BMJ Open

Prospective association between an individual dietary index based on the British Food Standards Agency Nutrient Profiling System and breast cancer risk

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-013718
Article Type:	Research
Date Submitted by the Author:	01-Aug-2016
Complete List of Authors:	Deschasaux, Mélanie; Equipe de Recherche en Epidemiologie Nutritionnelle, Julia, Chantal; Equipe de Recherche en Epidemiologie Nutritionnelle; Hopital Avicenne, Public Health Department Kesse-Guyot, Emmanuelle; Equipe de Recherche en Epidemiologie Nutritionnelle Lécuyer, Lucie; Equipe de Recherche en Epidemiologie Nutritionnelle Adriouch, Solia; Equipe de Recherche en Epidemiologie Nutritionnelle Méjean, Caroline; Equipe de Recherche en Epidemiologie Nutritionnelle Ducrot, Pauline; Equipe de Recherche en Epidemiologie Nutritionnelle Péneau, Sandrine; Equipe de Recherche en Epidemiologie Nutritionnelle Latino-Martel, Paule; Equipe de Recherche en Epidemiologie Nutritionnelle Fezeu, Léopold; Equipe de Recherche en Epidemiologie Nutritionnelle Fassier, Philippine; Equipe de Recherche en Epidemiologie Nutritionnelle Hercberg, Serge; Equipe de Recherche en Epidemiologie Nutritionnelle; Hopital Avicenne, Public Health Department Touvier, Mathilde; Equipe de Recherche en Epidemiologie Nutritionnelle
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Nutrition and metabolism
Keywords:	Breast tumours < ONCOLOGY, nutrient profiling system, nutrition policy, prospective study, food labelling

SCHOLARONE™
Manuscripts

Prospective association between an individual dietary index based on the British Food Standards Agency Nutrient Profiling System and breast cancer risk

Mélanie Deschasaux[1,2*], Chantal Julia[1,3], Emmanuelle Kesse-Guyot[1,2], Lucie Lécuyer[1,2], Solia Adriouch[1], Caroline Méjean[1], Pauline Ducrot[1], Sandrine Péneau[1], Paule Latino-Martel[1,2], Léopold Fezeu[1], Philippine Fassier[1,2], Serge Hercberg[1,2,3], Mathilde Touvier[1,2]

[1] Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), U1153 Inserm, U1125, Inra, Cnam, Paris 13 University, Nutritional Epidemiology Research Team (EREN), Bobigny, France

[2] French network for Nutrition And Cancer Research (NACRe network), www.inra.fr/nacre

[3] Public Health Department, Avicenne Hospital, Bobigny, France

Corresponding author: Mélanie Deschasaux, Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), Nutritional Epidemiology Research Team (EREN), Inserm U1153, Inra U1125, Cnam, Paris 13 University, SMBH Paris 13, 74, rue Marcel Cachin, F-93017, Bobigny Cedex, France; e-mail:m.deschasaux@eren.smbh.univ-paris13.fr; Telephone number: +33 1 48 38 89 44.

Word count: 3287

For peer review only

ABSTRACT

Objectives: French authorities are considering the implementation of a simplified front-of-pack nutrition labeling system on food products to help consumers make healthier food choices. One of the most documented candidates is the Five-Color Nutrition Label, based on the British Food Standards Agency Nutrient Profiling System (FSA-NPS). To assess its potential public health relevance, studies were conducted on the association between the nutritional quality of the diet, as measured at the individual level by an energy-weighted mean of all FSA-NPS scores of foods usually consumed (FSA-NPS DI), and the risk of chronic diseases. The present study aimed at investigating the relationship between the FSA-NPS DI and breast cancer risk in a large French prospective cohort.

Design: prospective study

Setting: population-based study, NutriNet-Santé cohort, France

Participants: 46,864 women aged ≥35y who completed at least three 24h-dietary records during their first 2y of follow-up among whom 555 incident breast cancers were diagnosed between 2009 and 2015.

Primary outcome measure: Associations between individual FSA-NPS DI and breast cancer risk were characterized by multivariable Cox proportional hazard models.

Results: A higher FSA-NPS DI (lower nutritional quality of the diet) was associated with increased breast cancer risk ($HR_{1\text{-point increment}}$=1.06 (1.02, 1.11), P=0.005; $HR_{Q5vs.Q1}$=1.52 (1.11, 2.08), P-trend=0.002). Similar trends were observed in pre- and post-menopausal women ($HR_{1\text{-point increment}}$=1.09 (1.01, 1.18) and 1.05 (1.00, 1.11) respectively).

Conclusions: These results suggested that unhealthy food choices are associated with an increase in breast cancer risk (by 52% for FSA-NPS DI ≥7.7 (Q5) vs. <4.1 (Q1)), supporting the potential public health relevance of developing front-of-pack nutrition labels based on this score.

Keywords: breast cancer, Nutrient Profiling System, nutrition policy, food labelling, prospective study

For peer review only

ARTICLE SUMMARY

Strengths and limitations of this study

- This study examined the association between an indicator of the overall nutritional quality of the diet based on the Food Standards Agency Nutrient Profiling System (FSA-NPS DI) and the incidence of breast cancer.
- This study was performed using data from the NutriNet-Santé study, a large prospective cohort with up-to-date assessment of dietary intakes.
- This study was conducted to assess the public health relevance of the implementation of simplified nutrition labels based on the FSA-NPS on the front-of-pack of food products to help consumers make healthier food choices (as envisioned in France).
- This study included volunteers involved in a long-term cohort study investigating the association between nutrition and health, with overall more health-conscious behaviors and higher professional and/or educational level compared to the general population so that unhealthy dietary behaviors may have been underrepresented.

For peer review only

INTRODUCTION

Breast cancer is the first female cancer worldwide, with 1.7 million new cases diagnosed in 2012, representing 25% of all cancers [1]. According to the estimations of the World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR), around one third of breast cancers could be avoided with appropriate diet, body fatness and physical activity [2].

Thus, nutrition is of particular interest as it is a modifiable individual factor that can be targeted by public health policies. In order to help consumers make healthier food choices, several scientific organizations worldwide have recommended the implementation of a simplified nutrition labeling system on the front-of-pack of food products [3-7]. In France, a five-color labeling system (Five-Color Nutrition Label, 5-CNL) based on the British Food Standards Agency Nutrient Profiling System (FSA-NPS) [8, 9] has been proposed to summarize the overall nutritional quality of food products [10]. The FSA-NPS allows the attribution of a single score to food products according to a limited number of input variables: content per 100g of energy, total sugar, saturated fatty acid (SFA), sodium, fruits and vegetables, dietary fibers and proteins. This scoring system was initially developed and validated in the UK, where it is used for advertising regulation [8, 9, 11, 12], and it has been adapted and validated in the French context [13-16]. At the individual level, the nutritional quality of the diet can be characterized with a dietary index (FSA-NPS DI) based on the FSA-NPS, which has been associated to food and nutrient intakes, nutritional status and adherence to the French nutritional recommendations [17, 18].

To evaluate the relevance and potential public health impact of the 5-CNL adoption, it is important to assess whether there is a relationship between the nutritional quality of the diet at the individual level, as graded by the FSA-NPS DI, and the occurrence of nutrition-related chronic diseases. To our knowledge, our group was the first to investigate the associations between the FSA-NPS DI and health outcomes. Using prospective designs, studies were conducted in the SU.VI.MAX cohort (13,017 participants, 1994-2007) on the associations between the FSA-NPS DI and 13-year weight gain and obesity onset [19], metabolic syndrome [20], cardiovascular diseases [21] and cancer [22]. A higher FSA-NPS DI, reflecting a lower nutritional quality of the diet, was associated with increased risk for all the studied outcomes and, in particular, with an increased risk of cancer overall [22]. No significant association with breast cancer risk was detected in this study [22], but the statistical power was limited for site-specific analyses (n=125 breast cancer cases).

For peer review only

Thus, our objective was to study the association between the FSA-NPS DI (an indicator of the overall nutritional quality of the diet based on a nutrient profiling system) and breast cancer risk, using data from the NutriNet-Santé study, a large prospective cohort with up-to-date assessment of dietary intakes.

For peer review only

METHODS

Study population

The NutriNet-Santé study is a French ongoing web-based cohort launched in 2009 with the objective to study the associations between nutrition and health as well as the determinants of dietary behaviors and nutritional status. This cohort has been previously described in details [23]. Participants aged over 18 years with access to the Internet are continuously recruited since May 2009 among the general population by means of vast multimedia campaigns. All questionnaires are completed online using a dedicated website (www.etude-nutrinet-sante.fr). The NutriNet-Santé study is conducted according to the Declaration of Helsinki guidelines and was approved by the Institutional Review Board of the French Institute for Health and Medical Research (IRB Inserm n°0000388FWA00005831) and the "Commission Nationale de l'Informatique et des Libertés" (CNIL n°908450/n°909216). Electronic informed consent is obtained from each participant (EudraCT no.2013-000929-31).

Data collection

At inclusion, participants fulfilled a set of five questionnaires related to socio-demographic and lifestyle characteristics [24] (e.g. occupation, educational level, smoking status, alcohol consumption, number of children), anthropometrics [25, 26] (e.g. height, weight), dietary intakes (see below), physical activity (validated IPAQ questionnaire) [27], and health status (e.g. personal and family history of diseases, medication use including hormonal treatment for menopause and oral contraception, menopausal status). Participants are invited to complete these five questionnaires every year as part of the follow-up.

Usual dietary intakes were assessed every six months through a series of three non-consecutive validated web-based 24h-dietary records, randomly assigned over a 2-week period (2 weekdays and 1 weekend day) [28-30]. Participants used a dedicated interface of the study website to declare all foods and beverages consumed during a 24h-period: three main meals (breakfast, lunch, dinner) or any other eating occasion. Portion sizes were estimated using validated photographs [31]. Mean daily energy, alcohol and nutrient intakes were estimated using a published French food composition table (>3300 items) [32]. Amounts consumed from composite dishes were estimated using French recipes validated by food and nutrition professionals. Dietary underreporting was identified on the basis of the method proposed by Black [33].

For peer review only

FSA-NPS DI computation

As described previously [9, 13, 34], the FSA-NPS score for all foods (processed and unprocessed) and beverages was computed taking into account nutrient content for 100g. FSA-NPS scores for foods and beverages are based on a discrete continuous scale from -15 (most healthy) to +40 (less healthy) (see **Supplemental file 1**). FSA-NPS score allocates points (0-10) for content in energy (kJ), total sugar (g), SFA (g) and sodium (mg). Points (0-5) are subtracted from the previous sum according to content in fruits and vegetables (%, including legumes and nuts), fibers and proteins. Specific modifications of the score for certain food groups were made, in order to maintain a high consistency with French recommendations, as proposed by the French High Council for Public Health (HCSP) [34].
In a second step, the FSA-NPS DI was computed at the individual level using arithmetic energy-weighted means with the following equation [17], in which FS_i represents the food (or beverage) score, and E_i represents energy intake from this food or beverage:

$$\text{FSA} - \text{NPS DI} = \frac{\sum_{i=1}^{n} FS_i E_i}{\sum_{i=1}^{n} E_i}$$

Thus, increasing FSA-NPS DI reflects decreasing nutritional quality in foods consumed.

Case ascertainment

Participants self-declared health events through the yearly health status questionnaire, through a specific check-up questionnaire for health events (every three months) or at any time through an interface on the study website. Following this declaration, participants are invited to send their medical records (diagnosis, hospitalization, etc.) and, if necessary, the study physicians contact the participants' treating physician or the medical structures to collect additional information. Then, data are reviewed by an independent physician expert committee for the validation of major health events. Cancer cases were classified using the International Chronic Diseases Classification, 10th Revision, Clinical Modification (ICD-10) (33). In this study, all first primary breast cancers diagnosed between the inclusion and August 2015 were considered as cases.

Statistical analyses

So far, 77,034 women without cancer at baseline provided at least three valid 24h-dietary records during their first two years of follow-up. Women aged <35y (n=29,249) were excluded because of a very low susceptibility to develop breast cancer and so were women with a null follow-up (n=921). Thus, 46,864 women were included in the analyses (see flowchart in **Supplementary file 2**).

