No of cases/person years	354/235 931	300/235 984	366/236 006	349/236 218	350/236 229	—	—
Model 1	1	0.82 (0.70 to 0.95)	0.97 (0.84 to 1.13)	0.92 (0.80 to 1.07)	0.90 (0.78 to 1.05)	0.59	0.99 (0.95 to 1.03)
Model 2	1	0.79 (0.68 to 0.93)	0.95 (0.82 to 1.11)	0.91 (0.78 to 1.07)	0.90 (0.76 to 1.06)	0.67	0.99 (0.95 to 1.03)
<b>Postmenopausal cases</b>							
Median intake (serving/day)	1.1	1.9	2.5	3.3	4.9	—	—
No of cases/person years	234/106 123	253/106 190	222/106 187	221/106 073	239/106 042	—	—
Model 1	1	1.05 (0.88 to 1.26)	0.93 (0.77 to 1.12)	0.93 (0.77 to 1.11)	0.99 (0.82 to 1.18)	0.58	0.99 (0.94 to 1.03)
Model 2	1	1.01 (0.84 to 1.21)	0.90 (0.74 to 1.09)	0.89 (0.73 to 1.09)	0.97 (0.80 to 1.19)	0.64	0.99 (0.94 to 1.04)

\*See table 7.

†P<sub>trend</sub> calculated with median intake of each variable in each fifth as continuous variable.



or negative status for either adolescent or early adulthood diets (table 11).

We also examined whether the association between fruit or vegetable intake and risk of breast cancer differed by levels of BMI at age 18. No significant interaction was observed between BMI at age 18 and either adolescent or early adulthood fruit and vegetable intake (data not shown). Also, we found no significant interaction between age at baseline and early adulthood fruit and vegetable intake (data not shown).

### Discussion

This prospective study provides evidence that greater consumption of fruits during adolescence is associated with lower risk of breast cancer. Furthermore, higher intake of fruits and vegetables rich in  $\alpha$  carotene during early adulthood was specifically associated with lower

risk. The associations with breast cancer differed significantly among individual fruits and vegetables: greater consumption of apple, banana, and grapes during adolescence, as well as oranges and kale during early adulthood, was significantly associated with a reduced risk.

### Strengths and weaknesses of the study

Our study has several strengths. To evaluate the importance of timing, we assessed the association between fruit and vegetable intake during specific life periods (adolescence, early adulthood, and cumulative average of the premenopausal period). The relatively large number of cases gave adequate power for us to detect modest differences in risk as well as examination of breast cancers by menopausal and hormone receptor status. Although a question was recently raised about

**Table 9 | Hazard ratios and 95% confidence intervals for breast cancer in two models\* according to fifths of intake of total fruit and vegetables during early adulthood among women in the Nurses' Health Study II**

	Fifth of intake					P <sub>trend</sub> †	Per serving/day
	1	2	3	4	5		
<b>All cases</b>							
Median intake (serving/day)	1.6	2.6	3.5	4.5	6.6	—	—
No of cases/person years	635/376 997	639/377 028	671/377 035	656/376 991	634/376 942	—	—
Model 1	1	0.98 (0.88 to 1.09)	1.02 (0.91 to 1.13)	0.97 (0.87 to 1.09)	0.93 (0.83 to 1.04)	0.17	0.99 (0.97 to 1.01)
Model 2	1	0.95 (0.85 to 1.06)	0.99 (0.88 to 1.11)	0.95 (0.85 to 1.07)	0.93 (0.82 to 1.06)	0.33	0.99 (0.97 to 1.01)
<b>Premenopausal cases</b>							
Median intake (serving/day)	1.6	2.6	3.4	4.5	6.5	—	—
No of cases/person years	346/235 935	313/236 063	364/236 124	341/236 115	355/236 132	—	—
Model 1	1	0.87 (0.74 to 1.01)	0.99 (0.86 to 1.15)	0.91 (0.78 to 1.06)	0.94 (0.81 to 1.09)	0.72	1.00 (0.97 to 1.02)
Model 2	1	0.85 (0.73 to 1.00)	0.98 (0.84 to 1.14)	0.90 (0.77 to 1.06)	0.95 (0.80 to 1.12)	0.87	1.00 (0.97 to 1.03)
<b>Postmenopausal cases</b>							
Median intake (serving/day)	1.6	2.6	3.5	4.6	6.7	—	—
No of cases/person years	230/106 264	246/106 040	227/106 137	237/106 137	229/106 037	—	—
Model 1	1	1.06 (0.88 to 1.27)	0.98 (0.81 to 1.18)	1.02 (0.85 to 1.22)	0.97 (0.81 to 1.17)	0.60	0.99 (0.96 to 1.02)
Model 2	1	1.02 (0.85 to 1.23)	0.96 (0.79 to 1.16)	0.99 (0.81 to 1.20)	0.97 (0.79 to 1.19)	0.71	0.99 (0.96 to 1.03)

