The American Journal of Clinical Nutrition 123 (2026) 101284

American Society for Nutrition
Excellence in Nutrition Research and Practice

The American Journal of CLINICAL NUTRITION

journal homepage: https://ajcn.nutrition.org/

Original Research Article

Consumption of organic compared with conventional fruits and vegetables in relation to cancer risk: findings from the NutriNet Santé cohort study

Justine Berlivet [1,2,*], Emma Meyer [1,2], Benjamin Alles [1,2], Denis Lairon [3], Serge Hercberg [1,2], Mathilde Touvier [1,2], Bernard Srour [1,2], Julia Baudry [1,2], Emmanuelle Kesse-Guyot [1,2]

[1] *Université Sorbonne Paris Nord and Université Paris Cité, Inserm, INRAE, CNAM, Centre for Research in Epidemiology and Statistics, Nutritional Epidemiology Research Team (EREN), Bobigny, France;* [2] *Nutrition and Cancer Research Network, Jouy-en-Josas, France;* [3] *Aix Marseille Université, Inserm, INRAE, C2VN, Marseille, France*

A B S T R A C T

Background: Regular fruits and vegetables (F&V) consumption is linked to a lower risk of certain cancers. However, F&V can contain pesticides, some of which may have carcinogenic properties. Organic foods are thought to reduce pesticide residue exposure, but the relationship between organic consumption and cancer risk is inconsistent. Notably, to our knowledge, no study has explored substituting conventional with organic F&V.

Objectives: This study examined the association between substituting conventional with organic F&V and cancer risk in the NutriNet-Santé cohort (~31,000 French adults), using detailed data from a food frequency questionnaire distinguishing between conventional and organic products.

Methods: Substitution of conventional F&V with organic F&V, at fixed total F&V consumption, was examined with multivariable Cox proportional hazards models per 100 g/d increment of substitution and by quintiles. Sensitivity analyses as well as marginal structural models were also implemented to improve causal inference.

Results: A total of 31,179 participants, 75% of whom were females, were included in the analyses. During follow-up {mean = 7.3 y [standard deviation = 3.1]}, 1718 cancer cases (284 cases of postmenopausal breast cancer) were registered >227,660 person-years. Substituting conventional with organic F&V was associated with a lower risk of overall cancer {hazard ratio (HR) [95% confidence interval (CI)] for a 100 g/d substitution = 0.98 (0.95, 1.00)} and a lower risk of postmenopausal breast cancer [HR (95% CI) for a 100 g/d substitution = 0.90 (0.85, 0.96)], whereas no association was detected for other cancer locations. When modeled by quintiles, an inverse association was found for postmenopausal breast cancer but not for overall cancer or other locations. When running marginal structural models and testing for additional adjustments, findings remained significant for postmenopausal breast cancer.

Conclusions: In the present work, substituting conventional F&V with organic ones was associated with a reduced risk of postmenopausal cancer. The specific role of organic F&V compared with conventional F&V needs further investigation in other contexts.

The NutriNet-Santé cohort is registered at clinicaltrials.gov as NCT03335644 (https://clinicaltrials.gov/study/NCT03335644).

Keywords: fruit and vegetables, organic food, cancer, pesticides, prospective study

Introduction

Cancers are multifactorial diseases, caused by genetic and environmental factors, as well as lifestyle behaviors, and are responsible for 10 million deaths worldwide [1,2]. A balanced diet, including a regular consumption of fruits and vegetables (F&V), has been shown to be protective against several types of cancers [3]. According to the World Cancer Research Fund, F&V consumption is associated with lower risk of upper aerodigestive tract cancer ("probable decreases risk") and lower risk of breast cancer (vegetables only, "limited—suggestive decreases risk") [3]. However, F&V are the main contributors of pesticide residue exposure through diet in the general population [4].

Certain pesticides have been classified as probable carcinogenic to humans by the International Agency for Research on Cancer [5]. Adverse associations related to pesticides imply many biological mechanisms [6–8]. Although the health implications of occupational pesticide exposure are well documented [9], dietary pesticide exposure

Abbreviations: CI, confidence interval; CU, consumption units; FFQ, food frequency questionnaire; F&V, fruit and vegetables; HR, hazard ratio; NHL, non-Hodgkin's lymphoma; PNNS-GS2, Programme National Nutrition Santé-Guidelines Score 2; Q, quintile.

* Corresponding author.
E-mail address: justine.berlivet@eren.smbh.univ-paris13.fr (J. Berlivet).

https://doi.org/10.1016/j.ajcnut.2026.101284
Received 14 October 2025; Received in revised form 6 March 2026; Accepted 16 March 2026; Available online 19 March 2026
0002-9165/

remains understudied. Organic food consumption has been related to lower pesticide exposure [9,10], bringing the hypothesis that organic food consumption could be beneficial [9,11–14]. Nonetheless, to our knowledge, only 3 studies have investigated the relationships between organic food consumption and cancer risk, yielding conflicting results. The British Million Women Study found a lower risk of non--Hodgkin's lymphoma (NHL), and a slightly higher risk of breast cancer among regular consumers of organic food products [15]. In a Danish study, a lower risk of stomach cancer and a slightly increased risk of NHL were observed among regular organic food consumers [16]. In the NutriNet-Santé cohort, we observed an inverse association between organic food consumption and cancer incidence, especially lymphoma and postmenopausal breast cancer [17]. In all aforementioned studies, organic food consumption was assessed through relatively rudimentary questions about the frequency of organic consumption.

In the past decade, some studies [18–21] carried out in the United States reported that exposure to F&V pesticide residues could dampen the protective role of F&V on risk of glioma (a brain tumor subtype) [20] and overall cancer-related mortality [19]. However, intake of highly pesticide-contaminated F&V was not related to cancer risk [18]. In a previous study, we found that postmenopausal breast cancer risk was associated with higher exposure to a specific pesticide mixture (highly correlated to chlorpyrifos, imazalil, malathion, and thiabendazole) among females with overweight or obesity [22].

In that context, we aimed to investigate the differential association between consumption of conventionally or organically produced F&V and cancer risk. Especially, we tested whether the substitution of conventional F&V with organic ones could be associated with a decrease in cancer risk in a large sample of French adults, using detailed data to evaluate F&V intake.

Methods

Study design and population

This study was conducted on a subpopulation of adults from the web-based prospective NutriNet-Santé cohort, which aims to study the complex relationships of dietary habits with health and diseases [23]. The participants are volunteers recruited from the general French population. The recruitment was launched in 2009 and is still ongoing.

Sociodemographic data (sex, age, monthly income, occupational category, and educational attainment), lifestyle characteristics (smoking status and physical activity level [24]), and anthropometric characteristics [25] are collected at inclusion and each year thereafter, using validated questionnaires [26,27].

This study was conducted in accordance with the Declaration of Helsinki, and all procedures were approved by the Institutional Review Board of the French Institute for Health and Medical Research (Inserm 0000388FWA00005831) and the National Commission on Informatics and Liberty (Commission Nationale de l'Informatique et des Libertés 908450 and 909216). Electronic informed consent was obtained from all participants. The NutriNet-Santé study is registered in clinicaltrials.gov (NCT03335644).

Dietary data

Dietary data were collected between June and December 2014 through a self-administered semiquantitative food frequency questionnaire (FFQ), aiming to assess organic (i.e., complying with official European Union standards and carrying the European label [28]) and conventional food consumption [29]. This dietary tool is based on a previously validated FFQ [30], upgraded by a 5-point scale on the consumption frequency as organic for each of the FFQ-item ($n = 264$) [30]. For each food item, participants reported the frequency of food consumed as organic by ticking the following modalities: "never," "rarely," "half-of-time," "often," or "always" in response to the question "How often was the product of organic origin?" Weights were allocated to each frequency, that is, 0%, 25%, 50%, 75% and 100%, respectively. The organic consumption for a single questionnaire item is determined by multiplying the quantity (portion   frequency) by the proportion consumed as organic (ranging from 0–1). Nutrient intakes were calculated using a published food composition database [31].

Underreporters and overreporters were defined as participants with a ratio between energy intake and energy requirement below or above cut-offs previously identified (0.35 and 1.93) corresponding to the 1st and 99th percentiles of the ratio distribution [29].

Overall dietary quality was assessed using the simplified Programme National Nutrition Santé-Guidelines Score 2 (sPNNS-GS2), a validated dietary quality score reflecting the adherence to the 2017 French dietary guidelines [32]. The score used in the present study is the simplified score, meaning that only the main benchmarks of the French dietary guidelines were considered [33]. Components, scoring, and weighting are detailed in Supplemental Method 1**.** Percentage of ultraprocessed food in the diet was estimated using the NOVA classification applied to the 24-h dietary records concomitant to the FFQ assessment. Furthermore, details are available in Supplemental Method 2**.** Using data from a previous work [34], we also evaluated the associations between estimated organic food or F&V consumption and the urinary concentrations of dialkylphosphates metabolites, to test whether organic food consumption was associated with a lower pesticide exposure among a subsample of individuals with available urinary biomarker measurements ($n = 289$) (Supplemental Method 3).

Ascertainment of cancer cases

Cancer ascertainment was based on a multisource approach. Participants self-report health events through annual health questionnaires, specific health check-up questionnaires administered every 6 mo, and at any time spontaneously via the secure online platform. Participants could also declare all new prescribed medications. Moreover, the NutriNet-Santé cohort is linked to the national health insurance system database to collect additional information regarding medication and consultations, and to the French national mortality registry to identify the occurrence and cause of deaths.

Medical records were obtained by study physicians to validate the cases. Cancer cases were classified using the International Classification of Diseases, 10th revision [35]. All first incident cancers were considered, except for basal cell carcinomas of the skin, which were not considered as cancer cases.

All primary cancers diagnosed between the date of completion of the FFQ (2014) and 20 November 2024 were considered as cases. Analyses were carried out for all cancers regardless of location, and then locations for which the number of cases exceeded 80 were explored separately: breast (premenopausal and postmenopausal), prostate, skin, and colon-rectum cancers. Of note, due to a few number of cases, lymphoma and upper aerodigestive tract cancers were not examined separately.

Covariates

Data on age, sex (male/female), highest graduation degree (high-school degree, high-school degree, and postsecondary degree), occupation (unemployed, farmer/merchant/craftworker/company director, employee/manual worker, intermediate profession, managerial staff, retired, and never employed), monthly income per household unit (see below), marital status (alone and cohabiting), smoking status (never, former, and current) as well as tobacco consumption, as packs-year were collected at enrolment and each year hereafter. Physical activity level (high, moderate, and low) was assessed at baseline and each year using the International Physical Activity Questionnaire [36].

In the present study, monthly income was calculated by household unit, that is, by dividing the household's total monthly income by the number of consumption units (CUs), using the National Institute of Statistics and Economic Studies guidelines. The first adult in the household was thus allocated 1 household unit, whereas other individuals aged 14 y or older were allocated 0.5 units, and children aged 14 y were allocated 0.3 units [37]. The following categories of monthly income per CU were used: 1200; 1200–1800; 1800–3700, and >3700 euros.

Validated anthropometric questionnaires provided information on height and weight [25]. Family history (including parents and siblings) of cancer was collected. For female participants, the number of biological children, age at menarche and at menopause, and any medication including oral contraception and hormonal treatment for menopause were also collected. All covariate data used were those collected closest to the FFQ.

Statistical analysis

Sample selection

Participants were individuals who completed the detailed FFQ administered in 2014. Individuals with a prevalent cancer and those without available follow-up data were subsequently excluded, as shown in the Supplemental Figure 1. The final analytical sample comprised 31,179 individuals.

Descriptive statistics

Participants were ranked and categorized into sex-specific quintiles of organic F&V consumption (modeled as residuals).

The sample's characteristics according to sex-specific quintiles of F&V consumed as organic were presented as mean (SD) or %. Tests for differences were computed using Mantel–Haenszel χ^2 test for dichotomous or ordinal variables, or linear contrasts from ANOVA for numeric variables.

The relationships among consumption, nutrient intake, and urinary pesticide concentrations are detailed in Supplemental Method 4.

Cancer risk modeling

The associations between exposures and the risk of cancer (overall and by location, including breast cancer in female participants and prostate cancer in male participants, as well as skin and colorectal cancer) were estimated using multivariable Cox proportional hazards models with age as the time scale, providing hazard ratios (HRs) and 95% confidence intervals (95% CIs). Each participant contributed person-time from baseline (completion of the FFQ) until the date of diagnosis, date of death, date of last completed questionnaire, or 20 November 2024, whichever occurred first. As recommended in the presence of competing events, we used cause-specific models [38]. For analyses by location, other cancer cases than the studied cancer were censored at the date of diagnosis. Missing data on covariates were dealt with multiple imputations using the SAS MI procedure [39], as described in Supplemental Method 5.

The Cox proportional hazards and linearity assumptions were formally tested and are presented graphically (Supplemental Figures 2 and 3) [40,41]. Regarding the risk of premenopausal breast cancer, the date of menopause, the date of diagnosis, date of death, date of last completed questionnaire, or 20 November 2024 was considered for censoring, whichever occurred first. For postmenopausal breast cancer, follow-up began at the date of menopause if the participant was already part of the cohort before reaching menopause.

Consumption modeling

Preliminary analyses were conducted to assess the relationship between conventional and organic F&V consumption and cancer risk (overall and by location). Models were adjusted for age (time scale), sex (when appropriate), educational attainment, occupation, monthly household income, marital status, energy intake without alcohol (kcal/d, continuous variable), alcohol consumption (g/d, continuous variable), physical activity, smoking status, number of cigarettes smoked (packs/y, continuous variable), BMI (kg/m^2, continuous variable), height (m, continuous variable), family history of cancer (yes/no), and intakes of dairy products, red meat, processed meat, added sugar, sodium, whole grain products and total lipids. For postmenopausal breast cancer analysis, the model was further adjusted for the number of biological children (continuous variable), age at first period (y, continuous variable), menopausal status (premenopausal/postmenopausal), and hormonal treatment for menopause use (yes/no).