For peer review only

For each woman, the FSA-NPS DI and usual dietary intakes were calculated taking into account all 24h-dietary records available in their first two years of follow-up. Associations between the FSA-NPS DI (continuous variable and quintiles) and breast cancer risk were characterized (HR and 95%CI) using multivariable Cox proportional hazard models with age as the primary time variable. We confirmed that the assumptions of proportionality were satisfied through examination of the log–log (survival) vs. log–time plots. Tests for linear trends were performed with the ordinal score on quintiles of FSA-NPS DI. Women contributed person-time to the Cox model until the date of cancer diagnosis, the date of last completed questionnaire, the date of death or August 2015, whichever occurred first. Women who reported a cancer other than breast cancer during the study period were included and censored at the date of diagnosis (except basal cell skin carcinoma, not considered as cancer).

Models were adjusted for age (time-scale), BMI (kg/m², continuous), height (cm, continuous), physical activity (high, moderate, low, computed following IPAQ recommendations [35]), smoking status (never smokers, former smokers, occasional smokers, smokers), number of dietary records (continuous), alcohol intake (g/d, continuous), energy intake (without alcohol, g/d, continuous), family history of cancer (yes/no), educational level (<high-school degree, <2 years after high-school degree, ≥2 years after high-school degree), number of biological children (continuous), menopausal status at baseline (pre-menopause, perimenopause, post-menopause), hormonal treatment for menopause (postmenopausal women, yes/no) and oral contraception use (premenopausal women, yes/no).

Sensitivity analyses were performed including only women that provided at least six 24h-dietary records during their first two years of follow-up or excluding cases diagnosed during their first year of follow-up. Analyses were also performed on invasive breast cancer cases only and according to menopausal status. For the latter, women contributed person-time until their age of menopause for premenopausal breast cancer or from their age of menopause for postmenopausal breast cancer. Age at menopause was determined using the yearly health status questionnaires available during the follow-up.

For all covariates except physical activity, ≤ 5% of values were missing and were imputed to the modal value. For physical activity (N=6,328 missing values), a “missing class” was introduced into the models.

All tests were two-sided, and $P<0.05$ was considered statistically significant. SAS version 9.4 (SAS Institute) was used for the analyses.

For peer review only

RESULTS

Between May 2009 and August 2015 (median follow-up time: 4.0y; 174,491 person-years), 555 incident breast cancer cases were diagnosed: 171 premenopausal and 384 postmenopausal; 71.4% ER+/PR+, 14.7% ER-/PR-, 13.6% ER+/PR-, 0.3% ER-/PR+ (data available for 361 cases); 83.6% invasive and 16.4% *in situ* (data available for 463 cases). Mean age at diagnosis was 56.6y (SD=9.2) and mean baseline-to-diagnosis time was 2.4y (SD=1.6). Mean number of dietary records per subject over their first two years of follow-up was 5.9 (SD=2.8).

In **Table 1,** the characteristics of participants at baseline are described overall and according to quintiles of the FSA-NPS DI. Mean FSA-NPS DI was 5.9±2.2 (min=-5.8; max=18.1). Women with a higher FSA-NPS DI (diet of lower nutritional quality), were more likely to be younger, to smoke, to have a higher educational level and to have higher energy or alcohol intakes.

Associations between the FSA-NPS DI and breast cancer risk overall and according to menopausal status are shown in **Table 2**. A direct association was observed between the FSA-NPS DI and breast cancer risk: $HR_{Q5vs.Q1}$=1.52 (95%CI 1.11-2.08), P-trend=0.002; $HR_{per\ 1\text{-}unit\ increment}$=1.06 (1.02-1.11), P=0.005. These associations were similarly observed in premenopausal women ($HR_{Q5vs.Q1}$=2.46 (1.27-4.75), P-trend=0.004; $HR_{per\ 1\text{-}unit\ increment}$=1.09 (1.01-1.18), P=0.03) and in postmenopausal women ($HR_{Q5vs.Q1}$=1.25 (0.85-1.84), P-trend=0.09; $HR_{per\ 1\text{-}unit\ increment}$=1.05 (1.00-1.11), P=0.06), although only trends were observed for the latter.

Similar results were observed when analyses excluded cases diagnosed during their first year of follow-up (425 cases/46,309 non-cases included; $HR_{Q5vs.Q1}$=1.54 (1.08-2.19), P-trend=0.007; $HR_{per\ 1\text{-}unit\ increment}$=1.07 (1.02-1.12), P=0.01) or when analyses were restricted to invasive breast cancers (387 cases/46,309 non-cases; $HR_{Q5vs.Q1}$=1.51 (1.03-2.22), P-trend=0.01; $HR_{per\ 1\text{-}unit\ increment}$=1.06 (1.01-1.12), P=0.03) or to women that provided at least 6 24h-dietary records during their first two years of follow-up (399 cases/25,439 non-cases; $HR_{Q5vs.Q1}$=1.63 (1.11-2.38), P-trend=0.006; $HR_{per\ 1\text{-}unit\ increment}$=1.08 (1.02-1.14), P=0.01) [data not tabulated].

For peer review only

DISCUSSION

In this prospective study conducted in a large sample of women from the French general population, a higher FSA-NPS DI, which reflects a diet composed of food products of lower nutritional quality, was associated with an increased risk of breast cancer.

In a previous study performed in the SU.VI.MAX cohort [22], we observed a direct association between the FSA-NPS DI and cancer risk overall but did not detect a significant association for breast cancer risk specifically, probably due to limited power in site-specific analyses (n=125 breast cancer cases, 13y-follow-up). To our knowledge, no other study investigated the relationship between breast cancer risk and a score that characterizes the nutritional quality of an individual diet based on a nutrient profiling system for foods/beverages consumed.

However a few studies have been conducted on other health outcomes in association with NPS-based dietary scores. While in this study, we used the FSA-NPS as a continuous score at the food/beverage level as a basis for the construction of the FSA-NPS DI at the individual level, the FSA-NPS was also recently used to define a variety score of "healthier" and "less healthy" foods/beverages (Ofcom binary cut-off used for advertising regulation in the UK [12]). This binary indicator was then studied in relation to mortality risk in the Whitehall II cohort [36]. The authors observed that a greater variety of healthier foods as defined with the FSA-NPS Ofcom binary cut-off was associated with a reduced risk of all-cause and cancer mortality while a greater variety of less healthy food was not associated with the studied outcomes. No association was observed when another nutrient profiling system, the SAIN, LIM [37, 38], was used [36].

To our knowledge, the Overall Nutritional Quality Index (ONQI-f) is the only other dietary score based on a nutrient profiling system that has been studied in relation to health outcomes [39]. It was tested in association with chronic diseases and mortality in the Nurses' Health Study and the Health Professionals Follow-up Study [39]. A higher ONQI-f, reflecting a higher nutritional quality of the diet, was associated with a decreased risk of cardiovascular diseases, diabetes and mortality but not with cancer. Some arguments may explain this lack of association observed with cancer: 1) since the ONQI-f is based on 30 nutrients among which few have shown a consistent association with cancer risk, its relevance regarding the cancer outcome may be lower than for other outcomes; 2) dietary intakes were assessed with an aggregated food frequency questionnaire (135-138 items), which provides less precise estimates than 24h-dietary records (as used in our study).

For peer review only

These studies are, to our knowledge, the only ones that investigated the associations between health outcomes and individual dietary indexes calculated from nutrient profiling systems at the food level. Other a priori scores have been designed based on the intake of specific food groups or nutrients and/or other information (e.g. body fatness, physical activity), but not based on a nutrient profiling system at the food/beverage level. These scores were studied prospectively in relation to breast cancer risk and provided relatively contrasted results: 1) scores measuring the adherence to a specific type of diet such as the Mediterranean diet score (no association in prospective cohorts, inverse association in case-control studies [40, 41]) or the Healthy Nordic Food Index (HNFI, no association [42]), 2) scores reflecting the adherence to general nutritional recommendations for the population such as the World Health Organization Healthy Diet Index, WHO HDI [43], the Alternate Healthy Eating Index, AHEI [44], the Recommended Food Score, RFS [44] or the Diet Quality Index revised, DQI-R, [44] (no association overall for these general scores), and 3) scores measuring the adherence to cancer-specific nutritional recommendations such as the WCRF/AICR adherence score (inverse associations [45, 46]) or the American Cancer Society, ACS cancer prevention guidelines score (inverse association [47]). Overall, these studies provided interesting insights into the relationships between nutrition and breast cancer risk. Although these a priori scores and the FSA-NPS DI included similar nutritional components, the approaches differed. The objective behind the FSA-NPS DI construction was not to obtain the best predictive score but to test specifically its association with breast cancer risk, as FSA-NPS is envisioned to serve as a basis for food labelling in the framework of public health policies in several countries such as France and Australia. The FSA-NPS displays several key advantages in a public health context: 1) it grades the nutritional quality of each food/beverage and thus takes into account the variation of nutritional quality between but also within food groups, 2) it has been designed in a perspective of prevention of a large range of chronic diseases (not only breast cancer), and 3) it is easy-to-compute for industrials and public health stakeholders.

Our results are consistent with current evidence regarding the association between nutrition and breast cancer, from epidemiological and mechanistic studies. Indeed, most of the input variables for the FSA-NPS are parameters for which associations with breast cancer have been established, either directly (e.g. dietary fibers [48]) or indirectly, through an association with body fatness which is a major risk factor of postmenopausal breast cancer [48-50] (e.g. energy content, total sugars and SFA as components of energy-dense foods; fruits and vegetables as components of low-energy foods).

For peer review only

Strengths of this study pertained to its prospective design, its large sample size, and the assessment of usual dietary intakes using repeated 24h-dietary records based on a recent food composition database with a large choice of items (>3300). The latter allowed a better insight into the food products consumed compared to studies that used a food frequency questionnaire (more aggregated food items). However, some limitations should be acknowledged. First, caution is needed regarding the extrapolation of these results to the entire French population since this study included volunteers involved in a long-term cohort study investigating the association between nutrition and health, with overall more health-conscious behaviors and higher professional and/or educational level compared to the general population. Thus, unhealthy dietary behaviors may have been underrepresented in this study, which may have weakened the observed associations. Next, statistical power was too limited to investigate the association between the FSA-NPS DI and breast cancer risk according to the characteristics of the tumors (ER/PR).

In conclusion, these results suggest that the consumption of food products of lower nutritional quality (higher FSA-NPS) may be associated with an increased risk of breast cancer. Women in the highest FSA-NPS DI quintile had a 52% increase in breast cancer risk compared to women with the lowest scores (first quintile). The ability of the FSA-NPS DI to predict disease risk (here breast cancer risk) suggests that the FSA-NPS is a valid system to characterize the nutritional quality of foodstuffs and to highlight products with a good nutritional profile that should be promoted and products with a lower nutritional profile that should not. Therefore, this study adds to the scientific evidence that supports the public health relevance of the implementation of front-of-pack nutrition labels based on this score (e.g. 5-CNL) in order to help consumers make healthier food choices.

For peer review only

ACKNOWLEDGEMENTS

The authors thank all the volunteers of the NutriNet-Santé cohort. We extend special thanks to Véronique Gourlet for the statistical analyses and to Nathalie Arnault, Stephen Besseau, Laurent Bourhis, Yasmina Chelghoum, Than Duong Van, Younes Esseddik, Paul Flanzy, Charlie Ménard, Mac Rakotondrazafy, Fabien Szabo, Roland Andrianasolo, Fatoumata Diallo, Cédric Agaesse, Claudia Chahine, Marion Genest and Ludivine Ursule for their technical contribution to the NutriNet-Santé study.

Authors' contribution: The authors' contributions were as follow – MD and MT: designed the research; SH, MT, CJ, EKG: conducted the research; MD and MT: supervised statistical analysis; MD and MT: wrote the paper; CJ, EKG, LL, SA, CM, PD, SP, PLM, LF, PF, SH: contributed to the data interpretation and revised each draft for important intellectual content. All authors read and approved the final manuscript. MD and MT had primary responsibility for the final content.