\*See table 7.

†P<sub>trend</sub> calculated with median intake of each variable in each fifth as continuous variable.

**Table 10 | Hazard ratios and 95% confidence intervals for breast cancer in two models\* according to fifths of intake of fruit juice during early adulthood among women in the Nurses' Health Study II**

	Fifth of intake					P <sub>trend</sub> †	Per serving/day
	1	2	3	4	5		
<b>All cases</b>							
Median intake (serving/day)	0	0.2	0.6	1.2	2.0	—	—
No of cases/person years	659/399 290	631/359 799	602/363 546	680/381 851	661/379 702	—	—
Model 1	1	1.11 (0.99 to 1.24)	1.04 (0.93 to 1.17)	1.13 (1.01 to 1.25)	1.13 (1.01 to 1.26)	0.05	1.05 (1.00 to 1.10)
Model 2	1	1.10 (0.98 to 1.23)	1.03 (0.92 to 1.15)	1.10 (0.98 to 1.23)	1.11 (0.99 to 1.25)	0.13	1.04 (0.99 to 1.10)
<b>Premenopausal cases</b>							
Median intake (serving/day)	0	0.2	0.6	1.2	2.1	—	—
No of cases/person years	325/237 972	333/237 215	355/233 474	352/233 467	353/237 779	—	—
Model 1	1	1.07 (0.91 to 1.24)	1.13 (0.97 to 1.32)	1.13 (0.97 to 1.32)	1.12 (0.97 to 1.31)	0.15	1.05 (0.98 to 1.12)
Model 2	1	1.07 (0.91 to 1.25)	1.12 (0.96 to 1.31)	1.12 (0.96 to 1.31)	1.12 (0.95 to 1.32)	0.21	1.04 (0.98 to 1.12)
<b>Postmenopausal cases</b>							
Median intake (serving/day)	0	0.2	0.5	1.2	2.0	—	—
No of cases/person (years)	262/122 468	207/88 494	225/106 537	247/113 306	227/99 537	—	—
Model 1	1	1.12 (0.93 to 1.34)	1.00 (0.83 to 1.19)	1.03 (0.87 to 1.23)	1.10 (0.92 to 1.32)	0.49	1.03 (0.95 to 1.12)
Model 2	1	1.11 (0.93 to 1.34)	0.98 (0.82 to 1.18)	1.02 (0.85 to 1.22)	1.08 (0.89 to 1.30)	0.68	1.02 (0.93 to 1.11)

\*See table 7.

†P<sub>trend</sub> calculated with median intake of each variable in each fifth as continuous variable.

**Table 11 | Adolescent and early adult intake of fruits and vegetables in relation to risk of breast cancer by estrogen and progesterone receptor status among women in Nurses' Health Study II**