The models examining the relationship between organic F&V consumption and cancer risk accounted for the intake of conventional F&V, and vice versa.

Substitution modeling

To examine the role of substituting conventional with organic F&V, substitution modeling was performed using the residual method [42]. This method enabled us to separate the contribution of organic F&V consumption from the total F&V intake by accounting for the latter.

Using this approach, the explanatory variable becomes the residual derived from a linear regression modeling the intake of organic F&V as a function of the total intake of F&V.

In a first step, the residual ε of the linear regression model (conducted by sex) of organic F&V consumption on the total F&V consumption (organic F&V + conventional F&V) was calculated as follows:

$$\text{Organic F\&V} = a0 + a1 \quad \text{total F\&V}) + \varepsilon \qquad (1)$$

Then, to facilitate the interpretation, a constant equal to the predicted value of organic F&V for the sample average of total F&V consumption $(\overline{total\ F\&V})$ was added to the residual:

$$\varepsilon^{*} = \varepsilon + [a0 + a1 \quad \overline{(total \quad \&V)}] \qquad (2)$$

By definition, ε*, the residual of the regression of organic F&V on total F&V (organic and conventional), was not correlated to the total consumption of F&V, and $(\overline{total\ F\&V})$was the mean consumption of total (organic and conventional) F&V in the study sample.

Both variables (total F&V and ε*) were then introduced into the final model aiming to assess the association between organic F&V consumption and cancer risk. The estimate for organic F&V therefore represents the substituting effect of organic with conventional F&V.

The association between the substitution of conventional F&V with organic ones and cancer risk (overall and by location) was computed. Substitution of conventional with organic F&V consumption was modeled as sex-specific quintiles (with the first as the reference) and as continuous variable using increments of 100 g/d (residual on the total F&V), that is, substitution of 100 g/d of conventionally produced by organically produced F&V. Model 1 was adjusted for the following baseline covariates: age (time scale), sex (except for breast cancer and prostate cancer), educational attainment, occupation, monthly household income, marital status, energy intake without alcohol (kcal/d, continuous variable), alcohol consumption (g/d, continuous variable), physical activity level, smoking status, number of cigarettes smoked (pack/y, continuous variable), BMI (kg/m^2, continuous variable), height (m, continuous variable), family history of cancer (yes/no), and level of adherence to food-based dietary guidelines (sPNNS-GS2, continuous variable). For postmenopausal breast cancer analysis, model 1 was further adjusted for the number of biological children (continuous variable), age at first period (y, continuous variable), menopausal status (premenopausal/postmenopausal), and hormonal treatment for menopause use (yes/no).

Model 2 was further adjusted for the proportion (in weight) of ultraprocessed food in the diet.

Model 3 was model 1 replacing sPNNS-GS2 with consumption of dairy products, red meat, processed meat, and intakes of sugar, sodium, whole grain products, and total lipids [all modeled as continuous variables (g/d)].

Sensitivity analyses

Moreover, we computed a standardized regression method using a counterfactual approach to derive survival curves, also known as marginal survival curves [43,44]. These standardized survival estimates provide a causal estimand that can be interpreted as the average counterfactual survival probabilities assuming each participant (with his/her own characteristics) in the study population was either exposed or unexposed. As a result, the distribution of all covariates will be similar in the 2 standardized probabilities, enabling comparisons to be made between the exposed and unexposed groups independently of confounding factors, thereby enhancing causal inference.

We also conducted several sensitivity analyses using various modeling approaches: *1*) removing early cases (from the first year) to reduce reverse causality (model 1); *2*) modifying the value assigned to the organic frequency “rarely" from 25% to 10% (model 1); *3*) using weight for 2009 census data including participants from mainland France only (model 1); *4*) using marginal structural models to enhance causal inference (Supplemental Method 6); *5*) using weight to compensate the selection bias using inverse probability weighting (model 1) (Supplemental Method 7); *6*) replacing sPNNS-GS2 by consumptions of conventional dairy products, processed meat, red meat, whole grain products and organic dairy products, processed meat, red meat, and whole grain products. All tests were 2-sided and *P* 0.05 was considered statistically significant. Statistical analyses were conducted using R version 4.0.4 and SAS version 8.3 (SAS Institute).

Results

A total of 31,179 participants were included in the analyses (as described in the flowchart in Supplemental Figure 1). The baseline mean (SD) age was 53 y (14 y), and 75% of participants were female. Mean (SD) follow-up was 7.3 y (3.1). During the follow-up, 1718 cancer cases (of which 284 postmenopausal breast, 269 prostate, 235 skin, and 146 colon rectum) were ascertained over 227,660 person-years.

The sample characteristics according to sex-specific quintiles of organic F&V consumption (modeled as residuals) are presented in Table 1. Higher organic F&V consumption was associated with higher physical activity level. The highest percentage of managerial staff was found in quintile (Q)3 and the lowest in Q1, whereas no association was observed with monthly income.

By construction, consumption of organic F&V as residuals was not linearly correlated to total F&V consumption. Participants in Q5 showed the highest sPNNS-GS2 and the lowest BMI. They also had the highest proportion of organic food consumption in their overall diet, with a mean of 68%.

Table 2 displays the associations between conventional or organic F&V consumption and cancer risk (overall and by cancer site) after multivariable adjustment including for corresponding conventional or organic F&V consumption. Regarding conventional F&V consumption, no association was found for overall cancer risk and for any site when the exposure was modeled continuously or by quintile, although for the latter, a trend was observed for overall cancer risk.

For organic F&V consumption, we observed a reduced risk of overall cancer and postmenopausal breast cancer, with a significant reduction of 3% and 10% for each 100 g/d increase in organic F&V consumption. Considering quintiles, a trend was detected only for postmenopausal breast cancer.

Substitution Cox models are presented in Table 3. Considered as quintiles, the substitution of conventional F&V with organic F&V was not associated with the risk of overall cancer [HR Q5 compared with Q1 = 0.89 (95% CI: 0.76, 1.03), *P* for trend = 0.48, model 1—main model]. A reduction of 2% (*P* = 0.05) was found when the exposure was modeled as a continuous variable, representing the replacement of 100 g/d of conventional by organic F&V. Moreover, when considering quintiles of exposure, the higher organic F&V substitution, the lower the risk of postmenopausal breast cancer [model 1, HRQ5 compared with Q1 = 0.59 (95% CI: 0.41, 0.85), *P*-trend = 0.01], with a risk reduction of 10% per 100 g/d increment of substitution. No association was detected for other cancer locations, regardless of the type of modeling (continuous or quintile modeling).

Furthermore, adjustment for the share of ultraprocessed food in the diet (model 2) or replacement of diet quality assessed by the sPNNS-GS2, by food groups and nutrient intakes (model 3) led to similar trends.

In the sensitivity analyses excluding early cases (Supplemental Table 1), no association was detected for overall cancer but the association for postmenopausal breast cancer remained.

Using the value of the “rarely” modality (allocating 10% instead of 25%) (Supplemental Table 2) yielded similar results. Additional sensitivity analyses are provided in Supplemental Table 3. In the model adjusted for census data, results showed a stronger association, with a –7% risk of overall cancer and –20% risk of postmenopausal breast cancer per 100 g/d increment (*P* 0.0001). The findings resulting from the marginal structural models were largely consistent. Then, the model that accounted for selection bias showed a significant association only for postmenopausal breast cancer. Finally, the model including organic and conventional consumption of food groups was similar to the main analyses.

The survival marginal curves for overall cancer, breast, colorectal, and skin cancer are presented in Figure 1. In line with the main model, the substitution of conventional F&V by organic ones analyzed as quintiles was associated with a reduced risk of both overall cancer and postmenopausal breast cancer.

TABLE 1
Characteristics of the participants according to sex-specific quintiles of organic fruit and vegetable consumption as residuals on total fruit and vegetables consumption, NutriNet-Santé cohort study, France, 2014 ($n = 31{,}179$)[1].

N	All	Q1	Q2	Q3	Q4	Q5
	31,179	6235	6236	6236	6236	6236
Cut-off of the variable organic fruit and vegetables (g/d)						
Female		71	71≤163	163≤232	232≤342	≥342
Male		45	45≤133	133≤200	200≤308	≥308
Sex (female) (%)	75.69	75.70	75.69	75.69	75.69	75.69
Age (y)	52.65 (14.02)	56.50 (13.12)	52.47 (14.27)	49.52 (14.54)	50.85 (13.92)	53.93 (13.13)
Education attainment (%)						
High-school degree	20.34	26.13	20.57	17.27	17.80	19.95
High-school degree	14.55	15.97	14.90	13.53	14.27	14.06
> High-school degree	65.11	57.90	64.53	69.19	67.93	65.99
Occupation (%)						
Unemployed	4.30	3.66	4.12	4.28	4.62	4.83
Farmer, merchant, craftworker, company director	1.94	1.30	1.67	1.78	2.08	2.85
Employee/manual worker	15.11	14.82	16.02	16.79	14.72	13.18
Intermediate profession	15.64	13.63	16.40	16.85	17.59	13.73
Managerial staff	22.16	15.51	21.66	26.17	24.98	22.47
Retired	33.76	43.77	33.39	26.64	29.49	35.52
Never employed	7.09	7.31	6.74	7.49	6.51	7.42
Monthly household income per CU (%)						
Unwilling to answer	11.02	11.47	11.10	10.87	10.23	11.45
1200 €	7.17	7.60	6.83	7.62	6.24	7.55
1200–>1800 €	23.31	23.18	22.90	23.59	23.75	23.14
1800–>3700 €	27.13	26.54	26.68	26.86	28.56	27.02
≥3700 €	31.36	31.21	32.49	31.06	31.22	30.84
Marital status (%)						
Living alone	16.28	15.96	16.82	17.61	16.08	14.93
Cohabiting	83.72	84.04	83.18	82.39	83.92	85.07
Physical activity (%)						
Low	21.80	19.58	23.88	27.35	22.77	15.51
Moderate	41.30	39.45	42.17	42.40	43.14	39.33
High	36.91	40.97	33.95	30.25	34.09	45.16
BMI (%)	24.11 (4.61)	24.94 (4.94)	24.32 (4.67)	24.08 (4.60)	23.89 (4.47)	23.33 (4.22)
Smoking status (%)						
Never smoker	49.32	47.81	49.29	49.94	50.22	49.31
Former smoker	39.45	43.10	39.38	36.93	36.53	41.32
Current smoker	11.23	9.09	11.32	13.13	13.25	9.36
Family history of cancer (yes) (%)	41.22	55.00	59.12	62.33	61.34	56.13
Number of cigarette pack-years	5.51 (11.66)	6.69 (13.23)	5.45 (11.52)	5.06 (11.17)	5.08 (11.17)	5.25 (10.97)
Energy intake (kcal/d)	1996.0 (629.9)	2222.9 (660.3)	1968.6 (591.9)	1854.4 (591.5)	1853.2 (592.1)	2081.0 (630.1)
Fruit and vegetables consumption (g/d)	555.58 (371.70)	811.92 (396.77)	467.93 (219.92)	355.16 (219.19)	403.44 (256.26)	739.50 (448.17)
Organic fruit and vegetables consumption (g/d)	203.16 (282.14)	52.89 (105.56)	72.89 (101.25)	95.62 (103.41)	200.68 (130.60)	593.69 (381.11)
sPNNS-GS2	2.74 (3.41)	2.78 (3.52)	2.35 (3.34)	2.06 (3.35)	2.57 (3.27)	3.95 (3.26)
Ultraprocessed food (%)	0.16 (0.08)	0.15 (0.07)	0.16 (0.08)	0.17 (0.09)	0.16 (0.08)	0.15 (0.07)
Organic food ratio	0.30 (0.27)	0.09 (0.11)	0.13 (0.14)	0.20 (0.16)	0.38 (0.19)	0.68 (0.20)
Alcohol consumption (g/d)	8.29 (12.34)	8.46 (13.78)	8.34 (11.86)	8.35 (12.61)	8.37 (12.18)	7.91 (11.10)

Abbreviations: Q, quintile; sPNNS-GS2, *simplified* Programme National Nutrition Santé-Guidelines Score 2.
[1] Values are means (SD) or %.As residuals as explained in Supplemental Method 1.

Discussion

In the present study conducted in a large sample of adult participants from the NutriNet-Santé cohort, substitution of conventional with organic F&V (for 100 g) was associated with a reduction of 2% of overall cancer risk, after a mean follow-up of 7.3 y, accounting for many confounding factors. This association was not observed when the exposure was modeled in quintiles. Noticeably, regardless of the way the exposure was modeled, we observed a reduced risk of postmenopausal breast cancer related to a higher substitution of conventional F&V with organic options (e.g., –10% for a 100 g/d increment substitution). Moreover, regardless of the sensitivity analyses performed, our findings exhibited robustness for postmenopausal breast cancer.

There is insufficient scientific literature about the long-term implications of organic food consumption on cancer risk [14,45,46]. In a previous work [17], we observed an inverse relationship between overall organic food consumption frequency and cancer risk. In contrast with the current study, organic food consumption frequency measurement was based on a few questions about 16 large food groups (e.g., fruits, vegetables, grains and legumes, etc.), so that it was not possible to consider food items within food groups, as no quantitative data were available. Two other studies [15,16] did not detect any associations between organic food consumption and overall cancer risk, but inverse associations were detected for some cancer locations, including NHL [15], and stomach cancer [16]. In contrast, a slight positive association was observed with breast cancer [15] and with lymphoma [16].