Funding: Mélanie Deschasaux, Philippine Fassier and Solia Adriouch were funded by PhD grants from the Région Ile-de-France (public funding: Cancéropôle Ile-de-France and CORDDIM). The NutriNet-Santé study was supported by the following public institutions: Ministère de la Santé, Institut de Veille Sanitaire (InVS), Institut National de la Prévention et de l'Education pour la Santé (INPES), Région Ile-de-France (CORDDIM), Institut National de la Santé et de la Recherche Médicale (INSERM), Institut National de la Recherche Agronomique (INRA), Conservatoire National des Arts et Métiers (CNAM) and Université Paris 13. The funders had no role in the design, implementation, analysis, or interpretation of the data.

Competing interests: The authors have no conflict of interest to disclose.

Data sharing: All relevant data are in the manuscript and its supporting files. No additional data available.

For peer review only

REFERENCES

(1) WHO/IARC. All Cancers: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012. GLOBOCAN 2012: 2016. From: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx

(2) WCRF/AICR. Cancer preventability estimates for food, nutrition, body fatness, and physical activity. WCRF/AICR: 2016. From: http://www.wcrf.org/int/cancer-facts-figures/preventability-estimates/cancer-preventability-estimates-diet-nutrition

(3) Food Standards Agency, Welsh Government, Scottish Government, et al. Front of Package Nutrition Labelling: Joint response to consultation. 2013.

(4) Gill T, King L, Vita P, et al. A 'state of the knowledge' assessment of comprehensive interventions that address the drivers of obesity. A Rapid Assessment Prepared for the National Health and Medical Research Council (NHMRC). 2010. Sydney, The Boden Institute of Obesity, Nutrition, Exercise and Eating Disorders, University of Sydney.

(5) Institute of Medecine. Front-of-Package Nutrition Rating Systems and Symbols: Promoting Healthier Choices. 2012. Washington, D.C., The National Academies Press.

(6) WHO. 2008-2013 action plan for the global strategy for the prevention and control of non-communicable diseases. 2009. Geneva, World Health Organization.

(7) WHO Europe. Action Plan for implementation of the European Strategy for the Prevention and Control of Noncommunicable Diseases 2012–2016. 2011. Copenhagen, World Health Organization Regional Office for Europe.

(8) Arambepola C, Scarborough P, Rayner M. Validating a nutrient profile model. *Public Health Nutr* 2008;11:371-378.

(9) Rayner M, Scarborough P, Stockley P, et al. Nutrient profiles: Development of Final Model. Final Report [online]. London, FSA . 2005.

(10) Hercberg S. Propositions pour un nouvel élan de la politique nutritionnelle française de santé publique dans le cadre de la stratégie nationale de santé. 1ère partie: mesures concernant la prévention nutritionnelle [in French]. 2013.

(11) Rayner M, Scarborough P, Stockley L. Nutrient profiles: Applicability of currently proposed model for uses in relation to promotion of food to children aged 5-10 and adults. [online]. London, Foods Standard Agency. 2005.

(12) Rayner M, Scarborough P, Lobstein T. The UK Ofcom Nutrient Profiling Model: Defining 'healthy' and 'unhealthy' foods and drinks for TV advertising to children. London, OfCom . 2009.

(13) Julia C, Kesse-Guyot E, Touvier M, et al. Application of the British Food Standards Agency nutrient profiling system in a French food composition database. *Br J Nutr* 2014;112:1699-1705.

For peer review only

(14) Julia C, Kesse-Guyot E, Ducrot P, et al. Performance of a five category front-of-pack labelling system - the 5-colour nutrition label - to differentiate nutritional quality of breakfast cereals in France. *BMC Public Health* 2015;15:179.

(15) Julia C, Ducrot P, Peneau S, et al. Discriminating nutritional quality of foods using the 5-Color nutrition label in the French food market: consistency with nutritional recommendations. *Nutr J* 2015;14:100.

(16) Haut Conseil de la Santé Publique. Avis relatif à l'information sur la qualité nutritionnelle des produits alimentaires. 2015. Paris, HCSP.

(17) Julia C, Touvier M, Mejean C, et al. Development and validation of an individual dietary index based on the British Food Standard Agency nutrient profiling system in a French context. *J Nutr* 2014;144:2009-2017.

(18) Julia C, Mejean C, Touvier M, et al. Validation of the FSA nutrient profiling system dietary index in French adults-findings from SUVIMAX study. *Eur J Nutr* 2015.

(19) Julia C, Ducrot P, Lassale C, et al. Prospective associations between a dietary index based on the British Food Standard Agency nutrient profiling system and 13-year weight gain in the SU.VI.MAX cohort. *Prev Med* 2015;81:189-194.

(20) Julia C, Fezeu LK, Ducrot P, et al. The Nutrient Profile of Foods Consumed Using the British Food Standards Agency Nutrient Profiling System Is Associated with Metabolic Syndrome in the SU.VI.MAX Cohort. *J Nutr* 2015;145:2355-2361.

(21) Adriouch S, Julia C, Kesse-Guyot E, et al. Prospective association between a dietary quality index based on a nutrient profiling system and cardiovascular disease risk. *Eur J Prev Cardiol* 2016.

(22) Donnenfeld M, Julia C, Kesse-Guyot E, et al. Prospective association between cancer risk and an individual dietary index based on the British Food Standards Agency Nutrient Profiling System. *Br J Nutr* 2015;114:1702-1710.

(23) Hercberg S, Castetbon K, Czernichow S, 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* 2010;10:242.

(24) Vergnaud AC, Touvier M, Mejean C, et al. Agreement between web-based and paper versions of a socio-demographic questionnaire in the NutriNet-Sante study. *Int J Public Health* 2011;56:407-417.

(25) Lassale C, Peneau S, Touvier M, et al. Validity of web-based self-reported weight and height: results of the Nutrinet-Sante study. *J Med Internet Res* 2013;15:e152.

(26) Touvier M, Mejean C, Kesse-Guyot E, et al. Comparison between web-based and paper versions of a self-administered anthropometric questionnaire. *Eur J Epidemiol* 2010;25:287-296.

(27) Craig CL, Marshall AL, Sjostrom M, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003;35:1381-1395.

For peer review only

(28) Lassale C, Castetbon K, Laporte F, et al. Correlations between Fruit, Vegetables, Fish, Vitamins, and Fatty Acids Estimated by Web-Based Nonconsecutive Dietary Records and Respective Biomarkers of Nutritional Status. *J Acad Nutr Diet* 2016;116:427-438.

(29) Lassale C, Castetbon K, Laporte F, et al. Validation of a Web-based, self-administered, non-consecutive-day dietary record tool against urinary biomarkers. *Br J Nutr* 2015;113:953-962.

(30) Touvier M, Kesse-Guyot E, Mejean C, et al. Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. *Br J Nutr* 2011;105:1055-1064.

(31) Le Moullec N, Deheeger M, Preziosi P, et al. Validation du manuel photo utilisé pour l'enquête alimentaire de l'étude SU.VI.MAX. [Validation of the food portion size booklet used in the SU.VI.MAX study] (in French). *Cah Nutr Diet* 1996;31:158-164.

(32) Arnault N, Caillot L, Castetbon K, et al. Table de composition des aliments, étude NutriNet-Santé. [Food composition table, NutriNet-Santé study] (in French). Paris: Les éditions INSERM/Economica; 2013.

(33) Black AE. Critical evaluation of energy intake using the Goldberg cut-off for energy intake:basal metabolic rate. A practical guide to its calculation, use and limitations. *Int J Obes Relat Metab Disord* 2000;24:1119-1130.

(34) Haut Conseil de la Santé Publique. Opinion on information regarding the nutritional quality of foodstuffs. 2015. Paris, HCSP.

(35) IPAQ Group. Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ). 2005.

(36) Masset G, Scarborough P, Rayner M, et al. Can nutrient profiling help to identify foods which diet variety should be encouraged? Results from the Whitehall II cohort. *Br J Nutr* 2015;113:1800-1809.

(37) Afssa. Définition de profils nutritionnels pour l'accès aux allégations nutritionnelles et de santé: propositions et arguments [The setting of nutrient profiles for access to nutrition an health claims: proposals and arguments] (in French). 2008. Paris, Agence française de sécurité des aliments.

(38) Darmon N, Vieux F, Maillot M, et al. Nutrient profiles discriminate between foods according to their contribution to nutritionally adequate diets: a validation study using linear programming and the SAIN,LIM system. *Am J Clin Nutr* 2009;89:1227-1236.

(39) Chiuve SE, Sampson L, Willett WC. The association between a nutritional quality index and risk of chronic disease. *Am J Prev Med* 2011;40:505-513.

(40) Schwingshackl L, Hoffmann G. Adherence to Mediterranean diet and risk of cancer: an updated systematic review and meta-analysis of observational studies. *Cancer Med* 2015;4:1933-1947.

(41) Schwingshackl L, Hoffmann G. Does a Mediterranean-Type Diet Reduce Cancer Risk? *Curr Nutr Rep* 2016;5:9-17.

For peer review only

(42) Li Y, Roswall N, Sandin S, et al. Adherence to a healthy Nordic food index and breast cancer risk: results from a Swedish cohort study. *Cancer Causes Control* 2015;26:893-902.

(43) Cade JE, Taylor EF, Burley VJ, et al. Does the Mediterranean dietary pattern or the Healthy Diet Index influence the risk of breast cancer in a large British cohort of women? *Eur J Clin Nutr* 2011;65:920-928.

(44) Fung TT, Hu FB, McCullough ML, et al. Diet quality is associated with the risk of estrogen receptor-negative breast cancer in postmenopausal women. *J Nutr* 2006;136:466-472.

(45) Harris HR, Bergkvist L, Wolk A. Adherence to the World Cancer Research Fund/American Institute for Cancer Research recommendations and breast cancer risk. *Int J Cancer* 2016;138:2657-2664.

(46) Nomura SJ, Inoue-Choi M, Lazovich D, et al. WCRF/AICR recommendation adherence and breast cancer incidence among postmenopausal women with and without non-modifiable risk factors. *Int J Cancer* 2016;138:2602-2615.

(47) Thomson CA, McCullough ML, Wertheim BC, et al. Nutrition and physical activity cancer prevention guidelines, cancer risk, and mortality in the women's health initiative. *Cancer Prev Res (Phila)* 2014;7:42-53.

(48) Latino-Martel P, Cottet V, Druesne-Pecollo N, et al. Alcoholic beverages, obesity, physical activity and other nutritional factors, and cancer risk: a review of the evidence. Crit Rev Oncol Hematol 99[308], 323. 2016.

(49) WCRF/AICR. Food, Nutrition, Physical Activity and the Prevention of Cancer: a Global Perspective. 2007. Washington, DC: AICR.

(50) WCRF/AICR. Systematic Literature Review - Continuous Update Project Report: The Associations between Food, Nutrition and Physical Activity and the Risk of Breast Cancer. 2010. Washington, DC: AICR.