Breast cancer subtype	No of cases	HR (95% CI)*	P value for heterogeneity
<b>Adolescent total fruit intake (serving/day)</b>			
Estrogen and progesterone receptor positive	722	0.91 (0.83 to 1.00)	0.02
Estrogen and progesterone receptor negative	165	0.70 (0.57 to 0.86)	
Estrogen receptor positive and progesterone receptor negative	111	1.03 (0.82 to 1.29)	
<b>Early adulthood total fruit intake (serving/day)</b>			
Estrogen and progesterone receptor positive	1634	0.95 (0.88 to 1.03)	0.78
Estrogen and progesterone receptor negative	435	0.98 (0.85 to 1.13)	
Estrogen receptor positive and progesterone receptor negative	251	0.90 (0.74 to 1.09)	
<b>Adolescent total vegetable intake (serving/day)</b>			
Estrogen and progesterone receptor positive	722	0.93 (0.87 to 1.00)	0.19
Estrogen and progesterone receptor negative	165	0.91 (0.79 to 1.05)	
Estrogen receptor positive and progesterone receptor negative	111	1.09 (0.93 to 1.28)	
<b>Early adulthood total vegetable intake (serving/day)</b>			
Estrogen and progesterone receptor positive	1634	0.97 (0.93 to 1.01)	0.91
Estrogen and progesterone receptor negative	435	0.97 (0.90 to 1.05)	
Estrogen receptor positive and progesterone receptor negative	251	0.99 (0.90 to 1.10)	

\*Multivariable model stratified by age in months at start of follow-up and calendar year of current questionnaire cycle and was simultaneously adjusted for race (white, non-white), family history of breast cancer in mother or sisters (yes, no), history of benign breast disease (yes, no), smoking (never, past, current 1-14/day, current 15-24/day, current  $\geq 25$ /day), height (<157, 157-165, 165-173,  $\geq 173$  cm), BMI at age 18 (<18.5, 18.5-22.5, 22.5-25, 25.0-30,  $\geq 30.0$ ), weight change since age 18 (continuous, missing indicator), age at menarche (<12, 12, 13,  $\geq 14$ ), parity and age at first birth (nulliparous, parity  $\leq 2$  and age at first birth <25, parity  $\leq 2$  and age at first birth 25-30, parity  $\leq 2$  and age at first birth  $\geq 30$ , parity 3-4 and age at first birth <25, parity 3-4 and age at first birth 25-30, parity 3-4 and age at first birth  $\geq 30$ , parity  $\geq 5$  and age at first birth <25, parity  $\geq 5$  and age at first birth  $\geq 25$ ), oral contraceptive use (never, past, current), alcohol intake (non-drinker, <5, 5-15,  $\geq 15$  g/day), and energy (fifth). In postmenopausal women, we additionally adjusted for hormone use (postmenopausal never users, postmenopausal past users, postmenopausal current users), age at menopause (<45, 45-46, 47-48, 49-50, 51-52,  $\geq 53$ ). Among all women, we additionally adjusted for hormone use and menopausal status (premenopausal, postmenopausal never users, postmenopausal past users, postmenopausal current users, unknown menopausal status) and age at menopause (premenopausal, unknown menopause, <45, 45-46, 47-48, 49-50, 51-52,  $\geq 53$ ). For adolescent fruit and vegetable intake, we additionally adjusted for adolescent alcohol intake (non-drinker, <5,  $\geq 5$  g/day) and adolescent energy intake (instead of adult energy intake).

the value of dietary assessments from self reported information,<sup>36</sup> and refuted elsewhere,<sup>37</sup> the questionnaire we used has documented validity,<sup>21-23</sup> as shown by strong correlations between nutrient intakes assessed by the food frequency questionnaire and by weighed diet records and by many different biomarkers of diet.<sup>21</sup> Also, higher intake of fruits and vegetables assessed by the questionnaire we used has predicted lower risks of diabetes<sup>22</sup> and coronary heart disease.<sup>23</sup> Further, intakes of specific fruits and vegetables assessed by this questionnaire robustly predict plasma carotenoid concentrations.<sup>38</sup>

Potential limitations also need to be considered. First, the participants were restricted to nurses and predominantly white women, which could reduce generalizability; however, it is unlikely that the biology underlying these associations differs by race or education.<sup>39,40</sup> Second, adolescent diet might be misclassified because assessments were done when women were aged 33-52. The associations, however, were largely independent of adult diet, and comparisons of diet recorded during adolescence and our questionnaire administered 10 years later showed good validity.<sup>29</sup> Further, because diet was assessed before diagnosis, misclassification would tend to obscure, not cause, associations. Third, residual confounding is possible as women who eat a lot of fruit and vegetables might have healthier lifestyles. Adjustment for potential confounders, including other aspects of diet, however, had minimal effects on associations. Fourth, our observation of borderline significant association between early adulthood intake of fruits and vegetables rich in  $\alpha$  carotene and breast cancer warrants