TABLE 2

Consumption of organic and conventional F&V in association with the risk of overall cancer and by locations, NutriNet-Santé cohort study, France, 2014–2024 ($n = 31{,}179$)[1].

Conventional F&V	Per 100 g/d increment	*P* value	Sex-specific quintiles[2]					*P*-trend[3]
			Q1	Q2	Q3	Q4	Q5	
All cancers								
n events	1718		289	341	337	351	400	
Person-years	227,660		44,908	45,793	45,879	46,071	45,009	
Model[4]	1.00 (0.97, 1.02)	0.70	1 (ref)	1.00 (0.85, 1.18)	0.90 (0.76, 1.06)	0.85 (0.71, 1.00)	0.88 (0.73, 1.06)	0.04
Postmenopausal breast cancer								
n events	284		44	52	53	57	78	
Person-years	90,645		14,730	15,853	17,512	20,385	22,165	
Model[4,5]	1.00 (0.94, 1.05)	0.92	1 (ref)	0.87 (0.58, 1.32)	0.76 (0.50, 1.16)	0.68 (0.44, 1.05)	0.80 (0.50, 1.26)	0.21
Prostate cancer								
n events	269		43	59	52	59	56	
Person-years	56,272		17,793	17,291	15,877	13,638	10,897	
Model[4]	0.95 (0.89, 1.01)	0.08	1 (ref)	1.14 (0.75, 1.71)	0.88 (0.57, 1.36)	0.88 (0.57, 1.36)	0.73 (0.45, 1.19)	0.09
Colorectal cancer								
n events	146		20	37	18	31	40	
Person-years	227,660		44,908	45,793	45,879	46,071	45,009	
Model[4]	1.03 (0.95, 1.11)	0.50	1 (ref)	1.61 (0.92, 2.82)	0.72 (0.37, 1.41)	1.13 (0.61, 2.09)	1.53 (0.79, 2.94)	0.62
Skin cancer								
n events	235		40	36	49	53	57	
Person-years	227,660		44,908	45,793	45,879	46,071	45,009	
Model[4]	1.01 (0.95, 1.07)	0.82	1 (ref)	0.76 (0.48, 1.21)	0.94 (0.60, 1.47)	0.90 (0.57, 1.42)	0.90 (0.55, 1.49)	0.98
Organic F&V			Q1	Q2	Q3	Q4	Q5	
All cancers								
n events	1718		357	318	354	361	328	
Person-years	227,660		46,832	43,298	46,726	45,922	44,881	
Model[4]	0.97 (0.95, 1.00)	0.04	1 (ref)	1.14 (0.98, 1.33)	1.15 (0.99, 1.34)	1.10 (0.94, 1.29)	1.03 (0.86, 1.23)	0.70
Postmenopausal breast cancer								
n events	284		66	54	58	63	43	
Person-years	90,645		16,002	15,446	17,304	19,955	21,938	
Model[4,5]	0.90 (0.84, 0.96)	0.002	1 (ref)	0.83 (0.57, 1.19)	0.80 (0.56, 1.16)	0.76 (0.52, 1.10)	0.49 (0.31, 0.77)	0.01
Prostate cancer								
n events	269		56	43	56	53	61	
Person-years	56,272		13,405	9056	11,445	11,235	11,132	
Model[4]	0.98 (0.93, 1.05)	0.61	1 (ref)	1.31 (0.88, 1.96)	1.40 (0.96, 2.05)	1.25 (0.84, 1.86)	1.41 (0.92, 2.17)	0.15
Colorectal cancer								
n events	146		29	29	28	27	33	
Person-years	227,660		46,832	43,298	46,726	45,922	44,881	
Model[4]	1.02 (0.94, 1.11)	0.66	1 (ref)	1.44 (0.85, 2.44)	1.25 (0.73, 2.14)	1.18 (0.67, 2.05)	1.72 (0.95, 3.10)	0.21
Skin cancer								
n events	235		39	50	49	55	42	
Person-years	227,660		46,832	43,298	46,726	45,922	44,881	
Model[4]	0.97 (0.91, 1.04)	0.44	1 (ref)	1.54 (1.01, 2.36)	1.36 (0.88, 2.10)	1.41 (0.92, 2.18)	1.06 (0.64, 1.75)	0.82

Abbreviations: CI, confidence interval; F&V, fruit and vegetables; HR, hazard ratio; Q, quintile.

[1] HR and 95% CI are derived from a multivariable Cox proportional hazard models.

[2] Sex-specific cut-offs for quintiles of conventional F&V (g/d) were 130, 130– 232, 232– 345, 345– 521, and ≥521 for females and 141, 141– 239, 239– 353, 353– 519, and ≥519 for males. Sex-specific cut-offs for quintiles of organic F&V (g/d) were 3, 3– 74, 74– 173, 173– 347, and ≥347 for females and 0, 0.05– 47, 47– 138, 138– 304, and ≥304 for males.

[3] *P* value of Wald's test for quintile as an ordinal variable.

[4] Models are a multivariable Cox proportional hazard model adjusted for age (time scale), sex (when appropriate), educational, occupation, monthly household income, marital status, energy intake without alcohol (continuous variable), alcohol consumption (continuous variable), physical activity, smoking status, number of cigarettes smoked (continuous variable), BMI (continuous variable), height (continuous variable), family history of cancer, and intakes of dairy products, red meat, processed meat, whole grain products, added sugar, sodium, and total lipids (continuous variable).

[5] In the postmenopausal breast cancer models, further adjustments were made for age at menarche (12 y, ≥12 y), menopausal treatment use (yes/no), and number of children (continuous).

These studies show varying results, especially for specific cancer sites. These inconsistencies may stem from discrepancies in study design, methodological factors, or possibly chance. In particular, Bradbury et al. [15] hypothesized that females reporting frequently consuming organic food may have healthier profiles, and more concerns about diet, lifestyle, and health. In particular, they could be more likely to attend breast cancer screening and to be diagnosed with breast cancer [15]. Although the elevated colorectal cancer risk observed among the highest consumers of F&V is unexpected and warrants careful consideration, this finding is likely attributable to statistical instability stemming from the limited number of cases in our cohort. This limitation reduces the robustness of quintile-based analyses when organic and conventional intakes are mutually adjusted, leading to wide CIs and unreliable coefficients at the extremes of the distribution.

TABLE 3

Substitution of conventional with organic fruit and vegetables in association with the risk of overall cancer and by locations, NutriNet-Santé cohort study, France, 2014–2024 (n = 31,179)[1].

	Per 100 g/d increment	P value	Sex-specific quintiles[2]					P-trend[3]
			Q1	Q2	Q3	Q4	Q5	
All cancers								
n events	1718		406	321	343	337	311	
Person-years	227,660		44,998	45,917	45,535	46,387	44,824	
Model 1[4]	0.98 (0.95, 1.00)	0.05	1 (ref)	0.85 (0.73, 0.99)	1.01 (0.86, 1.18)	0.96 (0.82, 1.13)	0.89 (0.76, 1.03)	0.48
Model 2[5]	0.98 (0.95, 1.00)	0.05	1 (ref)	0.85 (0.73, 1.00)	1.01 (0.86, 1.18)	0.96 (0.82, 1.13)	0.89 (0.76, 1.03)	0.47
Model 3[6]	0.98 (0.96, 1.00)	0.08	1 (ref)	0.86 (0.74, 1.01)	1.04 (0.88, 1.22)	0.99 (0.85, 1.16)	0.91 (0.78, 1.06)	0.74
Postmenopausal breast cancer								
n events	284		82	55	47	54	46	
Person-years	90,645		20,975	17,724	14,738	17,024	20,184	
Model 1[4,7]	0.90 (0.85, 0.96)	0.001	1 (ref)	0.69 (0.48, 0.99)	0.67 (0.45, 0.99)	0.70 (0.48, 1.01)	0.59 (0.41, 0.85)	0.01
Model 2[5,7]	0.90 (0.85, 0.96)	0.001	1 (ref)	0.69 (0.48, 0.99)	0.67 (0.45, 0.99)	0.70 (0.48, 1.01)	0.59 (0.41, 0.85)	0.01
Model 3[6,7]	0.91 (0.85, 0.96)	0.002	1 (ref)	0.69 (0.48, 0.99)	0.67 (0.45, 1.00)	0.70 (0.48, 1.02)	0.60 (0.41, 0.88)	0.02
Prostate cancer								
n events	269		55	54	56	46	58	
Person-years	56,272		11,179	11,423	11,233	11,288	11,149	
Model 1[4]	1.03 (0.97, 1.09)	0.38	1 (ref)	1.06 (0.71, 1.59)	1.28 (0.84, 1.93)	1.12 (0.73, 1.71)	1.36 (0.93, 1.98)	0.12
Model 2[5]	1.03 (0.97, 1.09)	0.38	1 (ref)	1.07 (0.71, 1.59)	1.29 (0.85, 1.95)	1.12 (0.73, 1.72)	1.35 (0.93, 1.98)	0.12
Model 3[6]	1.03 (0.97, 1.10)	0.30	1 (ref)	1.09 (0.73, 1.63)	1.31 (0.87, 1.99)	1.15 (0.75, 1.76)	1.43 (0.96, 2.11)	0.09
Colorectal cancer								
n events	146		36	27	25	29	29	
Person-years	227,660		44,998	45,917	45,535	46,387	44,824	
Model 1[4]	0.99 (0.92, 1.07)	0.81	1 (ref)	0.82 (0.48, 1.40)	0.86 (0.49, 1.52)	0.99 (0.58, 1.68)	1.02 (0.62, 1.68)	0.74
Model 2[5]	0.99 (0.92, 1.07)	0.81	1 (ref)	0.82 (0.48, 1.40)	0.86 (0.49, 1.51)	0.99 (0.58, 1.68)	1.02 (0.62, 1.68)	0.73
Model 3[6]	0.99 (0.92, 1.08)	0.87	1 (ref)	0.85 (0.50, 1.45)	0.95 (0.54, 1.67)	1.06 (0.62, 1.82)	1.07 (0.64, 1.79)	0.58
Skin cancer								
n events	235		62	37	49	43	44	
Person-years	227,660		44,998	45,917	45,535	46,387	44,824	
Model 1[4]	0.96 (0.90, 1.02)	0.21	1 (ref)	0.65 (0.42, 1.00)	0.97 (0.64, 1.47)	0.81 (0.53, 1.23)	0.74 (0.50, 1.10)	0.34
Model 2[5]	0.96 (0.90, 1.02)	0.21	1 (ref)	0.65 (0.42, 1.00)	0.97 (0.64, 1.48)	0.81 (0.53, 1.23)	0.74 (0.50, 1.10)	0.34
Model 3[6]	0.98 (0.92, 1.04)	0.48	1 (ref)	0.66 (0.43, 1.01)	1.00 (0.66, 1.51)	0.85 (0.56, 1.30)	0.82 (0.54, 1.23)	0.67

Abbreviations: CI, confidence interval; HR, hazard ratio; Q, quintile; sPNNS-GS2, *simplified* Programme National Nutrition Santé-Guidelines Score 2.

1 HR and 95% CI are derived from a multivariable Cox proportional hazard.

2 Sex-specific cut-offs for quintiles of organic fruit and vegetables (g/d) (as residuals) were 71, 71– 163,163– 232, 232– 342, and ≥342 for females and 45, 45– 133, 133– 200, 200– 308, and ≥308 for males.

3 P value of Wald's test for quintile as an ordinal variable.

4 Model 1 is a multivariable Cox proportional hazard model adjusted for age (time scale), sex (when appropriate), educational, occupation, monthly household income, marital status, energy intake without alcohol (continuous variable), alcohol consumption (continuous variable), physical activity, smoking status, number of cigarettes smoked (continuous variable), BMI (continuous variable), height (continuous variable), family history of cancer, sPNNS-GS2 (continuous variable) and total fruit and vegetables consumption.

5 Model 2 is model 1 further adjusted for ultraprocessed food consumption.

6 Model 3 is model 1 with the replacement of sPNNS-GS2 with consumption of dairy products, red meat, processed meat, whole grain products, sugar, sodium, and total lipids [all modeled as continuous variables (g/d)].

7 In the postmenopausal breast cancer models, further adjustments were made for age at menarche (12 y, ≥12 y), menopausal treatment use (yes/no), and number of children (continuous).

Replication in studies with larger colorectal cancer case counts would be required to draw more definitive conclusions.

Studying the relationships between organic food consumption and health is challenging and raises specific issues. First, numerous studies indicate that individuals consuming large amounts of organic food may have healthier profiles and higher overall F&V consumption than the average population [29,47,48], making it challenging to accurately model the relationship between health outcomes and especially organic F&V consumption. Additionally, distinguishing the health impacts of organic and conventional F&V is further complicated by the strong correlation between organic and total F&V intake. It emphasizes the need to distinguish additive from substitution mechanisms. To address this, we applied the residual modeling testing technique, commonly used in nutritional epidemiology [42]. Hence, we isolated the contribution of organic F&V consumption from the total F&V intake by adjusting for the latter. Many confounding factors were considered to focus on the exposure of interest, including sociodemographic, economic, lifestyle, and dietary factors. Furthermore, to attempt to mimic a randomized trial where each participant independently modifies his/her intake of organic F&V (moving between quintiles), we computed survival curves accounting for participants' fixed characteristics, improving causal inference [43,44].