For peer review only

Table 1 Baseline characteristics of the study population overall and according to quintiles of the FSA-NPS DI, NutriNet-Santé Cohort, France, 2009-2015

		Quintiles of the FSA-NPS DI					
	All women (n=46,864)	Q1 (n=9,349)	Q2 (n=9,395)	Q3 (n=9,387)	Q4 (n=9,415)	Q5 (n=9,318)	
	N (%) Mean±SD	N (%) Mean±SD	N (%) Mean±SD	N (%) Mean±SD	N (%) Mean±SD	N (%) Mean±SD	P-trend[a]
FSA-NPS DI	5.9±2.2	2.7±1.2	4.8±0.4	6.0±0.3	7.1±0.4	9.0±1.1	<.0001
Age, years	50.8±9.7	53.4±9.6	52.6±9.4	51.2±9.7	49.6±9.4	47.1±8.9	<.0001
Educational level							<.0001
< high-school degree	11269 (24.1)	2658 (28.4)	2345 (25.0)	2172 (23.1)	2083 (22.1)	2011 (21.6)	
≥high-school degree to < 2y after high-school degree	7834 (16.7)	1567 (16.8)	1579 (16.8)	1570 (16.7)	1556 (16.5)	1562 (16.8)	
≥ 2y after high-school degree	27761 (59.2)	5124 (54.8)	5471 (58.2)	5645 (60.1)	5776 (61.3)	5745 (61.6)	
Smoking status							<.0001
Non-smokers	22528 (48.1)	4630 (49.5)	4706 (50.1)	4615 (49.2)	4504 (47.8)	4073 (43.7)	
Former smokers	17904 (38.2)	3744 (40.0)	3640 (38.7)	3561 (37.9)	3527 (37.5)	3432 (36.8)	
Occasional smokers[b]	1622 (3.5)	257 (2.7)	302 (3.2)	336 (3.6)	350 (3.7)	377 (4.0)	
Smokers	4810 (10.3)	718 (7.7)	747 (7.9)	875 (9.3)	1034 (11.0)	1436 (15.4)	
Physical activity[c]							0.08
Low	13955 (34.4)	3312 (41.1)	2979 (36.4)	2800 (34.5)	2569 (31.5)	2295 (28.6)	
Moderate	17062 (42.1)	3224 (40.0)	3462 (42.4)	3487 (43.0)	3522 (43.2)	3367 (41.9)	
High	9519 (23.5)	1521 (18.9)	1732 (21.2)	1829 (22.5)	2066 (25.3)	2371 (29.5)	
BMI, kg/m²	24.1±4.8	24.5±4.9	24.1±4.7	23.9±4.5	23.9±4.6	24.3±5.2	<.0001
Height, cm	163.4±6.1	162.8±6.0	163.0±6.0	163.4±6.1	163.7±6.0	164.2±6.1	<.0001
Energy intake without alcohol, kcal/d	1710±385	1510±331	1648±334	1721±344	1792±370	1882±429	<.0001
Alcohol intake, g/d	6.5±9.1	4.5±7.7	5.9±8.4	6.8±9.0	7.4±9.5	7.9±10.5	<.0001
Number of biological children	1.8±1.2	1.8±1.2	1.8±1.1	1.8±1.1	1.8±1.1	1.8±1.2	<.0001
Family history of cancer (yes)	21158 (45.2)	4446 (47.6)	4393 (46.8)	4288 (45.7)	4185 (44.4)	3846 (41.3)	0.9
Menopausal status							0.5
Pre-menopause	23940 (51.1)	3767 (40.3)	4078 (43.4)	4637 (49.4)	5296 (56.2)	6162 (66.1)	
Perimenopause	3997 (8.5)	807 (8.6)	871 (9.3)	807 (8.6)	795 (8.4)	717 (7.7)	
Post-menopause	18927 (40.4)	4775 (51.1)	4446 (47.3)	3943 (42.0)	3324 (35.3)	2439 (26.2)	
Hormonal treatment for menopause use (yes)[d]	4068 (17.7)	1025(18.4)	978 (18.4)	806 (17.0)	732 (17.8)	527 (16.7)	0.04

[a] P value for the comparison between quintiles of FSA-NPS DI, by χ^2 tests from age-adjusted ordinal polytomous logistic regressions

[b] Occasional smokers smoke less than once a day

[c] Data available for 40,536 women

[d] Among women in peri- or post-menopause (n=22,924)

For peer review only

Table 2 Associations between the FSA-NPS DI and breast cancer risk, from multivariable Cox proportional hazards models[a], NutriNet-Santé Cohort, France, 2009-2015

FSA-NPS DI	N for cases/ non-cases	HR	95%CI	P-trend
Overall				
Continuous score	555/46,309	1.06	1.02, 1.11	0.005
Quintiles[b]				
Q1	82/9,267	1.00	(ref)	0.002
Q2	122/9,273	1.43	1.08, 1.90	
Q3	117/9,270	1.43	1.07, 1.91	
Q4	138/9,277	1.79	1.35, 2.38	
Q5	96/9,222	1.52	1.11, 2.08	
Premenopausal women				
Continuous score	171/23,483	1.09	1.01, 1.18	0.03
Quintiles[b]				0.004
Q1	12/3,667	1.00	(ref)	
Q2	28/3,982	1.92	0.97, 3.79	
Q3	31/4,558	1.89	0.96, 3.71	
Q4	52/5,204	2.76	1.45, 5.26	
Q5	48/6,072	2.46	1.27, 4.75	
Postmenopausal women				
Continuous score	384/27,188	1.05	1.00, 1.11	0.06
Quintiles[b]				0.09
Q1	70/6,416	1.00	(ref)	
Q2	94/6,173	1.36	0.99, 1.86	
Q3	86/5,578	1.37	0.99, 1.89	
Q4	86/5,028	1.57	1.13, 2.18	
Q5	48/3,993	1.25	0.85, 1.84	

[a] Models were adjusted for age (time-scale), BMI (kg/m², continuous), height (cm, continuous), physical activity (high, moderate, low), smoking status (never smokers, former smokers, occasional smokers, smokers), numbers of dietary records (continuous), alcohol intake (g/d, continuous), energy intake (without alcohol, g/d, continuous), family history of cancer (yes/no), educational level (<high-school degree, <2 years after high-school degree, ≥2 years after high-school degree), number of biological children (continuous), menopausal status at baseline (pre-menopause, perimenopause, post-menopause), hormonal treatment for menopause (postmenopausal women, yes/no) and oral contraception use (premenopausal women, yes/no).

[b] Cut-offs for quintiles of the FSA-NPS DI were 4.1/5.4/6.5/7.7

For peer review only

Supplemental file 1: FSA NPS score computation at food/beverage level

Points are allocated according to the nutrient content for 100g of foods or beverages.
Points are allocated for 'Negative' nutrients (A points) and can be balanced according to 'Positive' nutrients (C points).

A points
Total A points = (points for energy) + (points for saturated fat) + (points for total sugar) + (points for sodium)

Points	Energy (kJ)	Saturated Fat (g)	Total Sugars (g)	Sodium (mg)
0	≤ 335	≤ 1	≤ 4.5	≤ 90
1	> 335	> 1	> 4.5	> 90
2	> 670	> 2	> 9	> 180
3	> 1005	> 3	> 13.5	> 270
4	> 1340	> 4	> 18	> 360
5	> 1675	> 5	> 22.5	> 450
6	> 2010	> 6	> 27	> 540
7	> 2345	> 7	> 31	> 630
8	> 2680	> 8	> 36	> 720
9	> 3015	> 9	> 40	> 810
10	> 3350	> 10	> 45	> 900

C points
Total C points = (points for fruits and vegetables) + (points for fibers) + (points for proteins)

Points	Fruits, Vegetables (%)	Fiber (g) *	Protein (g)
0	≤ 40	≤ 0.7	≤ 1.6
1	> 40	> 0.7	> 1.6
2	> 60	> 1.4	> 3.2
3	-	> 2.1	> 4.8
4	-	> 2.8	> 6.4
5	> 80	> 3.5	> 8.0

*FSA score allocates different thresholds for fibers, depending on the measurement method used. We used NSP cut-offs to compute fibers score.

Overall score computation

- If Total A points <11, then FSA score =Total A points – Total C points
- If Total A points ≥11,
  - If points for fruits and vegetables =5, then FSA score =Total A points – Total C points
  - Else if points for fruits and vegetables <5, then FSA score = Total A points – (points for fiber + points for fruits and vegetables).

For 100g of a given food, the percentage of fruits and vegetables is obtained by summing up the amount (in grams) of all fruits, legumes and vegetables (including oleaginous fruits, dried fruits and olives) contained in this food.

For peer review only

Exceptions were made for cheese, fat, and drinks to better rank them according to their nutrient profile, consistently with nutritional recommendations:

Score computation for cheese

For cheese, the score takes in account the protein content, whether the A score reaches 11 or not, i.e.: FSA score =Total A points – Total C points

Score computation for fat

For fat, the grid for point attribution is based on the percentage of saturated fat among total lipids and has a six-point homogenous ascending step, as shown thereafter:

Points	**Saturated Fat/Lipids (%)**
0	< 10
1	< 16
2	< 22
3	< 28
4	< 34
5	< 40
6	< 46
7	< 52
8	< 58
9	< 64
10	≥ 64

Score computation for drinks

For drinks, the grids for point attribution regarding energy, total sugars and fruits and vegetables (%) were modified. The attribution of points for sugars takes into account the presence of sweeteners, in which case the grid maintains the total sugar score to 1 (instead of 0).

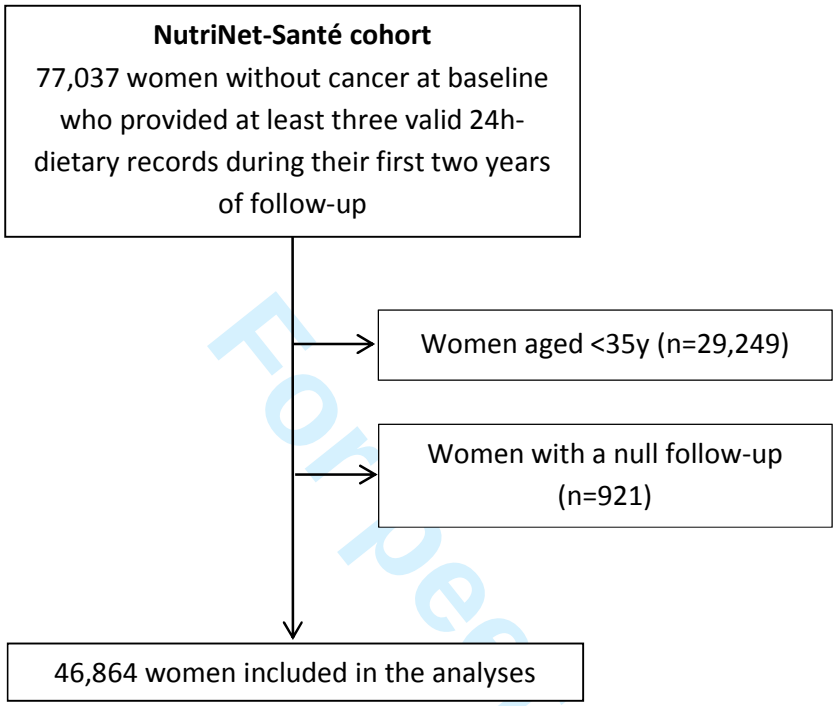
Points	**Energy (kJ)**	**Total Sugar (g)**	**Fruits, Vegetables (%)**
0	≤ 0	≤ 0	< 40
1	≤ 30	≤ 1.5	
2	≤ 60	≤ 3	> 40
3	≤ 90	≤ 4.5	
4	≤ 120	≤ 6	> 60
5	≤ 150	≤ 7.5	
6	≤ 180	≤ 9	
7	≤ 210	≤ 10.5	
8	≤ 240	≤ 12	
9	≤ 270	≤ 13.5	
10	> 270	> 13.5	> 80

Milk and vegetable milk are not concerned by this exception. Their scores are computed using the overall score computation system.

For peer review only

SUPPLEMENTARY FILE 2

Participants' flowchart



For peer review only

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cohort studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1; 2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	7
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	n.a
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-8
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6-8
Bias	9	Describe any efforts to address potential sources of bias	7-8
Study size	10	Explain how the study size was arrived at	7
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	7-8
		(*b*) Describe any methods used to examine subgroups and interactions	7-8
		(*c*) Explain how missing data were addressed	7-8
		(*d*) If applicable, explain how loss to follow-up was addressed	7-8
		(*e*) Describe any sensitivity analyses	8
Results			

For peer review only

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	7; 9
		(b) Give reasons for non-participation at each stage	7
		(c) Consider use of a flow diagram	Supplementary file 2
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9; Table 1
		(b) Indicate number of participants with missing data for each variable of interest	8
		(c) Summarise follow-up time (eg, average and total amount)	9
Outcome data	15*	Report numbers of outcome events or summary measures over time	9
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	9; Table 2
		(*b*) Report category boundaries when continuous variables were categorized	8; Table 2
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n.a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	9
Discussion			
Key results	18	Summarise key results with reference to study objectives	10
Limitations			
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	10-12
Generalisability	21	Discuss the generalisability (external validity) of the study results	10-12
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	13

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

For peer review only

BMJ Open

Prospective association between an individual dietary index based on the British Food Standards Agency Nutrient Profiling System and breast cancer risk