cautious interpretation. Fifth, while adolescent fruit consumption was inversely associated with risk, early adulthood fruit consumption was non-significantly inversely associated with risk. Whereas the results are not entirely consistent, they are in line with the hypothesis that breast tissue is more vulnerable to carcinogenic exposure in earlier life than later. Sixth, we made multiple comparisons (different food groups; premenopausal and postmenopausal subgroups, and subtype of tumors) in this analysis, and we cannot exclude the possibility of type I errors. The central finding of an association with fruits and vegetables, however, was the primary a priori hypothesis, the other analyses were directed to an understanding of this association.

#### Results in relation to other studies

Fruit and vegetable intake has been hypothesized to protect against breast cancer, but the overall evidence is conflicting. Although the European Prospective Investigation into Cancer and Nutrition (EPIC) study with about 330 000 women from 10 European countries found no association between consumption of fruits and vegetables and risk of breast cancer,<sup>16</sup> a recent pooled analysis including 993 466 women from 20 prospective cohort studies, which included NHSII, suggests that the associations with fruit and vegetable intake varied depended on breast cancer hormone receptor status; modest inverse associations were apparent only for vegetable intake and estrogen receptor negative breast cancer.<sup>17</sup> Our current results show that high fruit intake during adolescence is associated with reduced overall risk, with stronger results for both estrogen and progesterone receptors negative tumors.

### Biological plausibility

Several bioactive components in fruits and vegetables, such as carotenoids, vitamin C, flavonoids, fiber, magnesium, and potassium, could act through multiple mechanisms to promote a beneficial effect on risk of breast cancer.<sup>1-41</sup> Our own study using data from NHSII indicated that high intake of fiber during adolescence and early adulthood was associated with decreased risk in later life.<sup>3</sup> Furthermore, a meta-analysis of 16 prospective studies showed an inverse association between fiber intake and risk.<sup>41</sup> We found differences in the results for adolescent fruit intake when we additionally adjusted for fiber. The association, however, was still significant and suggests there might be other mechanisms in addition to fiber through which fruit consumption could reduce risk. Further, we noted inverse associations for intake of fruits and vegetables rich in  $\alpha$  carotene during early adulthood and risk of breast cancer. The Nurses' Health Study also reported a lower risk of estrogen receptor negative breast cancer with high intake of yellow/orange vegetables.<sup>42</sup> Although individual studies did not observe an association between dietary and circulating vitamin C or carotenoids and breast cancer,<sup>43-45</sup> a recent pooled analysis of eight prospective studies found that higher circulating levels of carotenoids were associated with decreased risk, with stronger associations for estrogen receptor negative tumors.<sup>2</sup> Further analyses of dietary  $\alpha$  and  $\beta$  carotene intake in another pooled analysis suggested protective effects of carotenoids for estrogen receptor negative breast cancer.<sup>46</sup> Our findings suggest that high intake of fruits and vegetables rich in carotenoids in early adulthood could be associated with a decreased risk of breast cancer.

### Conclusions

In summary, our findings suggest that high intake of fruit during adolescence could be associated with a lower risk of breast cancer. The consumption of fruits and vegetables rich in  $\alpha$  carotene might be particularly important. We also observed that the associations between adolescent fruit intake and breast cancer differed by hormone receptor status, with a stronger inverse association for both estrogen and progesterone receptor negative cancer. We noted that 2.9 servings/day fruit consumption (highest fifth) during adolescent compared with 0.5 serving/day (lowest fifth) was associated with about a 25% lower risk of breast cancer. If this apparent risk reduction is applied to lifetime risk of breast cancer, the absolute number of breast cancers that could potentially be prevented by higher intake of fruits would be substantial. Also, in decisions about fruit consumption, other consequences should be considered, and these include lower risks of diabetes and cardiovascular disease. Our findings are in line with cancer prevention recommendations<sup>47</sup> to consume more fruits and vegetables and support the adoption of these behaviors in early life. Further study of the relation between diet in early life and risk of breast cancer is needed.

The study sponsors were not involved in the study design and collection, analysis and interpretation of data, or the writing of the article or the decision to submit it for publication. The authors were independent from study sponsors. We thank the participants and staff of the NHS II for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY.