Importantly, in a large United States study [18], no association was observed between highly contaminated F&V consumption and overall cancer risk and most cancer locations, except glioma [20]. However, in the same cohorts, a marked reduction in the risk of death with elevated consumption of low-contaminated F&V was detected [19]. More specifically, heavy consumers of low-contaminated F&V (consuming >4 servings/d) experienced a 36% reduction in the risk of all-cause mortality compared with light consumers (consuming fewer than 1

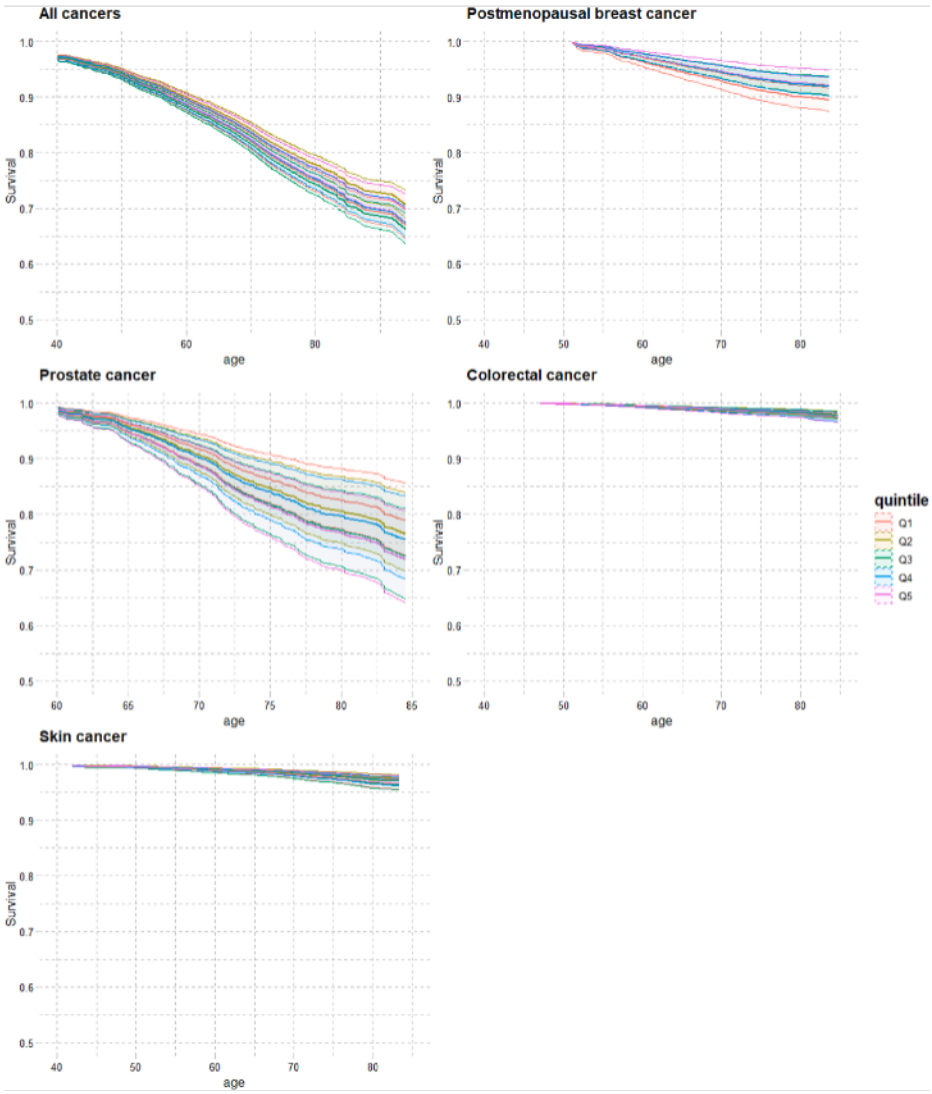


FIGURE 1. Survival marginal curves for overall cancer and by locations. The marginal survival curve is modeled using a standardized regression to evaluate the causal estimand and to calculate the counterfactual probability of developing cancer if all subjects in the population were hypothetically placed in the *i*th quintile of substitution, independent of the covariate profile. Q, quintile.

serving/d) (95% CI: 32%, 41%). However, increasing the consumption of highly contaminated F&V was not associated with reduced mortality risk neither with cancer mortality, suggesting that dietary exposure to pesticide residues may cancel the potential benefits of F&V consumption [19]. In another study conducted in the same NutriNet-Santé cohort sample, we found an inverse association between postmenopausal breast cancer risk and dietary exposure to a pesticide mixture highly correlated with substances authorized in organic farming [22].

Several underlying mechanisms could explain the postmenopausal breast cancer risk reduction associated with the substitution of conventional with organic F&V. Organic consumers experienced lower exposure to pesticide residues [34,49–51]. Pesticides can interact with genetic material, leading to DNA damage or chromosomal abnormalities, or inducing epigenetic modifications [52,53]. Xenobiotic chemicals can also interfere with various nuclear receptors [54], and may result in the onset of chronic diseases, including cancer [52,53]. Then, certain pesticides are known to be endocrine disruptors [52–54], even at low doses, corresponding to pesticide residue levels in food. Organophosphorus compounds can mimic estrogens, and also disrupt carbohydrate and lipid metabolism, whereas organochlorine insecticides, carbamates, and triazines can block androgen receptors, which may be involved in the association observed for breast cancer. Organophosphates inhibit thyroid hormone receptors, and pyrethroids may disrupt progesterone activity. Finally, although the evidence remains limited, gut microbiota may play a role in mediating the association between pesticide exposure and cancer risk [55]. All these mechanisms could be implied in carcinogenicity [53,56] and explain the strong negative association between organic F&V and postmenopausal breast cancer risk, considering that the statistical power for the other sites was limited. Nevertheless, some literature suggests that some biological compounds contained in F&V may mitigate the potential harms related to pesticide exposure, particularly by detoxification mechanisms [53,57].

Some limitations of our study should be acknowledged. First, the observational design restricts causal inference and does not allow for the full exclusion of residual confounding. However, we accounted for a wide range of confounding factors, including sociodemographic and socioeconomic variables, lifestyle, and dietary habits. We also conducted several sensitivity analyses, which demonstrated our findings' robustness. Additionally, the adjusted survival curves enabled the data to be interpreted as if they were from a randomized trial, enhancing causal inference. Second, the generalizability of our results to other populations is limited by the nature of our sample, which consisted mainly of female volunteers. This may have reduced the statistical power of analyses; however, we applied weighting techniques to improve representativeness. Furthermore, this study relies on self-reported data, particularly concerning dietary intakes, which are susceptible to measurement error and desirability bias. Nevertheless, a previous work indicated that organic and conventional consumers exhibited noticeable differences in the concentrations of certain urinary pesticide residues [34], supporting the reliability of our dietary data. Finally, as F&V consumption is a beneficial dietary factor regarding the risk of upper-digestive tract cancer, the number of cases in this study was too small to allow further analyses. Our study also has several strengths. The large sample included participants with various dietary patterns, particularly among heavy organic food consumers. This allowed us to assess the potential role of organic food in the diet with sufficient reliability and statistical power for many cancer sites. We evaluated organic food consumption in detail across a wide range of food items, enabling accurate estimation of the proportion of organic food consumed, overall and by specific food items, including F&V. Then, the use of residual adjustment helped us address the issue of correlation between total and organic F&V consumption. Finally, the prospective design of our study along with innovative modeling methods further strengthens causal inference.

In this study, substitution of conventionally grown F&V with organically grown F&V was associated with a lower risk of postmenopausal breast cancer. The mechanisms underlying these associations need further investigation.

Acknowledgments

We thank Cédric Agäesse (Dietician manager), Alexandre De Sa, Laure Legris, and Laura Chaud (dietitians); Selim Aloui (IT manager), Thi Hong Van Duong, Régis Gatibelza, Amelle Aitelhadj and Aladi Timera (computer scientists); Fabien Szabo de Edelenyi, PhD (data-management supervisor), Julien Allegre, Nathalie Arnault, Laurent Bourhis, and Nicolas Dechamp (data-manager/statisticians); Maria Gomes and Mirette Foham (participants' support); Paola Yvroud, MD (operational coordination); Nadia Khemache (HR and finance manager), Marie Ajanohun, and Tassadit Haddar (administrative support) for their technical support.

Author contributions

The authors' responsibilities were as follows – EK-G, JBaudry: designed the research; SH, MT, JBaudry, DL: conducted the research; JBerlivet, EK-G: performed statistical analysis and drafted the manuscript; EM, BA, DL, MT, BS, SH, JBaudry: contributed to the data interpretation and revised each draft for important intellectual content; EK-G: had full access to all the data in the study, takes responsibility for the integrity of the data and the accuracy of the data analysis, and is the guarantor; and all authors: read and approved the final manuscript.

Conflict of interest

The authors report no conflicts of interest.

Funding

The BioNutriNet project was supported by the French National Research Agency (Agence Nationale de la Recherche) in the context of the 2013 Programme de Recherche Systemes Alimentaires Durables (ANR-13-ALID-0001). The NutriNet-Santé cohort study is funded by the following public institutions: Ministere de la Santé, Santé Publique France, Institut National de la Santé et de la Recherche Médicale (Inserm), Institut National de Recherche pour l'agriculture, l'alimentation et l'environnement (INRAE), Conservatoire National des Arts et Métiers et Université Sorbonne Paris Nord. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data availability

Researchers from public institutions can submit a request to have access to the data for strict reproducibility analysis or for a new collaboration, including information on the institution and a brief description of the project to collaboration@etude-nutrinet-sante.fr. All requests will be reviewed by the steering committee of the NutriNet-Santé study. If the collaboration is accepted, a data access agreement will be necessary and appropriate authorizations from the competent administrative authorities may be needed. In accordance with existing regulations, no personal data will be accessible.

Declaration of generative AI and AI-assisted technologies in the writing process

The author(s) declare that no generative AI or AI-assisted technologies were used in the writing of this manuscript.

Copyright/license for publication statement

Transparency statement

Dr Kesse-Guyot (the guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ajcnut.2026.101284.

References

[1] J. Ferlay, M. Colombet, I. Soerjomataram, D.M. Parkin, M. Piñeros, A. Znaor, et al., Cancer statistics for the year 2020: an overview, Int. J. Cancer. (2021), https://doi.org/10.1002/ijc.33588.

[2] C. Schwedhelm, H. Boeing, G. Hoffmann, K. Aleksandrova, L. Schwingshackl, Effect of diet on mortality and cancer recurrence among cancer survivors: a systematic review and meta-analysis of cohort studies, Nutr. Rev 74 (2016) 737–748.

[3] WCRF/AICR, Cancer preventability estimates for diet, nutrition, body fatness, and physical activity [Internet], World Cancer Research Fund (2017). Cited date: January 2025. Available from: https://www.wcrf.org/dietandcancer/contents.

[4] HBM4EU, Substance report - Pesticides [Internet], 2021. Cited date: January 2025. Available from, https://www.hbm4eu.eu/wp-content/uploads/2022/07/Pesticides_Substance-report.pdf.

[5] IARC, Agents Classified by the IARC Monographs [Internet], 2024 [cited 16 October 2024]. Available from: https://monographs.iarc.who.int/agents-classified-by-the-iarc.

[6] S. Mostafalou, Persistent organic pollutants and concern over the link with insulin resistance related metabolic diseases, Rev. Environ. Contam. Toxicol. 238 (2016) 69–89.

[7] S. Mostafalou, M. Abdollahi, Pesticides: an update of human exposure and toxicity, Arch. Toxicol. 91 (2017) 549–599.

[8] M. Kumar, D.K. Sarma, S. Shubham, M. Kumawat, V. Verma, A. Prakash, et al., Environmental endocrine-disrupting chemical exposure: role in non-communicable diseases, Front. Public Health 8 (2020) 553850.

[9] B. Jiang, J. Pang, J. Li, L. Mi, D. Ru, J. Feng, et al., The effects of organic food on human health: a systematic review and meta-analysis of population-based studies, Nutr. Rev. 82 (9) (2024) 1151–1175.

[10] E. Hakme, M.E. Poulsen, A.D. Lassen, A comprehensive review on pesticide residues in human urine, J. Agric. Food Chem. 72 (2024) 17706–17729.

[11] A. Mie, E. Kesse-Guyot, E. Rembiałkowska, E. Andersen, J. Kahl, P. Grandjean, et al., Human health implications of organic food and organic agriculture [Internet], European Parliamentary Research Service, Scientific Foresight Unit (STOA), 2016, p. 88. Available from: https://www.europarl.europa.eu/RegData/etudes/STUD/2016/581922/EPRS_STU(2016)581922_EN.pdf.

[12] S. Hurtado-Barroso, A. Tresserra-Rimbau, A. Vallverdú-Queralt, R. M. Lamuela-Raventós, Organic food and the impact on human health, Crit. Rev. Food Sci. Nutr. 59 (4) (2019) 704–714.

[13] A. Rahman, P. Baharlouei, E.H.Y. Koh, D.G. Pirvu, R. Rehmani, M. Arcos, et al., A comprehensive analysis of organic food: evaluating nutritional value and impact on human health, Foods 13 (2024) 208.

[14] K.-A. Poulia, D.R. Bakaloudi, M. Alevizou, E. Papakonstantinou, A. Zampelas, M. Chourdakis, Impact of organic foods on chronic diseases and health perception: a systematic review of the evidence, Eur. J. Clin. Nutr. 79 (2) (2025) 90–103.

[15] K.E. Bradbury, A. Balkwill, E.A. Spencer, A.W. Roddam, G.K. Reeves, J. Green, et al., Organic food consumption and the incidence of cancer in a large prospective study of women in the United Kingdom, Br. J. Cancer. 110 (2014) 2321–2326.

[16] J.L.M. Andersen, K. Frederiksen, J. Hansen, C. Kyrø, K. Overvad, A. Tjønneland, et al., Organic food consumption and the incidence of cancer in the Danish diet, cancer and health cohort, Eur. J. Epidemiol. 38 (2023) 59–69.