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-013718.R1
Article Type:	Research
Date Submitted by the Author:	17-Nov-2016
Complete List of Authors:	Deschasaux, Mélanie; Equipe de Recherche en Epidemiologie Nutritionnelle, Julia, Chantal; Equipe de Recherche en Epidemiologie Nutritionnelle; Hopital Avicenne, Public Health Department Kesse-Guyot, Emmanuelle; Equipe de Recherche en Epidemiologie Nutritionnelle Lécuyer, Lucie; Equipe de Recherche en Epidemiologie Nutritionnelle Adriouch, Solia; Equipe de Recherche en Epidemiologie Nutritionnelle Méjean, Caroline; Equipe de Recherche en Epidemiologie Nutritionnelle Ducrot, Pauline; Equipe de Recherche en Epidemiologie Nutritionnelle Péneau, Sandrine; Equipe de Recherche en Epidemiologie Nutritionnelle Latino-Martel, Paule; Equipe de Recherche en Epidemiologie Nutritionnelle Fezeu, Léopold; Equipe de Recherche en Epidemiologie Nutritionnelle Fassier, Philippine; Equipe de Recherche en Epidemiologie Nutritionnelle Hercberg, Serge; Equipe de Recherche en Epidemiologie Nutritionnelle; Hopital Avicenne, Public Health Department Touvier, Mathilde; Equipe de Recherche en Epidemiologie Nutritionnelle
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Nutrition and metabolism
Keywords:	Breast tumours < ONCOLOGY, nutrient profiling system, nutrition policy, prospective study, food labelling

SCHOLARONE™
Manuscripts

Prospective association between an individual dietary index based on the British Food Standards Agency Nutrient Profiling System and breast cancer risk

Mélanie Deschasaux[1,2*], Chantal Julia[1,3], Emmanuelle Kesse-Guyot[1,2], Lucie Lécuyer[1,2], Solia Adriouch[1], Caroline Méjean[1], Pauline Ducrot[1], Sandrine Péneau[1], Paule Latino-Martel[1,2], Léopold K Fezeu[1], Philippine Fassier[1,2], Serge Hercberg[1,2,3], Mathilde Touvier[1,2]

[1] Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), U1153 Inserm, U1125, Inra, Cnam, Paris 13 University, Nutritional Epidemiology Research Team (EREN), Bobigny, France

[2] French network for Nutrition And Cancer Research (NACRe network), www.inra.fr/nacre

[3] Public Health Department, Avicenne Hospital, Bobigny, France

Corresponding author: Mélanie Deschasaux, Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), Nutritional Epidemiology Research Team (EREN), Inserm U1153, Inra U1125, Cnam, Paris 13 University, SMBH Paris 13, 74, rue Marcel Cachin, F-93017, Bobigny Cedex, France; e-mail:m.deschasaux@eren.smbh.univ-paris13.fr; Telephone number: +33 1 48 38 89 44.

Word count: 4000

For peer review only

ABSTRACT

Objectives: French authorities are considering the implementation of a simplified front-of-pack nutrition labeling system on food products to help consumers make healthier food choices. One of the most documented candidates is the Five-Color Nutrition Label, based on the British Food Standards Agency Nutrient Profiling System (FSA-NPS). The FSA-NPS is calculated for each food/beverage based on the amount per 100g of energy, total sugar, saturated fatty acid, sodium, dietary fibers, proteins, and % of fruits and vegetables. To assess its potential public health relevance, studies were conducted on the association between the nutritional quality of the diet, as measured at the individual level by an energy-weighted mean of all FSA-NPS scores of foods usually consumed (FSA-NPS DI), and the risk of chronic diseases. The present study aimed at investigating the relationship between the FSA-NPS DI and breast cancer risk in a large prospective cohort.

Design: prospective study

Setting: population-based, NutriNet-Santé cohort, France

Participants: 46,864 women aged ≥35y who completed at least three 24h-dietary records during their first 2y of follow-up among whom 555 incident breast cancers were diagnosed between 2009 and 2015.

Primary outcome measure: Associations between individual FSA-NPS DI and breast cancer risk were characterized by multivariable-adjusted Cox proportional hazard models.

Results: A higher FSA-NPS DI (lower nutritional quality of the diet) was associated with increased breast cancer risk ($HR_{1\text{-point increment}}$=1.06 (1.02, 1.11), P=0.005; $HR_{Q5vs.Q1}$=1.52 (1.11, 2.08), P-trend=0.002). Similar trends were observed in pre- and post-menopausal women ($HR_{1\text{-point increment}}$=1.09 (1.01, 1.18) and 1.05 (1.00, 1.11) respectively).

Conclusions: These results suggested that unhealthy food choices are associated with an increase in breast cancer risk (by 52% for FSA-NPS DI ≥7.7 (Q5) vs. <4.1 (Q1)), supporting the potential public health relevance of developing front-of-pack nutrition labels based on the FSA-NPS.

Keywords: breast cancer, Nutrient Profiling System, nutrition policy, food labelling, prospective study

For peer review only

ARTICLE SUMMARY

Strengths and limitations of this study

- This study examined the association between an indicator of the overall nutritional quality of the diet based on the Food Standards Agency Nutrient Profiling System (FSA-NPS DI) and the incidence of breast cancer.
- This study was performed using data from the NutriNet-Santé study, a large prospective cohort with up-to-date assessment of dietary intakes.
- This study was conducted to assess the public health relevance of the implementation of simplified nutrition labels based on the FSA-NPS on the front-of-pack of food products to help consumers make healthier food choices (as envisioned in France).
- This study included volunteers involved in a long-term cohort study investigating the association between nutrition and health, with overall more health-conscious behaviors and higher professional and/or educational level compared to the general population so that unhealthy dietary behaviors may have been underrepresented.

For peer review only

INTRODUCTION

Breast cancer is the most common female cancer worldwide, with 1.7 million new cases diagnosed in 2012, representing 25% of all cancers [1]. According to the estimations of the World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR), around one third of breast cancers could be avoided with appropriate diet, body fatness and physical activity [2].

Nutrition has therefore the potential to be a key factor in breast cancer prevention since it can be modified at the individual level and thus can be targeted by public health policies. To help consumers make healthier food choices, several scientific organizations worldwide have recommended the implementation of a simplified nutrition labeling system on the front-of-pack of food products [3-7]. In France, a five-color labeling system (Five-Color Nutrition Label, 5-CNL) based on the British Food Standards Agency Nutrient Profiling System (FSA-NPS) [8;9] has been proposed to summarize the overall nutritional quality of food products [10]. The FSA-NPS attributes a single score to food products based on a limited number of input variables: amount per 100g of energy, total sugar, saturated fatty acid (SFA), sodium, fruits and vegetables, dietary fibers and proteins. This scoring system was initially developed and validated in the UK, where it is used for advertising regulation [8;9;11;12], and it has been adapted and validated in the French context [13-16]. At the individual level, the nutritional quality of the diet can be characterized with a dietary index based on the FSA-NPS (FSA-NPS DI). The FSA-NPS DI has been associated to food and nutrient intakes, nutritional status and adherence to the French nutritional recommendations [17;18].

To evaluate the relevance and potential public health impact of the 5-CNL adoption, it is important to assess whether there is a relationship between the nutritional quality of food choices at the individual level, as graded by the FSA-NPS DI, and the occurrence of nutrition-related chronic diseases. To our knowledge, our group was the first to investigate the associations between the FSA-NPS DI and health outcomes. Using prospective designs, studies were conducted in the SU.VI.MAX cohort (13,017 participants, 1994-2007) on the associations between the FSA-NPS DI and 13-year weight gain/obesity onset [19], metabolic syndrome [20], cardiovascular diseases [21] and cancer [22]. A higher FSA-NPS DI, reflecting a diet of lower nutritional quality, was associated with an increased risk for all the studied outcomes and, in particular, with an increased risk of cancer overall [22]. No significant association with breast cancer risk was detected in this study [22], but the statistical power was limited for site-specific analyses (n=125 breast cancer cases).

For peer review only

Thus, our objective was to study the association between the FSA-NPS DI (an indicator of the nutritional quality of the diet based on a nutrient profiling system) and breast cancer risk, using data from NutriNet-Santé, a large prospective cohort with up-to-date assessment of dietary intakes.

For peer review only

METHODS

Study population

The NutriNet-Santé study is a French ongoing web-based cohort launched in 2009 with the objective to study the associations between nutrition and health as well as the determinants of dietary behaviors and nutritional status. This cohort has been previously described in details [23]. Participants aged ≥18y with access to the Internet are continuously recruited since May 2009 among the general population by means of vast multimedia campaigns. All questionnaires are completed online though a dedicated website (www.etude-nutrinet-sante.fr). The NutriNet-Santé study is conducted according to the Declaration of Helsinki guidelines and was approved by the Institutional Review Board of the French Institute for Health and Medical Research (IRB Inserm n°0000388FWA00005831) and the "Commission Nationale de l'Informatique et des Libertés" (CNIL n°908450/n°909216). Electronic informed consent is obtained from each participant (EudraCT no.2013-000929-31).

Data collection

At inclusion, participants fulfilled a set of five questionnaires on socio-demographic and lifestyle characteristics [24] (e.g. occupation, educational level, smoking status, alcohol consumption, number of children), anthropometrics [25;26] (e.g. height, weight), dietary intakes (see below), physical activity (validated IPAQ questionnaire) [27], and health status (e.g. personal and family history of diseases, medication use including hormonal treatment for menopause and oral contraception, menopausal status). Follow-up of participants began when participants answered their last baseline questionnaire. The date of completion of the last baseline questionnaire is thus used as inclusion date. Participants are then invited to complete these five baseline questionnaires every year as part of the follow-up.

Dietary intakes were assessed at baseline and every six months through series of three non-consecutive validated web-based 24h-dietary records, randomly assigned over a 2-week period (2 weekdays and 1 weekend day) [28-30]. Thus, over the first two years of follow-up, up to five series of three 24h-dietary records could have been completed. To be considered as valid, a series must have included at least two out of three 24h dietary records. Participants used a dedicated interface of the study website to declare all foods and beverages consumed during a 24h-period: three main meals (breakfast, lunch, dinner) or any other eating occasion. Portion sizes were estimated using validated photographs [31]. Mean daily energy, alcohol and nutrient intakes were estimated using a published French food composition table (>3300 items) [32] and a weighting for week days and week-end days. Amounts consumed from

For peer review only

composite dishes were estimated using French recipes validated by food and nutrition professionals. Dietary underreporting was identified on the basis of the method proposed by Black [33].

FSA-NPS DI computation

As described previously [9;13;34], the FSA-NPS score for all foods (processed and unprocessed) and beverages was computed based on the nutrient content for 100g. FSA-NPS scores for foods and beverages are based on a discrete continuous scale from -15 (most healthy) to +40 (less healthy) (**Supplemental file 1**). FSA-NPS score allocates points (0-10) for the amount of energy (kJ), total sugar (g), SFA (g) and sodium (mg). Points (0-5) are subtracted from the previous sum based on the amount of fruits and vegetables (%, including legumes and nuts), fibers (g) and proteins (g). Specific modifications of the score for particular food groups were made to maintain a high consistency with French nutritional recommendations, as proposed by the French High Council for Public Health (HCSP) [34].
In a second step, the FSA-NPS DI was computed at the individual level using arithmetic energy-weighted means with the following equation [17], in which FS_i represents the food (or beverage) score, and E_i represents the energy intake from this food or beverage (all 24h-dietary records from the first two years of follow-up were averaged to a mean 24-hour energy intake from this food/beverage):

$$\text{FSA} - \text{NPS DI} = \frac{\sum_{i=1}^{n} FS_i E_i}{\sum_{i=1}^{n} E_i}$$

Increasing FSA-NPS DI reflects decreasing nutritional quality of foods consumed.

Case ascertainment

Participants self-declared health events through the yearly health status questionnaire, through a specific check-up questionnaire for health events (every three months) or at any time through a dedicated interface on the study website. Following this declaration, participants are invited to send their medical records (diagnosis, hospitalization, etc.) and, if necessary, the study physicians contact the participants' treating physician or the medical structures to collect additional information. Then, data are reviewed by an independent physician expert committee which validates all major health events (such as cancers). Cancer cases were classified using the International Chronic Diseases Classification, 10th Revision, Clinical Modification (ICD-10) [35]. In this study, all first primary breast cancers diagnosed between the inclusion and August 2015 were considered as cases. Information on death and cause of

For peer review only

death was obtained through linkage to the national database on mortality of the French population [36].