**Contributors:** All authors designed the research, provided critical input in the writing of the manuscript, and read and approved the final version of manuscript. MSF carried out the analysis and wrote the manuscript. MSF and AHE are guarantors.

**Funding:** The study was supported by the National Institutes of Health Grants (R01 CA050385, UM1 CA176726) and a grant from The Breast Cancer Research Foundation. MSF was supported by the Japan Pharmaceutical Manufacturers Association.

**Competing interests:** All authors have completed the ICMJE uniform disclosure form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author) and declare: no support from any organization for the submitted work; no financial relationships with any organizations 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.

**Ethical approval:** The study protocol (IRB protocol No 1999-P-003389) was approved by the institutional review boards of Brigham and Women's Hospital and Harvard School of Public Health (Boston, MA). The completion of the self administered questionnaire was considered to imply informed consent.

**Data sharing:** No additional data available.

**Transparency:** The lead authors affirm 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.

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) 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 and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>.

- 1 World Cancer Research Fund/American Institute for Cancer Research. *Food, Nutrition, Physical Activity and the Prevention of Cancer: A Global Perspective*. AICR, 2007.
- 2 Eliassen AH, Hendrickson SJ, Brinton LA, et al. Circulating carotenoids and risk of breast cancer: pooled analysis of eight prospective studies. *J Natl Cancer Inst* 2012;104:1905-16. doi:10.1093/jnci/djs461.
- 3 Farvid MS, Eliassen AH, Cho E, Liao X, Chen WY, Willett WC. Dietary fiber intake in young adults and breast cancer risk. *Pediatrics* 2016;137:1-11. doi:10.1542/peds.2015-1226.
- 4 Shibata A, Paganini-Hill A, Ross RK, Henderson BE. Intake of vegetables, fruits, beta-carotene, vitamin C and vitamin supplements and cancer incidence among the elderly: a prospective study. *Br J Cancer* 1992;66:673-9. doi:10.1038/bjc.1992.336.
- 5 Rohan TE, Howe GR, Friedenreich CM, Jain M, Miller AB. Dietary fiber, vitamins A, C, and E, and risk of breast cancer: a cohort study. *Cancer Causes Control* 1993;4:29-37. doi:10.1007/BF00051711.
- 6 Verhoeven DT, Assen N, Goldbohm RA, et al. Vitamins C and E, retinol, beta-carotene and dietary fibre in relation to breast cancer risk: a prospective cohort study. *Br J Cancer* 1997;75:149-55. doi:10.1038/bjc.1997.25.
- 7 Zhang S, Hunter DJ, Forman MR, et al. Dietary carotenoids and vitamins A, C, and E and risk of breast cancer. *J Natl Cancer Inst* 1999;91:547-56. doi:10.1093/jnci/91.6.547.
- 8 Olsen A, Tjønneland A, Thomsen BL, et al. Fruits and vegetables intake differentially affects estrogen receptor negative and positive breast cancer incidence rates. *J Nutr* 2003;133:2342-7.
- 9 van Gils CH, Peeters PH, Bueno-de-Mesquita HB, et al. Consumption of vegetables and fruits and risk of breast cancer. *JAMA* 2005;293:183-93. doi:10.1001/jama.293.2.183.
- 10 Sonestedt E, Borgquist S, Ericson U, et al. Plant foods and oestrogen receptor alpha- and beta-defined breast cancer: observations from the Malmo Diet and Cancer cohort. *Carcinogenesis* 2008;29:2203-9. doi:10.1093/carcin/bgn196.
- 11 George SM, Park Y, Leitzmann MF, et al. Fruit and vegetable intake and risk of cancer: a prospective cohort study. *Am J Clin Nutr* 2009;89:347-53. doi:10.3945/ajcn.2008.26722.
- 12 Butler LM, Wu AH, Wang R, Koh WP, Yuan JM, Yu MC. A vegetable-fruit-soy dietary pattern protects against breast cancer among postmenopausal Singapore Chinese women. *Am J Clin Nutr* 2010;91:1013-9. doi:10.3945/ajcn.2009.28572.