[17] J. Baudry, K.E. Assmann, M. Touvier, B. Alles, L. Seconda, P. Latino-Martel, et al., Association of frequency of organic food consumption with cancer risk: findings from the NutriNet-Santé prospective cohort study, JAMA Intern. Med. 178 (2018) 1597–1606.

[18] H. Sandoval-Insausti, Y.-H. Chiu, D.H. Lee, S. Wang, J.E. Hart, L. Mínguez-Alarcón, et al., Intake of fruits and vegetables by pesticide residue status in relation to cancer risk, Environ. Int. 156 (2021) 106744.

[19] H. Sandoval-Insausti, Y.-H. Chiu, Y.-X. Wang, J.E. Hart, S.N. Bhupathiraju, L. Mínguez-Alarcón, et al., Intake of fruits and vegetables according to pesticide residue status in relation to all-cause and disease-specific mortality: results from three prospective cohort studies, Environ. Int. 159 (2022) 107024.

[20] D.J. Cote, A.M. Bever, Y.-H. Chiu, H. Sandoval-Insausti, S.A. Smith-Warner, J.E. Chavarro, et al., Pesticide residue intake from fruit and vegetable consumption and risk of glioma, Am. J. Epidemiol. 191 (2022) 825–833.

[21] Y.-H. Chiu, H. Sandoval-Insausti, S.H. Ley, S.N. Bhupathiraju, R. Hauser, E. B. Rimm, et al., Association between intake of fruits and vegetables by pesticide residue status and coronary heart disease risk, Environ. Int. 132 (2019) 105113.

[22] P. Rebouillat, R. Vidal, J.-P. Cravedi, B. Taupier-Letage, L. Debrauwer, L. Gamet-Payrastre, et al., Prospective association between dietary pesticide exposure profiles and postmenopausal breast-cancer risk in the NutriNet-Santé cohort, Int. J. Epidemiol. 50 (2021) 1184–1198.

[23] S. Hercberg, K. Castetbon, S. Czernichow, A. Malon, C. Mejean, E. Kesse, et al., The Nutrinet-Sante Study: a web-based prospective study on the relationship between nutrition and health and determinants of dietary patterns and nutritional status, BMC Public Health 10 (2010) 242.

[24] M. Hagströmer, P. Oja, M. Sjöström, The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity, Public Health Nutr 9 (6) (2006) 755–762.

[25] C. Lassale, S. Péneau, M. Touvier, C. Julia, P. Galan, S. Hercberg, et al., Validity of web-based self-reported weight and height: results of the Nutrinet-Santé study, J. Med. Internet Res. 15 (2013) e152.

[26] A.-C. Vergnaud, M. Touvier, C. Méjean, E. Kesse-Guyot, C. Pollet, A. Malon, et al., Agreement between web-based and paper versions of a socio-demographic questionnaire in the NutriNet-Santé study, Int, J. Public Health. 56 (2011) 407–417.

[27] M. Touvier, C. Mejean, E. Kesse-Guyot, C. Pollet, A. Malon, K. Castetbon, et al., Comparison between web-based and paper versions of a self-administered anthropometric questionnaire, Eur. J. Epidemiol. 25 (2010) 287–296.

[28] Council Regulation (EC) No 834/2007 of 28 June 2007 on organic production and labelling of organic products and repealing Regulation (EEC) No 2092/91 [Internet]. Cited date: January 2025. Available from: https://eur-lex.europa.eu/eli/reg/2007/834/oj.

[29] J. Baudry, C. Méjean, B. Alles, S. Péneau, M. Touvier, S. Hercberg, et al., Contribution of organic food to the diet in a large sample of French adults (the NutriNet-Santé cohort study), Nutrients 7 (2015) 8615–8632.

[30] E. Kesse-Guyot, K. Castetbon, M. Touvier, S. Hercberg, P. Galan, Relative validity and reproducibility of a food frequency questionnaire designed for French adults, Ann. Nutr. Metab. 57 (2010) 153–162.

[31] Etude Nutrinet-Santé, Table de composition des aliments de l'étude Nutrinet-Santé (Nutrinet-Santé Study Food Composition Database), Economica, Paris, 2013. Cited date: January 2025.

[32] High Council of Public Health, Statement related to the revision of the 2017-2021 French Nutrition and Health Programme's dietary guidelines for adults [Internet] Paris, Haut Conseil de la Santé Publique, 2017 February. Available from: https://www.hcsp.fr/explore.cgi/avisrapportsdomaine?clefr=653.

[33] D. Chaltiel, M. Adjibade, V. Deschamps, M. Touvier, S. Hercberg, C. Julia, et al., Programme National Nutrition Santé – guidelines score 2 (PNNS-GS2): development and validation of a diet quality score reflecting the 2017 French dietary guidelines, Br. J. Nutr. 122 (2019) 331–342.

[34] J. Baudry, L. Debrauwer, G. Durand, G. Limon, A. Delcambre, R. Vidal, et al., Urinary pesticide concentrations in French adults with low and high organic food consumption: results from the general population-based NutriNet-Santé, J. Expo. Sci. Environ. Epidemiol. 29 (3) (2019) 366–378.

[35] WHO, ICD-10, International classification of diseases and related health problems, 10th revision, edition, World Health Organization, Geneva, 2010.

[36] P.C. Hallal, C.G. Victora, Reliability and validity of the International Physical Activity Questionnaire (IPAQ), Med. Sci. Sports Exerc. 36 (2004) 556.

[37] INSEE, Definitions and methods [Internet], 2009 [cited 23 February 2015]. Available from: http://www.insee.fr/en/methodes/.

[38] P.C. Austin, D.S. Lee, J.P. Fine, Introduction to the analysis of survival data in the presence of competing risks, Circulation 133 (2016) 601–609.

[39] R.M. Yucel, State of the multiple imputation software, J. Stat. Softw. 45 (2011) v45/i01.

[40] D. Schoenfeld, Partial residuals for the proportional hazards regression model, Biometrika 69 (1982) 239–241.

[41] L. Desquilbet, F. Mariotti, Dose-response analyses using restricted cubic spline functions in public health research, Stat. Med. 29 (2010) 1037–1057.

[42] W. Willett, M.J. Stampfer, Total energy intake: implications for epidemiologic analyses, Am. J. Epidemiol. 124 (1986) 17–27.

[43] E. Syriopoulou, T. Wästerlid, P.C. Lambert, T.M.-L. Andersson, Standardised survival probabilities: a useful and informative tool for reporting regression models for survival data, Br. J. Cancer. 127 (2022) 1808–1815.

[44] A. Sjölander, Estimation of causal effect measures with the R-package stdReg, Eur. J. Epidemiol. 33 (2018) 847–858.

[45] N. Komati, J.-P. Cravedi, J.-M. Lecerf, L.P. Belzunces, D. Tailliez, C. Chambrier, et al., Potential health benefits of a diet rich in organic fruit and vegetables versus a diet based on conventional produce: a systematic review, Nutr. Rev. 83 (3) (2025) e1101–e1114.

[46] X. Theodoridis, A. Papaemmanouil, N. Papageorgiou, A.V. Georgakou, I. Kalaitzopoulou, M. Stamouli, et al., The level of adherence to organic food consumption and risk of cancer: a systematic review and meta-analysis, Life (Basel) 15 (2025) 160.

[47] L. Oates, M. Cohen, L. Braun, Characteristics and consumption patterns of Australian organic consumers, J. Sci. Food Agric. 92 (2012) 2782–2787.

[48] S.B. Petersen, M.A. Rasmussen, M. Strom, T.I. Halldorsson, S.F. Olsen, Sociodemographic characteristics and food habits of organic consumers–a study from the Danish National Birth Cohort, Public Health Nutr 16 (2013) 1810–1819.
[49] C.L. Curl, Characterizing dietary exposure to organophosphate pesticides, incorporating information on organic food consumption, for use in epidemiological research [Internet] [Thesis], 2014. Cited date: January 2025. Available from, https://digital.lib.washington.edu/researchworks/handle/1773/25395.
[50] C. Lu, K. Toepel, R. Irish, R.A. Fenske, D.B. Barr, R. Bravo, Organic diets significantly lower children's dietary exposure to organophosphorus pesticides, Environ. Health Perspect. 114 (2006) 260–263.
[51] L. Oates, M. Cohen, L. Braun, A. Schembri, R. Taskova, Reduction in urinary organophosphate pesticide metabolites in adults after a week-long organic diet, Environ. Res. 132 (2014) 105–111.
[52] S. Mostafalou, M. Abdollahi, Pesticides and human chronic diseases: evidences, mechanisms, and perspectives, Toxicol. Appl. Pharmacol. 268 (2013) 157–177.
[53] M. Sajad, S. Shabir, S.K. Singh, R. Bhardwaj, W.F. Alsanie, A.S. Alamri, et al., Role of nutraceutical against exposure to pesticide residues: power of bioactive compounds, Front. Nutr. 11 (2024) 1342881.
[54] C. Lau, B.D. Abbott, J.C. Corton, M.L. Cunningham, PPARs and xenobiotic-induced adverse effects: relevance to human health, PPAR Res 2010 (2011) 954639.
[55] X. Yuan, Z. Pan, C. Jin, Y. Ni, Z. Fu, Y. Jin, Gut microbiota: an underestimated and unintended recipient for pesticide-induced toxicity, Chemosphere 227 (2019) 425–434.
[56] R. McKinlay, J.A. Plant, J.N. Bell, N. Voulvoulis, Endocrine disrupting pesticides: implications for risk assessment, Environ. Int. 34 (2008) 168–183.
[57] A. Jabłońska – Trypuć, J. Wiater, Protective effect of plant compounds in pesticides toxicity, J. Environ. Health Sci. Eng. 20 (2022) 1035–1045.

Supplementary material
Consumption of organic versus conventional fruit and vegetables in relation to cancer risk: findings from the NutriNet-Santé cohort study - Justine Berlivet et al.

Table des matières

Supplemental Method 1: PNNS-GS2 computation

The simplified version of the PNNS-GS2 score was used in the Cox proportional hazards models, named sPNNS-GS2 score, which includes only the main benchmarks of the recommendations (excluding the components on organic foods, oily fish, white ham, recommended oils and the proportion of vegetable fats). The sPNNS-GS2 comprises 13 components (1). The components, thresholds, scoring system and weighting are summarized in the Table below.

Dietary components	Recommendation	Criteria	Score
Fruit and	At least 5 serv/d,	[0 - 3.5[	0
vegetables	with 1 max as juice and 1	[3.5 - 5[	0.5
(weight=3)	max as dried	[5 - 7.5[	1
		≥7.5	2
Nuts	A handful/d	0	0
(weight=1)		]0 – 0.5[	0.5
		[0.5- 1.5[	1
		≥1.5	0
Pulses	At least 2 serv/w	0 /w	0
(weight=1)		]0-2[/w	0.5
		≥2 /w	1
Whole-grain	Every day	0	0
Food		]0 - 1[	0.5
(weight=2)		[1 - 2[	1
		≥2	1.5
Milk and	2 serv/d	[0 - 0.5[	0
dairy products		[0.5 - 1.5[	0.5
(weight=1)		[1.5 - 2.5[	1
		≥2.5	0
Red meat	Limit consumption	≥750 g/w	-2
(weight=2)		[500 - 750[g/w	-1
		[0 - 500[g/w	0
Processed meat	Limit consumption	≥300 g/w	-2
(weight=3)		[150 - 300[g/w	-1
		[0 - 150[g/w	0
Fish and	2 serv/w	[0 - 1.5[serv /w	0
Seafood		[1.5 - 2.5[serv /w	1
(weight=2)		[2.5 - 3.5[serv /w	0.5
		≥3.5 serv /w	0
Added fat	Avoid overeating	>16% of EIWA [1]	0
(weight=2)		≤16% of EIWA	1.5
Sugary foods	Limit consumption	≥15% of EIWA	-2
(weight=3)		[10-15[% of EIWA	-1
		<10 % of EIWA	0
Sweet drinks	Limit consumption	≥ 750mL mL/d	-2
Beverages		[250 - 750[mL/d	-1
(weight=3)		]0 - 250[mL/d	-0.5
		0 mL/d	0
Alcoholic	Limit consumption	>200 g/d	-2
Beverages		]100-200] g/d	-1
(weight=3)		]0-100] g/d	0
		0 g/d	0.5
Salt	Limit consumption	>12 g/d	-2
(weight=3)		]10-12] g/d	-1
		]8-10] g/d	-0.5
		]6-8] g/d	0
		≤6 g/d	1

Abbreviations: d, day; serv, serving; w week;

[1]EIWA, energy intake without alcohol

Supplemental Method 2: Ultra-processed food consumption exposure
As 24h dietary records are more suitable than FFQ to assess consumption of ultra-processed food (UPF), in the present study, we selected the 24h records at the same period of the FFQ assessment to specifically assess ultra-processed food (UPF) consumption.

The methodology for the classification of foods according to NOVA presented in this supplemental has been described in a previous publication (2).
All food and beverage items of the NutriNet-Santé composition dataset (n>3,500) were categorized into one of the four NOVA groups, a food classification system based on the extent and purpose of industrial food processing (3–5).