Statistical analyses

So far, 77,034 women without cancer at baseline provided at least three valid 24h-dietary records during their first two years of follow-up. Women aged <35y at baseline (n=29,249) were excluded due to a very low susceptibility to develop breast cancer in these women [37] and a potentially limited influence of nutrition on breast cancers diagnosed in young women. Women with a null follow-up were also excluded from the analyses (i.e. women for whom baseline questionnaires were the last completed questionnaires, n=921), thus leaving 46,864 women included in the analyses (flowchart in **Supplementary file 2**).

For each woman, the FSA-NPS DI and usual dietary intakes were calculated using all 24h-dietary records available in their first two years of follow-up. Associations between the FSA-NPS DI (continuous variable and quintiles) and breast cancer risk were characterized (HR and 95%CI) using multivariable Cox proportional hazards models with age as the primary time variable. We confirmed that the assumptions of proportionality were satisfied through examination of the log–log (survival) vs. log–time plots. Tests for linear trends were performed with the ordinal score on quintiles of FSA-NPS DI. Women contributed person-time to the model until the date of cancer diagnosis, the date of last completed questionnaire, the date of death or August 2015, whichever occurred first. Women who reported a cancer other than breast cancer during the study period were included and censored at the date of diagnosis (except basal cell skin carcinoma, not considered as cancer).

Models were adjusted for classic risk factors for breast cancer: age (time-scale), BMI (kg/m², continuous), height (cm, continuous), physical activity (high, moderate, low, computed following IPAQ recommendations [38]), smoking status (never smokers, former smokers, occasional smokers, smokers), number of dietary records (continuous), alcohol intake (g/d, continuous), energy intake (without alcohol, g/d, continuous), family history of cancer (yes/no), educational level (<high-school degree, <2 years after high-school degree, ≥2 years after high-school degree), number of biological children (continuous), menopausal status at baseline (pre-menopause, perimenopause, post-menopause), hormonal treatment for menopause (postmenopausal women, yes/no) and oral contraception use (premenopausal women, yes/no).

Interaction analysis was conducted between BMI and the FSA-NPS DI and stratified analyses were performed by overweight status (BMI < vs. ≥25kg/m^2).

For peer review only

Sensitivity analyses were performed including only women that provided at least six 24h-dietary records during their first two years of follow-up or excluding cases diagnosed during their first year of follow-up. Analyses were also performed on invasive breast cancer cases only and by hormonal receptor status of the tumors. Analyses were also performed by menopausal status. Women contributed person-time to the "pre-menopause model" until their age of menopause and to the "post-menopause model" from their age of menopause. Age at menopause was determined using the yearly health status questionnaires available during the follow-up.

For all covariates except physical activity, ≤ 5% of values were missing and were imputed to the modal value. For physical activity (N=6,328 missing values), a "missing class" was introduced into the models.

All tests were two-sided, and $P<0.05$ was considered statistically significant. SAS version 9.4 (SAS Institute) was used for the analyses.

For peer review only

RESULTS

Between May 2009 and August 2015 (median follow-up time: 4.0y; 174,491 person-years), 555 incident breast cancer cases were diagnosed: 171 premenopausal and 384 postmenopausal; 71.4% ER+/PR+, 14.7% ER-/PR-, 13.6% ER+/PR-, 0.3% ER-/PR+ (data available for 361 cases); 83.6% invasive and 16.4% *in situ* (data available for 463 cases). Mean age at diagnosis was 56.6y (SD=9.2) and mean baseline-to-diagnosis time was 2.4y (SD=1.6). Mean number of dietary records per participant over their first two years of follow-up was 5.9 (SD=2.8).

In **Table 1,** the characteristics of participants at baseline are described overall and by quintiles of the FSA-NPS DI. Mean FSA-NPS DI was 5.9±2.2 (min=-5.8; max=18.1). Women with a higher FSA-NPS DI (diet of lower nutritional quality), were more likely to be young, to smoke, to have a higher educational level and to have higher energy or alcohol intakes. As expected, women in the lowest quintiles of FSA-NPS DI (diet of higher nutritional quality) had overall healthier food intakes: higher intakes of fiber, fruits, vegetables, legume, fish and lower intakes of red and processed meat and lipids.

Compared to women that provided at least three 24h-dietary records over their first two years of follow-up, women that did not (15,918 women with a non-null follow-up) were younger, pre-menopause, were more likely to be overweight/obese, to smoke, to practice physical activity and were less likely to have a family history of cancer or to take a hormonal treatment for menopause [data not tabulated].

Associations between the FSA-NPS DI and breast cancer risk overall and by menopausal status are shown in **Table 2**. A direct association was observed between the FSA-NPS DI and breast cancer risk: $HR_{Q5vs.Q1}$=1.52 (95%CI 1.11-2.08), P-trend=0.002; $HR_{per\ 1\text{-}unit\ increment}$=1.06 (1.02-1.11), P=0.005. These associations were similarly observed in premenopausal women ($HR_{Q5vs.Q1}$=2.46 (1.27-4.75), P-trend=0.004; $HR_{per\ 1\text{-}unit\ increment}$=1.09 (1.01-1.18), P=0.03) and in postmenopausal women ($HR_{Q5vs.Q1}$=1.25 (0.85-1.84), P-trend=0.09; $HR_{per\ 1\text{-}unit\ increment}$=1.05 (1.00-1.11), P=0.06), although the associations seemed stronger for premenopausal women and only trends were observed for postmenopausal women (P-interaction=0.06).

Analyses performed by overweight status showed that associations tended to be stronger in non-overweight women (368 cases/ 31,401 non-cases, $HR_{Q5vs.Q1}$=1.97 (95%CI 1.31-2.96), P-trend=0.0007; $HR_{per\ 1\text{-}unit\ increment}$=1.09 (1.03-1.15), P=0.003) compared to overweight/obese women (187 cases/14,908 non-cases, $HR_{Q5vs.Q1}$=1.02 (95%CI 0.61-1.73), P-trend=0.6; $HR_{per\ 1\text{-}}$

For peer review only

$_{\text{unit increment}}$=1.03 (0.95-1.11), P=0.5), but the interaction was not statistically significant (P=0.07).

Information regarding hormone receptor status was not available for all cases (ER status: 361 cases, PR status: 362 cases, ER/PR status: 361 cases). Significant direct associations between the FSA-NPS DI and breast cancer risk were observed for breast cancer types PR- (102 cases/ 46,762 non-cases) and ER+/PR- (49 cases/ 46,815 non-cases). For ER+ tumours, the linear trend was not statistically significant (P=0.07, 307 cases/46,557 non-cases) but compared to women in the lowest quintile of FSA-NPS DI, those with higher scores had an increased breast cancer risk (e.g. $HR_{Q5vs.Q1}$=1.60 (1.04-1.46)). Associations were non-significant for the other hormone receptor status (**Supplementary file 3**). However, these exploratory findings should be considered with caution due to limited statistical power for analyses by cancer sub-types.

Similar results were observed when analyses excluded cases diagnosed during their first year of follow-up (425 cases/46,309 non-cases included; $HR_{Q5vs.Q1}$=1.54 (1.08-2.19), P-trend=0.007; $HR_{\text{per 1-unit increment}}$=1.07 (1.02-1.12), P=0.01) or when analyses were restricted to invasive breast cancers (387 cases/46,309 non-cases; $HR_{Q5vs.Q1}$=1.51 (1.03-2.22), P-trend=0.01; $HR_{\text{per 1-unit increment}}$=1.06 (1.01-1.12), P=0.03).

Results were also similar when analyses were restricted to women that provided at least 6 24h-dietary records during their first two years of follow-up (399 cases/25,439 non-cases; $HR_{Q5vs.Q1}$=1.63 (1.11-2.38), P-trend=0.006; $HR_{\text{per 1-unit increment}}$=1.08 (1.02-1.14), P=0.01) [data not tabulated].

Finally, similar but weaker trends were observed when women aged <35y at baseline were included in the analyses (585 cases/ 74,617 non-cases, $HR_{Q5vs.Q1}$=1.17 (95%CI 0.83-1.64), P-trend=0.1; $HR_{\text{per 1-unit increment}}$=1.05 (1.01-1.10), P=0.02).

For peer review only

DISCUSSION

In this prospective study conducted in a large sample of women from the French general population, a higher FSA-NPS DI, which reflects a diet composed of food products of lower nutritional quality, was associated with an increased risk of breast cancer.

In a previous study performed in the SU.VI.MAX cohort [22], we observed a direct association between the FSA-NPS DI and cancer risk overall but did not detect a significant association for breast cancer risk, probably due to limited power in site-specific analyses (n=125 breast cancer cases, 13y-follow-up). To our knowledge, no other study investigated the relationship between breast cancer risk and a score that characterizes the nutritional quality of an individual's diet based on a nutrient profiling system at the level of foods/beverages consumed.

However a few studies have been conducted on the association between NPS-based dietary scores and other health outcomes. While in this study, we used the FSA-NPS as a continuous score at the food/beverage level as a basis for the construction of the FSA-NPS DI at the individual level, the FSA-NPS was also recently used to define a variety score of "healthier" and "less healthy" foods/beverages (Ofcom binary cut-off used for advertising regulation in the UK [12]). This binary indicator was then studied in relation to mortality in the Whitehall II cohort [39]. The authors observed that a greater variety of healthier foods, as defined with the FSA-NPS Ofcom binary cut-off, was associated with a reduced all-cause and cancer mortality while a greater variety of less healthy food was not associated with the studied outcomes. No association was observed when another nutrient profiling system, the SAIN, LIM [40;41], was used [39].

To our knowledge, the Overall Nutritional Quality Index (ONQI-f) is the only other dietary score based on a nutrient profiling system that has been studied in relation to health outcomes [42]. It was tested in association with chronic diseases and mortality within the Nurses' Health Study and the Health Professionals Follow-up Study [42]. A higher ONQI-f, reflecting a higher nutritional quality of the diet, was associated with a decreased risk of cardiovascular diseases, diabetes and mortality but not was not associated with cancer. Some arguments may explain this lack of association: 1) the ONQI-f is based on 30 nutrients among which few have shown a consistent association with cancer risk, thus, its relevance regarding the cancer outcome may be lower than for other outcomes; 2) dietary intakes were assessed with an aggregated food frequency questionnaire (135-138 items), which provides less precise estimates than 24h-dietary records (as used in our study).

For peer review only

These studies are, to our knowledge, the only ones that investigated the associations between health outcomes and individual dietary indexes derived from nutrient profiling systems at the food level. Other a priori scores have been designed based on the intake of specific food groups or nutrients and/or other information (e.g. body fatness, physical activity), but not based on a nutrient profiling system at the food/beverage level. These scores were studied prospectively in relation to breast cancer risk and provided relatively contrasted results: 1) scores measuring the adherence to a specific type of diet such as the Mediterranean diet score (no association in prospective cohorts, inverse association in case-control studies [43-45]) or the Healthy Nordic Food Index (HNFI, no association [46]), 2) scores reflecting the adherence to general nutritional recommendations for the population such as the World Health Organization Healthy Diet Index, WHO HDI [47], the Alternate Healthy Eating Index, AHEI [45;48], the Recommended Food Score, RFS [48], the Diet Quality Index revised, DQI-R, [48] or the Dietary Approaches to Stop Hypertension (DASH) [45] (no association overall), and 3) scores measuring the adherence to cancer-specific nutritional recommendations such as the WCRF/AICR adherence score (inverse associations [49;50]) or the American Cancer Society (ACS) cancer prevention guidelines score (inverse association [51]). In these studies, differences according to hormonal receptor status of the tumors have been suggested, with inconsistent results. Indeed, inverse associations between a "healthier" diet and breast cancer risk were particularly observed in ER- type (AHEI, RFS, aMed) [48], ER-/PR+ type (Mediterranean diet score) [43], and ER-/PR-/HER2+ type (DASH) [45], but also with ER+/PR+ type (WCRF/AICR adherence score) [49] and ER+/PR- type (''healthy/Mediterranean'' pattern) [52]. In our study, information regarding hormonal receptor status of the tumors was only partially available and the statistical power was limited in the analyses (**Supplementary file 3**), thus preventing to derive firm conclusions.