- 13 Boggs DA, Palmer JR, Wise LA, et al. Fruit and vegetable intake in relation to risk of breast cancer in the Black Women's Health Study. *Am J Epidemiol* 2010;172:1268-79. doi:10.1093/aje/kwq293.
- 14 Löf M, Sandin S, Lajou P, Trichopoulos D, Adami HO, Weiderpass E. Fruit and vegetable intake and risk of cancer in the Swedish women's lifestyle and health cohort. *Cancer Causes Control* 2011;22:283-9. doi:10.1007/s10552-010-9696-1.
- 15 Suzuki R, Iwasaki M, Hara A, et al. Japan Public Health Center-based Prospective Study Group. Fruit and vegetable intake and breast cancer risk defined by estrogen and progesterone receptor status: the Japan Public Health Center-based Prospective Study. *Cancer Causes Control* 2013;24:2117-28. doi:10.1007/s10552-013-0289-7.
- 16 Bradbury KE, Appleby PN, Key TJ. Fruit, vegetable, and fiber intake in relation to cancer risk: findings from the European Prospective Investigation into Cancer and Nutrition (EPIC). *Am J Clin Nutr* 2014;100(Suppl 1):394S-8S. doi:10.3945/ajcn.113.071357.
- 17 Jung S, Spiegelman D, Baglietto L, et al. Fruit and vegetable intake and risk of breast cancer by hormone receptor status. *J Natl Cancer Inst* 2013;105:219-36. doi:10.1093/jnci/djs635.
- 18 Land CE, Tokunaga M, Koyama K, et al. Incidence of female breast cancer among atomic bomb survivors, Hiroshima and Nagasaki, 1950-1990. *Radiat Res* 2003;160:707-17. doi:10.1667/RR3082.
- 19 Swerdlow AJ, Barber JA, Hudson GV, et al. Risk of second malignancy after Hodgkin's disease in a collaborative British cohort: the relation to age at treatment. *J Clin Oncol* 2000;18:498-509.
- 20 Wahner-Roedler DL, Nelson DF, Croghan IT, et al. Risk of breast cancer and breast cancer characteristics in women treated with supradiaphragmatic radiation for Hodgkin lymphoma: Mayo Clinic experience. *Mayo Clin Proc* 2003;78:708-15. doi:10.4065/78.6.708.
- 21 Willett WC, Lenart E. Reproducibility and validity of food frequency questionnaires. In: *Nutritional Epidemiology*. Oxford University Press, 2013: 96-141.
- 22 Muraki I, Imamura F, Manson JE, et al. Fruit consumption and risk of type 2 diabetes: results from three prospective longitudinal cohort studies. *BMJ* 2013;347:f5001. doi:10.1136/bmj.f5001.
- 23 Joshupura KJ, Hu FB, Manson JE, et al. The effect of fruit and vegetable intake on risk for coronary heart disease. *Ann Intern Med* 2001;134:1106-14. doi:10.7326/0003-4819-134-12-200106190-00010.
- 24 Steinmetz KA, Potter JD, Folsom AR. Vegetables, fruit, and lung cancer in the Iowa Women's Health Study. *Cancer Res* 1993;53:536-43.
- 25 Bhupathiraju SN, Wedick NM, Pan A, et al. Quantity and variety in fruit and vegetable intake and risk of coronary heart disease. *Am J Clin Nutr* 2013;98:1514-23. doi:10.3945/ajcn.113.066381.
- 26 Nutrient Database for Standard Reference, Release 14: Department of Agriculture ARS, 2001.
- 27 Holland GWA, Unwin ID, Buss DH, Paul AA, Dat S. *The Composition of Foods*. The Royal Society of Chemistry and Ministry of Agriculture, Fisheries and Food, 1991.
- 28 Maruti SS, Feskanich D, Rockett HR, Colditz GA, Sampson LA, Willett WC. Validation of adolescent diet recalled by adults. *Epidemiology* 2006;17:226-9. doi:10.1097/01.ede.0000198181.86685.49.
- 29 Maruti SS, Feskanich D, Colditz GA, et al. Adult recall of adolescent diet: reproducibility and comparison with maternal reporting. *Am J Epidemiol* 2005;161:89-97. doi:10.1093/aje/kwi019.