The UPF group of the NOVA classification (NOVA4) is the primary focus of this study. Products in this group undergo industrial processes that include, for instance, hydrogenation, hydrolysis, extruding, molding, reshaping, and pre-processing by frying. Flavoring agents, colorings, emulsifiers, humectants, non-sugar sweeteners and other cosmetic additives are often added to these products to imitate the sensorial properties of unprocessed or minimally-processed foods and their culinary preparations. The UPF group is defined by opposition to the other NOVA groups: “unprocessed or minimally-processed foods” (fresh, dried, grounded, chilled, frozen, pasteurized or fermented staple foods such as fruit, vegetables, pulses, rice, pasta, eggs, meat, fish or milk), “processed culinary ingredients” (salt, vegetable oils, butter, sugar and other substances extracted from foods and used in kitchens to transform unprocessed or minimally-processed foods into culinary preparations) and “processed foods” (canned vegetables with added salt, sugar-coated dry fruits, meat products only preserved by salting, cheese and freshly made unpackaged breads, and other products manufactured with the addition of salt, sugar or other substances of the “processed culinary ingredients” group). As previously described (2), home-made and artisanal food preparations were identified and broken down using standardized recipes, and the NOVA classification was applied to their ingredients. Examples of such products as well as examples of distinctions between UPF and products from other NOVA categories are provided below:
Poultry and fish nuggets and sticks and other reconstituted meat products transformed with addition of preservatives other than salt (e.g. nitrites); instant noodles and dehydrated soups; carbonated diet and regular sodas; chocolate with emulsifiers, chewing gums and candies with dyes (confectionery); margarine; instant desserts; most breakfast ‘cereals’, ‘energy’ bars; ‘energy’ drinks; flavored milk drinks; sweet desserts made from fruit with added sugars, artificial flavors and texturizing agents; cooked seasoned vegetables with ready-made sauces; vegetable patties (meat substitutes) containing food additives; ‘health’ and ‘slimming’ products such as powdered or ‘fortified’ meal and dish substitutes.

For instance, salted-only red or white meats are considered as “processed foods” whereas smoked or cured meats with added nitrites and conservatives, such as sausages and ham are classified as “ultra-processed foods”.
Similarly, canned salted vegetables are considered as “processed foods” whereas industrial cooked or fried seasoned vegetables, marinated in industrial sauces with added flavourings are considered as “ultra-processed foods”.
Flavoured breakfast cereals with added emulsifiers, texturizing agents and/or colorants were included in the UPF group. Homemade granola, oatmeal, rye and barley flakes without additives were not considered as ultra-processed.

Regarding soups, canned liquid soups with added salts, herbs and spices are considered as “processed foods” while instant dry soup mixes are considered as “ultra-processed foods”.

The NOVA classification was applied to the foods consumed in the NutriNet-Santé study. To do this, we assigned 5 categories: NOVA 1,2,3,4 and a category for recipes. Indeed, if many mixed foods are classified in NOVA4 some have to be broken down (40% of food items). We used generic recipes to assign NOVA categories to the ingredients of these recipes.

Unprocessed or minimally processed foods (NOVA 1)	fresh, dried, milled, refrigerated, frozen, pasteurized or fermented staple foods such as fruit, vegetables, pulses, rice, pasta, eggs, meat, fish or milk
Processed culinary ingredients (NOVA 2)	salt, vegetable oils, butter, sugar and other substances extracted from food and used in kitchens to transform unprocessed or minimally processed food into culinary preparations
Processed foods (NOVA 3)	canned vegetables with added salt, dried fruit coated with sugar, meat products preserved only by salting, unpackaged fresh cheeses and breads, and other products manufactured with the addition of salt, sugar or other substances from the "processed culinary ingredients" group
Ultra-processed foods (NOVA 4)	Foods treated by various industrial processes (hydrogenation, hydrolysis, extrusion, moulding, reshaping and pre-treatment by frying). Addition of flavouring agents, colourings, emulsifiers, humectants, non-sugar sweeteners and other cosmetic additives.

Supplemental Method 3: Molar sum of DAP metabolites in µmol/g creatinine according to quartiles of share of organic food in the diet and of organic and conventional fruit and vegetable consumption in a subsample of the study with available pesticide biomarkers, NutriNet-Santé cohort, France, 2011-2014 (n=289)[1]

	Q1	Q2	Q3	Q4	P-trend[2]
Share of organic food in the diet	0.00 (0.00)	0.06 (0.06)	0.57 (0.04)	0.78 (0.09)	<.0001
MPs	0.19 (0.42)	0.19 (0.32)	0.18 (0.27)	0.15 (0.32)	0.52
EPs	0.25 (0.47)	0.20 (0.46)	0.12 (0.25)	0.13 (0.31)	0.04
DAPs	0.43 (0.66)	0.39 (0.61)	0.30 (0.40)	0.29 (0.45)	0.06
Organic fruit and vegetable consumption (g/d)	0.00 (0.00)	31.24 (25.10)	218.64 (77.18)	619.79 (252.95)	<.0001
MPs	0.19 (0.41)	0.20 (0.38)	0.16 (0.33)	0.15 (0.18)	0.40
EPs	0.25 (0.46)	0.18 (0.47)	0.09 (0.14)	0.16 (0.35)	0.06
DAPs	0.44 (0.65)	0.39 (0.64)	0.25 (0.37)	0.32 (0.44)	0.06
Conventional fruit and vegetable consumption (g/d)	61.51 (38.74)	189.34 (42.49)	388.23 (79.78)	848.20 (373.30)	<.0001
MPs	0.15 (0.32)	0.18 (0.32)	0.12 (0.17)	0.25 (0.46)	0.22
EPs	0.08 (0.15)	0.15 (0.32)	0.21 (0.38)	0.26 (0.56)	0.003
DAPs	0.23 (0.38)	0.33 (0.43)	0.33 (0.47)	0.51 (0.77)	0.004

Abbreviations: DAPs, dialkylphosphates; EPs, diethylphosphates; MPs, dimethylphosphates; concentratition are in µmol/g creatinine. For a complete description of this analysis, see (6).
[1]Values are means (SD)
[2]P-value refers to contrast from ANOVA across quartiles

Supplemental Method 4: Spearman correlation between fruit and vegetables consumption, nutrients and urinary pesticide residues concentration[1,2]

	total fruit & vegetables	organic fruit & vegetables	conventional fruit & vegetables	Beta-carotene	Vitamin C	Fibers	EPs	DAPs	MPs
total fruit & vegetables	1.00	**0.45**	**0.59**	**0.81**	**0.68**	**0.80**	0.08	0.08	0.05
organic fruit & vegetables		1.00	**-0.30**	**0.42**	**0.30**	**0.44**	-0.07	-0.06	0.03
conventional fruit & vegetables			1.00	**0.44**	**0.41**	**0.41**	**0.14**	**0.13**	0.04
Beta-carotene				1.00	**0.60**	**0.72**	0.01	-0.02	-0.04
Vitamin C					1.00	**0.61**	0.08	0.06	0.02
Fibers						1.00	0.05	0.04	0.00
EPs							1.00	**0.73**	**0.30**
DAPs								1.00	**0.78**
MPs									1.00

Abbreviations: DAPs, dialkylphosphates; EPs, diethylphosphates; MPs, dimethylphosphates;
[1]n=31,179 for all variables except for urinary pesticide residues concentration (n=289)
[2] bold value are P<0.05

Supplemental Method 5: Multiple Imputation

Multiple Imputation by Chained Equations (MICE) is a widely used method for handling missing values in datasets (7,8). This approach involves generating multiple complete (or "imputed") datasets, analysing each dataset separately, and then combining the results to draw final conclusions.

The process for imputing missing values involves two steps. First, simple methods such as using the mean for numeric variables or the most common category for categorical variables are used. Then, variables with missing values are imputed sequentially using regression models. For each variable with missing values, a regression model is built using the other variables as predictors. For example, if a variable Y has missing values, a regression model is fitted with Y as the dependent variable and the other variables as predictors. The missing values of Y are then imputed using this model.

The stepwise imputation process is repeated for a number of iterations, allowing the imputations to converge to stable values. This process is repeated several times (10 times) to create multiple complete datasets. Each imputed dataset is slightly different, reflecting the uncertainty associated with missing values. Each complete dataset is analysed separately using appropriate statistical methods (e.g., regression models, hypothesis testing). The results of the separate analyses are combined to produce final estimates and confidence intervals using rules specific to multiple imputation, such as Rubin's rules (9).

Supplemental Method 6: Marginal Structural Modelling

We employed marginal structural modelling to construct counterfactual models and thus facilitating causal inference (10,11). Original structural models involve a multi-step estimation procedure designed to control the effects of confounding variables. These models considered two weights based on inverse probability weighting, incorporating the likelihood of exposure and the likelihood of censoring, which are combined as follows:

$$\mathrm{SW}^{\mathrm{E,C}} = \mathrm{SW}^{\mathrm{X}} \times \mathrm{SW}^{\mathrm{C}} = \frac{\mathrm{f}(\mathrm{E}_0)}{\mathrm{f}(\mathrm{E}_0|\mathrm{Z}_0)} \times \frac{\Pr[\mathrm{C}=0|\mathrm{E}_0]}{\Pr[\mathrm{C}=0|\mathrm{E}_0,\mathrm{Z}_0]}$$

The combined weights $\mathrm{SW}^{\mathrm{E,C}}$ are calculated by multiplying the stabilized inverse probability of exposure weight (SW^{E}) and the stabilized inverse probability of censoring weight (SW^{C}). These probabilities were derived from linear and logistic regressions, with f(x) representing a probability density function under the assumption of Gaussian distribution (12). The variables E, Z and C were defined as follows: E represents exposure, Z is a vector of covariates, and C is the indicator variable for censoring during the follow-up. Both the numerators were used for stabilisation process and were derived from distinct models. The probability of exposure was estimated using the covariates of the model 1.

Assuming there were no measurement errors during the study, no unmeasured confounding factors, and that the models used for estimating weights were correctly specified, applying the combined weights to the study participants will create a pseudo-population. This pseudo-population will ensure that the distribution of exposure is not influenced by any confounding factors.

Supplemental Method 7: Inverse Probability Weighting
Inverse probability weighting is a statistical method that allows us to partially account for selection bias that is related to missing data at follow-up. The strategy of the method is to attribute more weight to individuals who have similar characteristics as those individuals who were excluded from analyses (13,14). Its implementation includes the following steps:
1.Creation of a logistic regression model that predicts the probability of being included into the study sample with a maximum of accuracy (i.e. with a maximal area under the curve-value), and a Hosmer-Lemeshow-test indicating adequate fit (i.e. $p > 0.05$).
2.Extraction of the predicted probabilities for inclusion into the study sample for each participant.
3.Verification that the highest 10% predicted probabilities do not exceed 50% of the sum of all probabilities.
4.Computation of the inverse of the predicted probabilities (1/probability).
5.Computation of 'stabilized weights': multiplication of each "1/probability" with the sampling proportion (i.e. the proportion of included participants among all participants of the 'source' population);
6.Inclusion of these stabilized weights into the desired statistical analysis (using, for example, the 'weight'- statement of SAS proc logistic®).
The reference population was the entire NutriNet-Santé cohort (in 2014), i.e. 112,468 participants.
In the case of our analyses, the probability to be included in the study sample notably corresponded to the probability to have follow-up data. The basic logistic regression model comprised 'inclusion in our final study sample (yes/no)' as the dependent variable, and the exposure variable of interest as an independent variable. Further independent variables to be included were chosen from the different available sociodemographic, lifestyle, and health-variables on the basis of descriptive tables comparing included and excluded participants: variables that showed a significant (or nearly significant) difference between included and excluded participants were included in the logistic regression model. Moreover, interaction terms were tested in order to improve the predictive potential of the model.

Supplemental Table 1: Substitution of conventional with organic fruits and vegetables in association with the risk of cancer, after removing early cases (n= 241), NutriNet-Santé cohort, France, 2014–2024 (n = 30,938)[1]

	Per 100g increment	P-value	Sex-specific quintiles					P-trend[2]
			Q1	Q2	Q3	Q4	Q5	
All cancers								
n Events			343	271	297	291	275	
Person-Years			44,964	45,891	45,512	46,361	44,804	
Model 1 [3]	0.99 (0.96-1.01)	0.23	1 (ref)	0.86 (0.72-1.01)	1.05 (0.88-1.24)	1.00 (0.85-1.19)	0.94 (0.80-1.11)	0.89
Model 2 [4]	0.98 (0.96-1.01)	0.23	1 (ref)	0.86 (0.72-1.01)	1.05 (0.88-1.25)	1.00 (0.85-1.19)	0.94 (0.80-1.10)	0.9
Model 3 [5]	0.98 (0.96-1.01)	0.23	1 (ref)	0.87 (0.73-1.03)	1.08 (0.91-1.28)	1.02 (0.86-1.22)	0.95 (0.80-1.12)	0.8
Postmenopausal Breast cancer								
n Events			71	47	42	49	42	
Person-Years			20,968	17,720	14,733	17,022	20,181	
Model 1 [3,6]	0.92 (0.86-0.98)	0.01	1 (ref)	0.71 (0.48-1.05)	0.73 (0.48-1.11)	0.77 (0.52-1.13)	0.63 (0.43-0.93)	0.05
Model 2 [4,6]	0.92 (0.86-0.98)	0.01	1 (ref)	0.71 (0.48-1.05)	0.73 (0.48-1.11)	0.77 (0.52-1.13)	0.63 (0.43-0.93)	0.05
Model 3 [5,6]	0.92 (0.86-0.98)	0.01	1 (ref)	0.70 (0.48-1.04)	0.73 (0.48-1.11)	0.76 (0.51-1.13)	0.63 (0.42-0.94)	0.05
Prostate cancer								
n Events			49	43	52	41	52	
Person-Years			11,170	11,411	11,228	11,281	11,139	
Model 1 [3]	1.03 (0.96-1.11)	0.36	1 (ref)	0.69 (0.42-1.14)	1.20 (0.75-1.92)	1.04 (0.65-1.68)	1.13 (0.72-1.77)	0.26
Model 2 [4]	1.03 (0.97-1.10)	0.31	1 (ref)	0.98 (0.63-1.51)	1.37 (0.89-2.11)	1.17 (0.75-1.83)	1.40 (0.94-2.09)	0.06
Model 3 [5]	1.03 (0.97-1.10)	0.31	1 (ref)	0.98 (0.63-1.51)	1.39 (0.90-2.14)	1.18 (0.75-1.84)	1.39 (0.93-2.08)	0.07
Colorectal cancer								
n Events			33	21	23	26	25	
Person-Years			44,964	45,891	45,512	46,361	44,804	
Model 1 [3]	0.99 (0.91-1.08)	0.82	1 (ref)	0.66 (0.37-1.17)	0.81 (0.45-1.46)	0.92 (0.52-1.61)	0.96 (0.57-1.64)	0.74
Model 2 [4]	0.99 (0.91-1.08)	0.83	1 (ref)	0.66 (0.37-1.17)	0.81 (0.45-1.46)	0.92 (0.52-1.61)	0.97 (0.57-1.64)	0.73
Model 3 [5]	0.99 (0.90-1.08)	0.79	1 (ref)	0.67 (0.38-1.21)	0.88 (0.49-1.59)	0.96 (0.55-1.70)	0.97 (0.56-1.68)	0.7
Skin cancer								
n Events			44	28	42	37	41	
Person-Years			44,964	45,891	45,512	46,361	44,804	
Model 1 [3]	1.01 (0.94-1.08)	0.79	1 (ref)	0.68 (0.42-1.12)	1.17 (0.73-1.86)	0.97 (0.61-1.55)	0.99 (0.64-1.52)	0.6
Model 2 [4]	1.01 (0.94-1.08)	0.79	1 (ref)	0.68 (0.42-1.12)	1.17 (0.73-1.87)	0.97 (0.61-1.55)	0.99 (0.64-1.52)	0.6
Model 3 [5]	1.03 (0.96-1.11)	0.36	1 (ref)	0.69 (0.42-1.14)	1.20 (0.75-1.92)	1.04 (0.65-1.68)	1.13 (0.72-1.77)	0.26