Overall, these studies involving a priori scores provided interesting insights into the relationships between nutrition and breast cancer risk. Although these a priori scores and the FSA-NPS DI included similar nutritional components, the approaches differed, making the comparison between our study and previous findings not straightforward (even though our results were in line with those obtained with scores measuring the adherence to cancer-specific nutritional recommendations [49-51]). The FSA-NPS DI is not primarily built at the individual level but is rather derived from a nutrient profiling system at the food level (FSA-NPS) thus taking into account the nutritional quality of each food/beverage consumed and not only of the overall diet or overall consumption of food groups. In addition, the objective

For peer review only

behind the FSA-NPS DI construction was not to obtain the best predictive score for breast cancer but to specifically test its association with breast cancer risk, as the FSA-NPS is envisioned to serve as a basis for food labelling in the framework of public health policies in several countries such as France and Australia. The FSA-NPS displays several key advantages in a public health context: 1) it grades the nutritional quality of each food/beverage and thus reflects the variation of nutritional quality between but also within food groups, 2) it has been designed in a perspective of prevention of a large range of chronic diseases (not only breast cancer), and 3) it is easy-to-compute for industrials and public health stakeholders.

Our results are consistent with current evidence from epidemiological and mechanistic studies regarding the association between nutrition and breast cancer. Most of the input variables for the FSA-NPS are indeed parameters for which associations with breast cancer have been established either directly (e.g. dietary fibers [53]) or indirectly, through an association with body fatness, a major risk factor for postmenopausal breast cancer [53-55] (e.g. energy content, total sugars and SFA as components of energy-dense foods; fruits and vegetables as components of low-energy foods).

In our study, although similar trends were observed in pre- and post-menopausal women for the association between the FSA-NPS DI and breast cancer risk, this association was nonetheless stronger in pre-menopausal women. This may be explained by the fact that women pre-menopause were more likely to score high on the FSA-NPS DI, thus resulting in a clearer/stronger association: mean±SD FSA-NPS DI was 6.3±2.3 in women pre-menopause (median:6.4, 25th-75th percentiles: 4.9-7.8) and 5.5±2.1 in women post-menopause (median:5.5, 25th-75th percentiles: 4.1-6.9).

Strengths of this study pertained to its prospective design, its large sample size, and the assessment of usual dietary intakes using repeated 24h-dietary records based on a recent food composition database with a large choice of items (>3300). The latter allowed a better insight into the food products consumed and their intrinsic nutritional quality compared to studies that used a food frequency questionnaire (more aggregated food items). However, some limitations should be acknowledged. First, caution is needed regarding the extrapolation of these results to the entire French population since this study included volunteers involved in a long-term cohort study investigating the association between nutrition and health, with overall more health-conscious behaviors and higher professional and/or educational level compared to the general population. Thus, unhealthy dietary behaviors may have been underrepresented in this study, which may have weakened the observed associations. Next, information

For peer review only

regarding cancer stage was not available. Finally, as usually done in nutritional epidemiology, dietary intakes were estimated based on averaged intakes from all 24h-dietary records collected over the first two years of follow-up. Although diet may change over time, it is usually hypothesized that this estimation reflects general eating behavior throughout the adult life [56]. This very classical method allowed us to obtain a reliable estimation of usual dietary intakes, while respecting the prospective design (i.e. estimation of usual dietary intakes prior to cancer diagnosis). Indeed, breast cancer is a disease with relatively long latency so that the involvement of nutritional factors is supposed to be based on long-term processes. Thus, it is important to guarantee sufficient delay between nutritional exposure and cancer outcome. This is why we tested a model (sensitivity analysis) where cancer cases diagnosed during the first year of follow-up were excluded (similar results). In our study, although the follow-up time was appropriate to perform etiological analyses, it did not necessarily guarantee this sufficient delay. Hence, our estimation of usual dietary intakes may reflect dietary protective and risk factors that may have played a role in the first steps of carcinogenesis (initiation) but also later in the carcinogenic process (progression). Nonetheless, previous studies with longer follow-up observed associations between diet and breast cancer risk, suggesting that nutritional factors could play a role in cancer initiation and not only in cancer progression [45;48-52].

In conclusion, these results suggest that the consumption of food products of lower nutritional quality (higher FSA-NPS) may be associated with an increased risk of breast cancer. Women in the highest FSA-NPS DI quintile had a 52% increase in breast cancer risk compared to women with the lowest scores (first quintile). The ability of the FSA-NPS DI to predict disease risk (here breast cancer risk) suggests that the FSA-NPS is a valid system to characterize the nutritional quality of foodstuffs and to highlight products with a good nutritional profile that should be promoted and products with a lower nutritional quality that should not. Therefore, this study adds to the scientific evidence that supports the public health relevance of the implementation of front-of-pack nutrition labels based on the FSA-NPS (e.g. 5-CNL) in order to help consumers make healthier food choices.

For peer review only

ACKNOWLEDGEMENTS

The authors thank all the volunteers of the NutriNet-Santé cohort. We extend special thanks to Véronique Gourlet for the statistical analyses and to Nathalie Arnault, Stephen Besseau, Laurent Bourhis, Yasmina Chelghoum, Than Duong Van, Younes Esseddik, Paul Flanzy, Charlie Ménard, Mac Rakotondrazafy, Fabien Szabo, Roland Andrianasolo, Fatoumata Diallo, Cédric Agaesse, Claudia Chahine, Marion Genest and Ludivine Ursule for their technical contribution to the NutriNet-Santé study.

Authors' contribution: The authors' contributions were as follow – MD and MT: designed the research; SH, MT, CJ, EKG: conducted the research; MD and MT: supervised statistical analysis; MD and MT: wrote the paper; CJ, EKG, LL, SA, CM, PD, SP, PLM, LF, PF, SH: contributed to the data interpretation and revised each draft for important intellectual content. All authors read and approved the final manuscript. MD and MT had primary responsibility for the final content.

Funding: Mélanie Deschasaux, Philippine Fassier and Solia Adriouch were funded by PhD grants from the Région Ile-de-France (public funding: Cancéropôle Ile-de-France and CORDDIM). The NutriNet-Santé study was supported by the following public institutions: Ministère de la Santé, Institut de Veille Sanitaire (InVS), Institut National de la Prévention et de l'Education pour la Santé (INPES), Région Ile-de-France (CORDDIM), Institut National de la Santé et de la Recherche Médicale (INSERM), Institut National de la Recherche Agronomique (INRA), Conservatoire National des Arts et Métiers (CNAM) and Université Paris 13. The funders had no role in the design, implementation, analysis, or interpretation of the data.

Competing interests: The authors have no conflict of interest to disclose.

Data sharing: All relevant data are in the manuscript and its supporting files. No additional data available.

For peer review only

REFERENCES

(1) WHO/IARC. All Cancers: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012. GLOBOCAN 2012 [2016 Available from: URL:http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx

(2) WCRF/AICR. Cancer preventability estimates for food, nutrition, body fatness, and physical activity. WCRF/AICR [2016 Available from: URL:http://www.wcrf.org/int/cancer-facts-figures/preventability-estimates/cancer-preventability-estimates-diet-nutrition

(3) Food Standards Agency, Welsh Government, Scottish Government, Department of Health. Front of Package Nutrition Labelling: Joint response to consultation. 2013.

(4) Gill T, King L, Vita P, Caterson I, Colagiuri S, Colagiuri R et al. A 'state of the knowledge' assessment of comprehensive interventions that address the drivers of obesity. A Rapid Assessment Prepared for the National Health and Medical Research Council (NHMRC). 2010. Sydney, The Boden Institute of Obesity, Nutrition, Exercise and Eating Disorders, University of Sydney.

(5) Institute of Medecine. Front-of-Package Nutrition Rating Systems and Symbols: Promoting Healthier Choices. 2012. Washington, D.C., The National Academies Press.

(6) WHO. 2008-2013 action plan for the global strategy for the prevention and control of non-communicable diseases. 2009. Geneva, World Health Organization.

(7) WHO Europe. Action Plan for implementation of the European Strategy for the Prevention and Control of Noncommunicable Diseases 2012–2016. 2011. Copenhagen, World Health Organization Regional Office for Europe.

(8) Arambepola C, Scarborough P, Rayner M. Validating a nutrient profile model. *Public Health Nutr* 2008; 11(4):371-378.

(9) Rayner M, Scarborough P, Stockley P, Boxer A. Nutrient profiles: Development of Final Model. Final Report [online]. London, FSA . 2005.

(10) Hercberg S. Propositions pour un nouvel élan de la politique nutritionnelle française de santé publique dans le cadre de la stratégie nationale de santé. 1ère partie: mesures concernant la prévention nutritionnelle [in French]. 2013.

(11) Rayner M, Scarborough P, Stockley L. Nutrient profiles: Applicability of currently proposed model for uses in relation to promotion of food to children aged 5-10 and adults. [online]. London, Foods Standard Agency. 2005.

(12) Rayner M, Scarborough P, Lobstein T. The UK Ofcom Nutrient Profiling Model: Defining 'healthy' and 'unhealthy' foods and drinks for TV advertising to children. London, OfCom . 2009.

(13) Julia C, Kesse-Guyot E, Touvier M, Mejean C, Fezeu L, Hercberg S. Application of the British Food Standards Agency nutrient profiling system in a French food composition database. *Br J Nutr* 2014; 112(10):1699-1705.

For peer review only

(14) Julia C, Kesse-Guyot E, Ducrot P, Peneau S, Touvier M, Mejean C et al. Performance of a five category front-of-pack labelling system - the 5-colour nutrition label - to differentiate nutritional quality of breakfast cereals in France. *BMC Public Health* 2015; 15:179.

(15) Julia C, Ducrot P, Peneau S, Deschamps V, Mejean C, Fezeu L et al. Discriminating nutritional quality of foods using the 5-Color nutrition label in the French food market: consistency with nutritional recommendations. *Nutr J* 2015; 14:100.

(16) Haut Conseil de la Santé Publique. Avis relatif à l'information sur la qualité nutritionnelle des produits alimentaires. 2015. Paris, HCSP.

(17) Julia C, Touvier M, Mejean C, Ducrot P, Peneau S, Hercberg S et al. Development and validation of an individual dietary index based on the British Food Standard Agency nutrient profiling system in a French context. *J Nutr* 2014; 144(12):2009-2017.

(18) Julia C, Mejean C, Touvier M, Peneau S, Lassale C, Ducrot P et al. Validation of the FSA nutrient profiling system dietary index in French adults-findings from SUVIMAX study. *Eur J Nutr* 2015.

(19) Julia C, Ducrot P, Lassale C, Fezeu L, Mejean C, Peneau S et al. Prospective associations between a dietary index based on the British Food Standard Agency nutrient profiling system and 13-year weight gain in the SU.VI.MAX cohort. *Prev Med* 2015; 81:189-194.

(20) Julia C, Fezeu LK, Ducrot P, Mejean C, Peneau S, Touvier M et al. The Nutrient Profile of Foods Consumed Using the British Food Standards Agency Nutrient Profiling System Is Associated with Metabolic Syndrome in the SU.VI.MAX Cohort. *J Nutr* 2015; 145(10):2355-2361.

(21) Adriouch S, Julia C, Kesse-Guyot E, Mejean C, Ducrot P, Peneau S et al. Prospective association between a dietary quality index based on a nutrient profiling system and cardiovascular disease risk. *Eur J Prev Cardiol* 2016.

(22) Donnenfeld M, Julia C, Kesse-Guyot E, Mejean C, Ducrot P, Peneau S et al. Prospective association between cancer risk and an individual dietary index based on the British Food Standards Agency Nutrient Profiling System. *Br J Nutr* 2015; 114(10):1702-1710.

(23) Hercberg S, Castetbon K, Czernichow S, Malon A, Mejean C, Kesse E 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* 2010; 10:242.

(24) Vergnaud AC, Touvier M, Mejean C, Kesse-Guyot E, Pollet C, Malon A et al. Agreement between web-based and paper versions of a socio-demographic questionnaire in the NutriNet-Sante study. *Int J Public Health* 2011; 56(4):407-417.

(25) Lassale C, Peneau S, Touvier M, Julia C, Galan P, Hercberg S et al. Validity of web-based self-reported weight and height: results of the Nutrinet-Sante study. *J Med Internet Res* 2013; 15(8):e152.

For peer review only

(26) Touvier M, Mejean C, Kesse-Guyot E, Pollet C, Malon A, Castetbon K et al. Comparison between web-based and paper versions of a self-administered anthropometric questionnaire. *Eur J Epidemiol* 2010; 25(5):287-296.