- 30 Colditz GA, Stampfer MJ, Willett WC, et al. Reproducibility and validity of self-reported menopausal status in a prospective cohort study. *Am J Epidemiol* 1987;126:319-25. doi:10.1093/aje/126.2.319.
- 31 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.
- 32 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177-88. doi:10.1016/0197-2456(86)90046-2.
- 33 Cochran WG. The combination of estimates from different experiments. *Biometrics* 1954;10:101-29. doi:10.2307/3001666.
- 34 Hefti MM, Hu R, Knoblauch NW, et al. Estrogen receptor negative/progesterone receptor positive breast cancer is not a reproducible subtype. *Breast Cancer Res* 2013;15:R68. doi:10.1186/bcr3462.
- 35 Lunn M, McNeil D. Applying Cox regression to competing risks. *Biometrics* 1995;51:524-32. doi:10.2307/2532940.
- 36 Archer E, Pavea G, Lavie CJ. The inadmissibility of what we eat in America and NHANES dietary data in nutrition and obesity research and the scientific formulation of national dietary guidelines. *Mayo Clin Proc* 2015;90:911-26. doi:10.1016/j.mayocp.2015.04.009.
- 37 Subar AF, Freedman LS, Toozé JA, et al. Addressing current criticism regarding the value of self-report dietary data. *J Nutr* 2015;145:2639-45. doi:10.3945/jn.115.219634.
- 38 Hendrickson SJ, Willett WC, Rosner BA, Eliassen AH. Food predictors of plasma carotenoids. *Nutrients* 2013;5:4051-66. doi:10.3390/nu5104051.
- 39 Chlebowski RT, Anderson GL, Aragaki AK, Prentice R. Breast Cancer and Menopausal Hormone Therapy by Race/Ethnicity and Body Mass Index. *J Natl Cancer Inst* 2015;108:djv327. doi:10.1093/jnci/djv327.
- 40 Gram IT, Park SY, Kolonel LN, et al. Smoking and risk of breast cancer in a racially/ethnically diverse population of mainly women who do not drink alcohol: The MEC Study. *Am J Epidemiol* 2015;182:917-25. doi:10.1093/aje/kwv092.
- 41 Aune D, Chan DS, Greenwood DC, et al. Dietary fiber and breast cancer risk: a systematic review and meta-analysis of prospective studies. *Ann Oncol* 2012;23:1394-402. doi:10.1093/annonc/mdr589.
- 42 Fung TT, Hu FB, McCullough ML, Newby PK, Willett WC, Holmes MD. Diet quality is associated with the risk of estrogen receptor-negative breast cancer in postmenopausal women. *J Nutr* 2006;136:466-72.
- 43 Hu F, Wu Z, Li G, et al. The plasma level of retinol, vitamins A, C and  $\alpha$ -tocopherol could reduce breast cancer risk? A meta-analysis and meta-regression. *J Cancer Res Clin Oncol* 2015;141:601-14. doi:10.1007/s00432-014-1852-7.
- 44 Nagel G, Linseisen J, van Gils CH, et al. Dietary beta-carotene, vitamin C and E intake and breast cancer risk in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Breast Cancer Res Treat* 2010;119:753-65. doi:10.1007/s10549-009-0444-8.
- 45 Maillard V, Kuriki K, Lefebvre B, et al. Serum carotenoid, tocopherol and retinol concentrations and breast cancer risk in the E3N-EPIC study. *Int J Cancer* 2010;127:1188-96. doi:10.1002/ijc.25138.
- 46 Zhang X, Spiegelman D, Baglietto L, et al. Carotenoid intakes and risk of breast cancer defined by estrogen receptor and progesterone receptor status: a pooled analysis of 18 prospective cohort studies. *Am J Clin Nutr* 2012;95:713-25. doi:10.3945/ajcn.111.014415.
- 47 Blanchard CM, Courneya KS, Stein K. American Cancer Society's SCS-II. Cancer survivors' adherence to lifestyle behavior recommendations and associations with health-related quality of life: results from the American Cancer Society's SCS-II. *J Clin Oncol* 2008;26:2198-204. doi:10.1200/JCO.2007.14.6217.

© BMJ Publishing Group Ltd 2016

## Appendix: Supplementary tables A-H