[1] HR (Hazard Ratio) and 95% CI (95% confidence interval) are derived from multivariable Cox proportional hazard, Q: Quintile.
[2] P-value of Wald test for quintile as an ordinal variable

[3] Model 1 is a multivariable Cox proportional hazard model adjusted for age (time-scale), sex (when appropriate), educational, occupation, monthly household income, marital status, energy intake without alcohol (continuous variable), alcohol consumption (continuous variable), physical activity, smoking status, number of cigarettes smoked (continuous variable), body mass index (continuous variable), height (continuous variable), family history of cancer, sPNNS-GS2 (continuous variable) and total fruit and vegetables consumption
[4] Model 2 is Model 1 further adjusted for ultra-processed food consumption
[5] Model 3 is Model 1 with the replacement of sPNNS-GS2 with intakes of dairy products, red meat, processed meat, wholegrain products, sugar, sodium, and total lipids (all modeled as continuous variables (g/d))
[6] In the postmenopausal breast cancer models, further adjustments were made for age at menarche (<12 years, ≥12 years), menopausal treatment use (yes/no), and number of children (continuous)

Supplemental Table 2. Substitution of conventional with organic fruits and vegetables in association with the risk of cancer while replacing % organic for "rarely" modalities from 25% to 10%, NutriNet-Santé cohort study, France, 2014–2024 (n = 31,179)[1]

	Per 100g increment	P-value	Sex-specific quintiles					P-trend[2]
			Q1	Q2	Q3	Q4	Q5	
All cancers								
n Events			406	332	335	332	313	
Person-Years			45,005	46,169	45,376	46,276	44,834	
Model 1 [3]	0.98 (0.96-1.00)	0.04	1 (ref)	0.85 (0.73-1.00)	1.00 (0.86-1.18)	0.96 (0.82-1.13)	0.89 (0.76-1.03)	0.48
Model 2 [4]	0.98 (0.95-1.00)	0.04	1 (ref)	0.85 (0.73-1.00)	1.01 (0.86-1.18)	0.96 (0.82-1.13)	0.89 (0.76-1.03)	0.47
Model 3 [5]	0.98 (0.95-1.00)	0.06	1 (ref)	0.86 (0.74-1.01)	1.03 (0.88-1.21)	0.99 (0.84-1.16)	0.90 (0.77-1.05)	0.63
Postmenopausal Breast cancer								
n Events			85	53	50	47	49	
Person-Years			21,396	18,150	14,363	16,658	20,078	
Model 1 [3,6]	0.91 (0.86-0.97)	0.003	1 (ref)	0.60 (0.41-0.88)	0.67 (0.45-1.00)	0.57 (0.39-0.85)	0.61 (0.42-0.87)	0.01
Model 2 [4,6]	0.91 (0.86-0.97)	0.003	1 (ref)	0.60 (0.41-0.87)	0.67 (0.45-1.00)	0.57 (0.39-0.85)	0.61 (0.42-0.87)	0.01
Model 3 [5,6]	0.91 (0.86-0.97)	0.01	1 (ref)	0.61 (0.42-0.88)	0.68 (0.45-1.01)	0.58 (0.39-0.86)	0.61 (0.42-0.89)	0.01
Prostate cancer								
n Events			56	57	53	45	58	
Person-Years			11,121	11,625	11,051	11,345	11,130	
Model 1 [3]	1.01 (0.96-1.07)	0.66	1 (ref)	1.07 (0.72-1.59)	1.26 (0.83-1.92)	1.12 (0.73-1.71)	1.36 (0.93-1.99)	0.12
Model 2 [4]	1.01 (0.95-1.07)	0.67	1 (ref)	1.07 (0.72-1.60)	1.28 (0.84-1.95)	1.12 (0.73-1.72)	1.35 (0.93-1.98)	0.13
Model 3 [5]	1.02 (0.96-1.08)	0.62	1 (ref)	1.09 (0.73-1.63)	1.30 (0.85-1.99)	1.13 (0.74-1.75)	1.39 (0.94-2.06)	0.11
Colorectal cancer								
n Events			36	26	27	29	28	
Person-Years			45,005	46,169	45,376	46,276	44,834	
Model 1 [3]	0.99 (0.92-1.07)	0.83	1 (ref)	0.81 (0.47-1.38)	0.87 (0.49-1.52)	0.98 (0.58-1.68)	1.01 (0.61-1.66)	0.74
Model 2 [4]	0.99 (0.92-1.07)	0.83	1 (ref)	0.81 (0.47-1.38)	0.86 (0.49-1.52)	0.98 (0.58-1.68)	1.01 (0.61-1.67)	0.74
Model 3 [5]	0.99 (0.92-1.07)	0.86	1 (ref)	0.84 (0.49-1.44)	0.96 (0.54-1.69)	1.06 (0.62-1.81)	1.04 (0.62-1.74)	0.63
Skin cancer								
n Events			61	47	41	43	43	
Person-Years			45,005	46,169	45,376	46,276	44,834	
Model 1 [3]	0.96 (0.90-1.02)	0.19	1 (ref)	0.66 (0.42-1.04)	0.96 (0.62-1.47)	0.80 (0.53-1.23)	0.75 (0.50-1.11)	0.33
Model 2 [4]	0.96 (0.90-1.02)	0.19	1 (ref)	0.66 (0.42-1.04)	0.96 (0.62-1.48)	0.80 (0.53-1.23)	0.75 (0.50-1.11)	0.33
Model 3 [5]	0.98 (0.92-1.04)	0.48	1 (ref)	0.67 (0.43-1.06)	0.98 (0.64-1.52)	0.85 (0.56-1.31)	0.84 (0.56-1.26)	0.72

[1]HR (Hazard Ratio) and 95% CI (95% confidence interval) are derived from multivariable Cox proportional hazard, Q: Quintile.

[2] P-value of Wald test for quintile as an ordinal variable

[3] Model 1 is a multivariable Cox proportional hazard model adjusted for age (time-scale), sex (when appropriate), educational, occupation, monthly household income, marital status, energy intake without alcohol (continuous variable), alcohol consumption (continuous variable), physical activity, smoking status, number of cigarettes smoked

(continuous variable), body mass index (continuous variable), height (continuous variable), family history of cancer, sPNNS-GS2 (continuous variable) and total fruit and vegetables consumption

[4] Model 2 is Model 1 further adjusted for ultra-processed food consumption

[5] Model 3 is Model 1 with the replacement of sPNNS-GS2 with intakes of dairy products, red meat, processed meat, wholegrain products, sugar, sodium, and total lipids (all modeled as continuous variables (g/d))

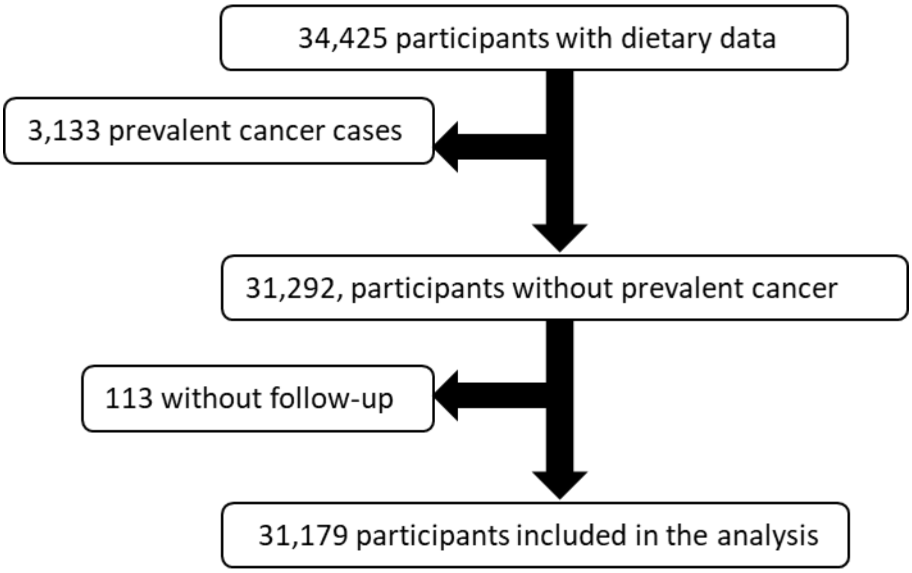
[6] In the postmenopausal breast cancer models, further adjustments were made for age at menarche (<12 years, ≥12 years), menopausal treatment use (yes/no), and number of children (continuous)

Supplemental Table 3. Substitution of conventional with organic fruits and vegetables in association with the risk of overall cancer, NutriNet-Santé cohort study, sensitivity analyses, France, 2014–2024 (n = 31,179)[1]

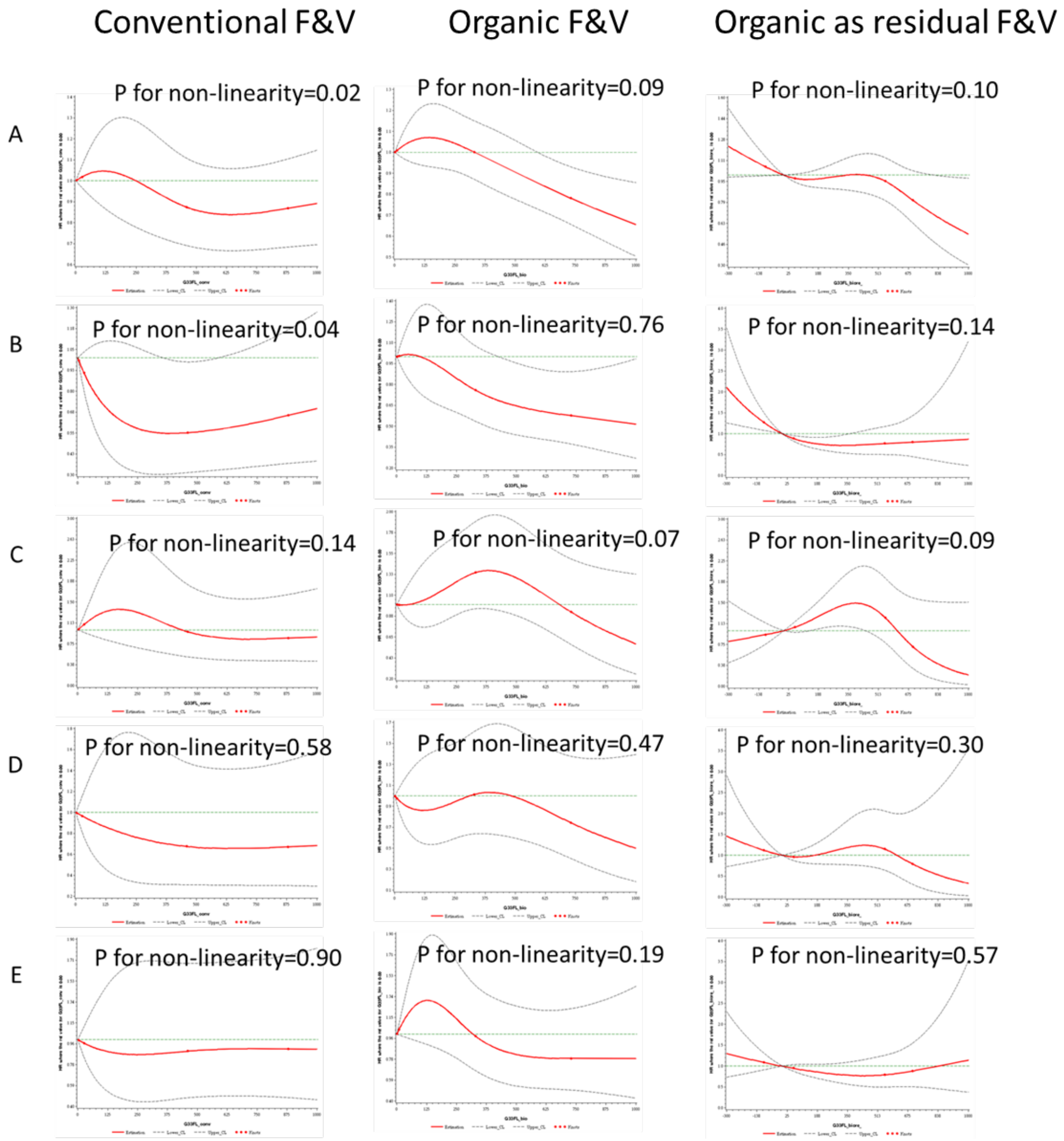
	Per 100g increment	P-value	Sex-specific quintiles					P-trend[3]
			Q1	Q2	Q3	Q4	Q5	
Sensitivity analysis [2,4]								
All cancers	0.93 (0.90-0.96)	<0.0001	1 (ref)	0.75 (0.63-0.90)	0.87 (0.72-1.05)	0.63 (0.51-0.78)	0.58 (0.48-0.70)	<0.0001
Postmenopausal Breast cancer	0.80 (0.72-0.88)	<0.0001	1 (ref)	1.11 (0.68-1.79)	0.39 (0.21-0.73)	0.38 (0.20-0.72)	0.49 (0.28-0.87)	0.0001
Sensitivity analysis 2[2,5]								
All cancers	0.97 (0.95-0.99)	0.02	1 (ref)	0.88 (0.76-1.02)	1.10 (0.95-1.28)	1.03 (0.89-1.19)	0.89 (0.76-1.03)	0.60
Postmenopausal Breast cancer	0.90 (0.85-0.95)	0.0002	1 (ref)	0.76 (0.54-1.08)	0.77 (0.54-1.12)	0.81 (0.57-1.15)	0.56 (0.39-0.82)	0.01
Sensitivity analysis 3 [2,6]								
All cancers	0.98 (0.96-1.01)	0.21	1 (ref)	1.03 (0.86-1.24)	1.12 (0.92-1.35)	1.06 (0.88-1.28)	0.93 (0.78-1.12)	0.59
Postmenopausal Breast cancer	0.90 (0.83-0.97)	0.004	1 (ref)	0.92 (0.60-1.40)	0.79 (0.49-1.26)	0.76 (0.49-1.20)	0.55 (0.35-0.88)	0.01
Sensitivity analysis 4[2,7]								
All cancers	0.97 (0.94-0.99)	0.02	1 (ref)	0.85 (0.73-1.00)	1.02 (0.86-1.20)	0.96 (0.81-1.14)	0.87 (0.72-1.04)	0.46
Postmenopausal Breast cancer	0.92 (0.85-0.99)	0.03	1 (ref)	0.71 (0.49-1.02)	0.71 (0.47-1.06)	0.76 (0.50-1.14)	0.69 (0.44-1.09)	0.17

[1] HR (Hazard Ratio) and 95% CI (95% confidence interval) are derived from multivariable Cox proportional hazard, Q: Quintile.
[2] Cox proportional hazard model adjusted for age (time-scale), sex (when appropriate), educational, occupation, monthly household income, marital status, energy intake without alcohol (continuous variable), alcohol consumption (continuous variable), physical activity, smoking status, number of cigarettes smoked (continuous variable), body mass index (continuous variable), height (continuous variable), family history of cancer, sPNNS-GS2 (continuous variable) and total fruit and vegetables consumption. Models for the risk of postmenopausal breast cancer were further adjusted for age at menarche, menopausal treatment use, and number of children (continuous)
[3] P-value of Wald test for quintile as an ordinal variable
[4] Sensitivity analysis 1 is a model weighted for 2009 census data (including only participants from the French mainland with available variable for weighting, n=30,948, weighted number of cancer cases n=1,194, weighted number of postmenopausal breast cancer cases n=158)
[5] Sensitivity analysis 2 is a marginal structural model
[6] Sensitivity analysis 3 is a model weighted for selection bias using inverse probability weighting
[7]Sensitivity analysis 4 is a multivariable Cox proportional with the replacement of sPNNS-GS2 with consumption of organic dairy products, red meat, processed meat, wholegrain products, conventional dairy products, red meat, processed meat, wholegrain products, and intake of sugar, sodium, and total lipids (continuous variables)

Supplemental Figure 1: Flowchart of the sample



Supplemental Figure 2: Linearity of the association between conventional, organic, and organic as residuals (substitution model) fruits and vegetables and risk of cancers[1]



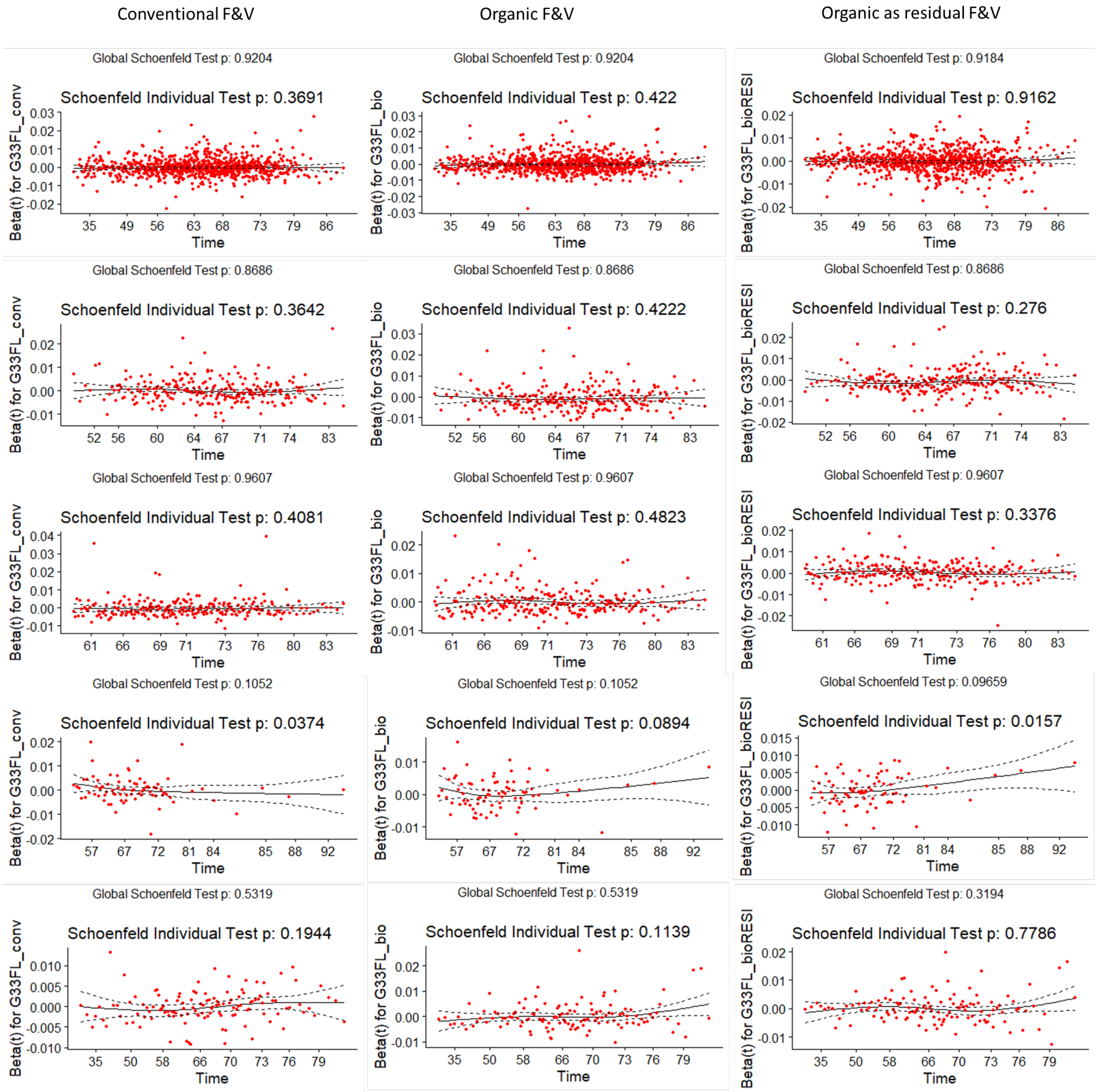
Panel A is cancer all sites, panel B is postmenopausal breast cancer, panel C is prostate cancer, panel D is colorectal cancer, and panel E is skin cancer

Organic fruit and vegetables consumption modelled as residuals (substitution) on total fruits and vegetables consumption corresponds to substitution of conventional with organic ones

[1]The associations with cancer risk are modelled by multivariable Cox proportional hazard model using restricted cubic splines and adjusted for age (time-scale), sex (when appropriate), educational, occupation, monthly household income, marital status, energy intake without alcohol (continuous variable), alcohol consumption (continuous variable), physical activity, smoking status, number of cigarettes smoked (continuous variable), body mass index (BMI) (continuous variable), height (continuous variable), family history of cancer. Models for the risk of postmenopausal breast cancer were further adjusted for age at menarche (<12 years, ≥12 years), menopausal treatment use (yes/no), number of children (continuous)

When analyzing conventional F&V and cancer risk (first column), models are adjusted for organic F&V consumption; conversely, when assessing organic F&V and cancer risk (second column), they are adjusted for conventional F&V intake. The models evaluating organic fruit and vegetables as residuals (substitution) are further adjusted for overall fruit and vegetable consumption (third column)

Supplemental Figure 3: Schoenfeld residuals for the association between conventional, organic, and organic as residuals (substitution model) fruits and vegetables and risk of cancers[1]



Panel A is cancer all sites, panel B is postmenopausal breast cancer, panel C is prostate cancer, panel D is colorectal cancer, and panel E is skin cancer

[1]The associations with cancer risk are modelled by multivariable Cox proportional hazard model using restricted cubic splines and adjusted for age (time-scale), sex (when appropriate), educational, occupation, monthly household income, marital status, energy intake without alcohol (continuous variable), alcohol consumption (continuous variable), physical activity, smoking status, number of cigarettes smoked (continuous variable), body mass index (BMI) (continuous variable), height (continuous variable), family history of cancer. Models for the risk of postmenopausal breast cancer were further adjusted for age at menarche (<12 years, ≥12 years), menopausal treatment use (yes/no), number of children (continuous). The models for the association with organic and conventional fruit and vegetable consumption are further associated for organic fruit and vegetable consumption when assessing conventional F&V and cancer risk (first column), and for conventional fruit and vegetable consumption when assessing organic F&V and cancer risk (second column) and intakes of dairy products, red meat, processed meat, wholegrain products, added sugar, sodium, and total lipids (continuous variable). The model for the association with organic fruit and vegetables as residuals (substitution) is further adjusted for the total consumption of fruit and vegetables and the sPNNS-GS2

References

1. Chaltiel D, Adjibade M, Deschamps V, Touvier M, Hercberg S, Julia C, Kesse-Guyot E. Programme National Nutrition Santé - guidelines score 2 (PNNS-GS2): development and validation of a diet quality score reflecting the 2017 French dietary guidelines. Br J Nutr 2019;122:331–42.
2. Srour B, Fezeu LK, Kesse-Guyot E, Allès B, Debras C, Druesne-Pecollo N, Chazelas E, Deschasaux M, Hercberg S, Galan P, et al. Ultraprocessed Food Consumption and Risk of Type 2 Diabetes Among Participants of the NutriNet-Santé Prospective Cohort. JAMA Intern Med 2020;180:283–91.
3. Monteiro CA, Cannon G, Levy RB, Moubarac JC, Jaime PC, Martins AP, Canella DS, Louzada ML, Parra DC. NOVA. The star shines bright. World Nutr 2016;7:28–38.
4. Monteiro CA, Cannon G, Moubarac JC, Levy RB, Louzada ML, Jaime PC. The UN Decade of Nutrition, the NOVA food classification and the trouble with ultra-processing. Public Health Nutr 2017;1–13.
5. Moubarac JC, Parra DC, Cannon G, Monteiro CA. Food Classification Systems Based on Food Processing: Significance and Implications for Policies and Actions: A Systematic Literature Review and Assessment. CurrObesRep 2014;3:256–72.
6. Baudry J, Debrauwer L, Durand G, Limon G, Delcambre A, Vidal R, Taupier-Letage B, Druesne-Pecollo N, Galan P, Hercberg S, et al. Urinary pesticide concentrations in French adults with low and high organic food consumption: results from the general population-based NutriNet-Santé. J Expo Sci Environ Epidemiol 2018;1.
7. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM, Carpenter JR. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ British Medical Journal Publishing Group; 2009;338:b2393.
8. Graham JW. Missing data analysis: making it work in the real world. Annu Psychol 2009;60:549–76.
9. Rubin. Multiple Imputation for Nonresponse in Surveys [Internet] [Internet]. 1.ed Wiley. 1987 [cited 2025 May 30]. Available from: https://onlinelibrary.wiley.com/doi/book/10.1002/9780470316696
10. Robins JM, Hernan J. Causal Inference: What If (the book) [Internet]. Chapman&Hall/CRC. Boca Raton; 2020. Available from: https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/
11. Hernan MA, Robins JM. Estimating causal effects from epidemiological data. J Epidemiol Community Health 2006;60:578–86.
12. Naimi AI, Moodie EEM, Auger N, Kaufman JS. Constructing Inverse Probability Weights for Continuous Exposures: A Comparison of Methods. Epidemiology 2014;25:292–9.
13. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. AmJEpidemiol 2008;168:656–64.
14. Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. StatMethods MedRes 2011;