(27) Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003; 35(8):1381-1395.

(28) Lassale C, Castetbon K, Laporte F, Deschamps V, Vernay M, Camilleri GM et al. Correlations between Fruit, Vegetables, Fish, Vitamins, and Fatty Acids Estimated by Web-Based Nonconsecutive Dietary Records and Respective Biomarkers of Nutritional Status. *J Acad Nutr Diet* 2016; 116(3):427-438.

(29) Lassale C, Castetbon K, Laporte F, Camilleri GM, Deschamps V, Vernay M et al. Validation of a Web-based, self-administered, non-consecutive-day dietary record tool against urinary biomarkers. *Br J Nutr* 2015; 113(6):953-962.

(30) Touvier M, Kesse-Guyot E, Mejean C, Pollet C, Malon A, Castetbon K et al. Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. *Br J Nutr* 2011; 105(7):1055-1064.

(31) Le Moullec N, Deheeger M, Preziosi P, Montero P, Valeix P, Rolland-Cachera M et al. Validation du manuel photo utilisé pour l'enquête alimentaire de l'étude SU.VI.MAX. [Validation of the food portion size booklet used in the SU.VI.MAX study] (in French). *Cah Nutr Diet* 1996; 31:158-164.

(32) Arnault N, Caillot L, Castetbon K, Coronel S, Deschamps V, Fezeu L et al. Table de composition des aliments, étude NutriNet-Santé. [Food composition table, NutriNet-Santé study] (in French). Paris: Les éditions INSERM/Economica; 2013.

(33) Black AE. Critical evaluation of energy intake using the Goldberg cut-off for energy intake:basal metabolic rate. A practical guide to its calculation, use and limitations. *Int J Obes Relat Metab Disord* 2000; 24(9):1119-1130.

(34) Haut Conseil de la Santé Publique. Opinion on information regarding the nutritional quality of foodstuffs. 2015. Paris, HCSP.

(35) WHO. ICD-10, International classification of diseases and related health problems. 10th revision. 2010. Geneva, Switzerland: World Health Organization.

(36) Inserm. CépiDC. Interrogation des données sur les causes de décès de 1979 à 2013. Inserm [2016 Available from: URL:http://www.cepidc.inserm.fr/inserm/html/index2.htm

(37) INCa. Incidence et mortalité estimées par classe d'âge et par localisation cancéreuse en 2012. 2016.

(38) IPAQ Group. Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ). 2005.

For peer review only

(39) Masset G, Scarborough P, Rayner M, Mishra G, Brunner EJ. Can nutrient profiling help to identify foods which diet variety should be encouraged? Results from the Whitehall II cohort. *Br J Nutr* 2015; 113(11):1800-1809.

(40) Afssa. Définition de profils nutritionnels pour l'accès aux allégations nutritionnelles et de santé: propositions et arguments [The setting of nutrient profiles for access to nutrition an health claims: proposals and arguments] (in French). 2008. Paris, Agence française de sécurité des aliments.

(41) Darmon N, Vieux F, Maillot M, Volatier JL, Martin A. Nutrient profiles discriminate between foods according to their contribution to nutritionally adequate diets: a validation study using linear programming and the SAIN,LIM system. *Am J Clin Nutr* 2009; 89(4):1227-1236.

(42) Chiuve SE, Sampson L, Willett WC. The association between a nutritional quality index and risk of chronic disease. *Am J Prev Med* 2011; 40(5):505-513.

(43) Schwingshackl L, Hoffmann G. Adherence to Mediterranean diet and risk of cancer: an updated systematic review and meta-analysis of observational studies. *Cancer Med* 2015; 4(12):1933-1947.

(44) Schwingshackl L, Hoffmann G. Does a Mediterranean-Type Diet Reduce Cancer Risk? *Curr Nutr Rep* 2016; 5:9-17.

(45) Hirko KA, Willett WC, Hankinson SE, Rosner BA, Beck AH, Tamimi RM et al. Healthy dietary patterns and risk of breast cancer by molecular subtype. *Breast Cancer Res Treat* 2016; 155(3):579-588.

(46) Li Y, Roswall N, Sandin S, Strom P, Adami HO, Weiderpass E. Adherence to a healthy Nordic food index and breast cancer risk: results from a Swedish cohort study. *Cancer Causes Control* 2015; 26(6):893-902.

(47) Cade JE, Taylor EF, Burley VJ, Greenwood DC. Does the Mediterranean dietary pattern or the Healthy Diet Index influence the risk of breast cancer in a large British cohort of women? *Eur J Clin Nutr* 2011; 65(8):920-928.

(48) Fung TT, Hu FB, McCullough ML, Newby PK, Willett WC, Holmes MD. Diet quality is associated with the risk of estrogen receptor-negative breast cancer in postmenopausal women. *J Nutr* 2006; 136(2):466-472.

(49) Harris HR, Bergkvist L, Wolk A. Adherence to the World Cancer Research Fund/American Institute for Cancer Research recommendations and breast cancer risk. *Int J Cancer* 2016; 138(11):2657-2664.

(50) Nomura SJ, Inoue-Choi M, Lazovich D, Robien K. WCRF/AICR recommendation adherence and breast cancer incidence among postmenopausal women with and without non-modifiable risk factors. *Int J Cancer* 2016; 138(11):2602-2615.

(51) Thomson CA, McCullough ML, Wertheim BC, Chlebowski RT, Martinez ME, Stefanick ML et al. Nutrition and physical activity cancer prevention guidelines, cancer risk, and mortality in the women's health initiative. *Cancer Prev Res (Phila)* 2014; 7(1):42-53.

For peer review only

(52) Cottet V, Touvier M, Fournier A, Touillaud MS, Lafay L, Clavel-Chapelon F et al. Postmenopausal breast cancer risk and dietary patterns in the E3N-EPIC prospective cohort study. *Am J Epidemiol* 2009; 170(10):1257-1267.

(53) Latino-Martel P, Cottet V, Druesne-Pecollo N, Pierre FH, Touillaud M, Touvier M et al. Alcoholic beverages, obesity, physical activity and other nutritional factors, and cancer risk: a review of the evidence. Crit Rev Oncol Hematol 99[308], 323. 2016.

(54) WCRF/AICR. Food, Nutrition, Physical Activity and the Prevention of Cancer: a Global Perspective. 2007. Washington, DC: AICR.

(55) WCRF/AICR. Systematic Literature Review - Continuous Update Project Report: The Associations between Food, Nutrition and Physical Activity and the Risk of Breast Cancer. 2010. Washington, DC: AICR.

(56) Willett WC. Nutritional Epidemiology, 2nd ed. New York: Oxford University Press; 1998.

For peer review only

Table 1 Baseline characteristics of the study population overall and according to quintiles of the FSA-NPS DI, NutriNet-Santé Cohort, France, 2009-2015

		Quintiles of the FSA-NPS DI					
	All women (n=46,864)	Q1 (n=9,349)	Q2 (n=9,395)	Q3 (n=9,387)	Q4 (n=9,415)	Q5 (n=9,318)	
	N (%) Mean±SD	N (%) Mean±SD	N (%) Mean±SD	N (%) Mean±SD	N (%) Mean±SD	N (%) Mean±SD	P-trend[a]
FSA-NPS DI	5.9±2.2	2.7±1.2	4.8±0.4	6.0±0.3	7.1±0.4	9.0±1.1	<.0001
Age, years	50.8±9.7	53.4±9.6	52.6±9.4	51.2±9.7	49.6±9.4	47.1±8.9	<.0001
Educational level							<.0001
< high-school degree	11269 (24.1)	2658 (28.4)	2345 (25.0)	2172 (23.1)	2083 (22.1)	2011 (21.6)	
≥high-school degree to < 2y after high-school degree	7834 (16.7)	1567 (16.8)	1579 (16.8)	1570 (16.7)	1556 (16.5)	1562 (16.8)	
≥ 2y after high-school degree	27761 (59.2)	5124 (54.8)	5471 (58.2)	5645 (60.1)	5776 (61.3)	5745 (61.6)	
Smoking status							<.0001
Non-smokers	22528 (48.1)	4630 (49.5)	4706 (50.1)	4615 (49.2)	4504 (47.8)	4073 (43.7)	
Former smokers	17904 (38.2)	3744 (40.0)	3640 (38.7)	3561 (37.9)	3527 (37.5)	3432 (36.8)	
Occasional smokers[b]	1622 (3.5)	257 (2.7)	302 (3.2)	336 (3.6)	350 (3.7)	377 (4.0)	
Smokers	4810 (10.3)	718 (7.7)	747 (7.9)	875 (9.3)	1034 (11.0)	1436 (15.4)	
Physical activity[c]							0.08
Low	13955 (34.4)	3312 (41.1)	2979 (36.4)	2800 (34.5)	2569 (31.5)	2295 (28.6)	
Moderate	17062 (42.1)	3224 (40.0)	3462 (42.4)	3487 (43.0)	3522 (43.2)	3367 (41.9)	
High	9519 (23.5)	1521 (18.9)	1732 (21.2)	1829 (22.5)	2066 (25.3)	2371 (29.5)	
BMI, kg/m²	24.1±4.8	24.5±4.9	24.1±4.7	23.9±4.5	23.9±4.6	24.3±5.2	<.0001
Weight status							<.0001
Normal-weight (BMI<25kg/m^2)	31,769 (67.8)	5929 (63.4)	6406 (68.2)	6558 (69.9)	6550 (69.6)	6326 (67.9)	
Overweight (25≤BMI<30kg/m^2)	9975 (21.3)	2270 (24.3)	2002 (21.3)	1971 (21.0)	1924 (20.4)	1808 (19.4)	
Obese (BMI≥30kg/m^2)	5120 (10.9)	1150 (12.3)	987 (10.5)	858 (9.1)	941 (10.0)	1184 (12.7)	
Height, cm	163.4±6.1	162.8±6.0	163.0±6.0	163.4±6.1	163.7±6.0	164.2±6.1	<.0001
Number of biological children	1.8±1.2	1.8±1.2	1.8±1.1	1.8±1.1	1.8±1.1	1.8±1.2	<.0001
Family history of cancer (yes)	21158 (45.2)	4446 (47.6)	4393 (46.8)	4288 (45.7)	4185 (44.4)	3846 (41.3)	0.9
Menopausal status							0.5
Pre-menopause	23940 (51.1)	3767 (40.3)	4078 (43.4)	4637 (49.4)	5296 (56.2)	6162 (66.1)	
Perimenopause	3997 (8.5)	807 (8.6)	871 (9.3)	807 (8.6)	795 (8.4)	717 (7.7)	
Post-menopause	18927 (40.4)	4775 (51.1)	4446 (47.3)	3943 (42.0)	3324 (35.3)	2439 (26.2)	
Hormonal treatment for menopause use (yes)[d]	4068 (17.7)	1025(18.4)	978 (18.4)	806 (17.0)	732 (17.8)	527 (16.7)	0.04
Energy intake without alcohol, kcal/d	1710±385	1510±331	1648±334	1721±344	1792±370	1882±429	<.0001
Alcohol intake, g/d	6.5±9.1	4.5±7.7	5.9±8.4	6.8±9.0	7.4±9.5	7.9±10.5	<.0001

For peer review only

Lipid intake, g/d	76.2±22.6	58.5±17.4	70.2±17.4	76.9±18.2	83.0±19.7	92.4±23.9	<.0001
Protein intake, g/d	76.0±18.3	78.1±20.6	76.3±17.7	75.8±17.1	75.6±17.4	74.4±18.4	<.0001
Carbohydrate intake, g/d	94.9±31.9	88.0±34.5	94.6±31.5	96.3±30.4	98.1±30.9	97.3±31.0	<.0001
Fiber intake, g/d	19.4±6.5	22.4±7.9	20.5±6.2	19.4±5.7	18.4±5.4	16.6±5.2	<.0001
Fruit intake, g/d	247.8±152.3	303.9±185.3	271.1±145.8	249.6±138.2	226.7±130.9	187.4±128.0	<.0001
Vegetable intake, g/d	236.6±113.3	295.8±138.5	255.6±105.7	234.9±98.7	215.2±92.1	181.4±91.0	<.0001
Legume intake, g/d	11.6±21.4	16.8±29.4	12.7±21.3	11.0±19.1	9.7±17.9	7.6±15.7	<.0001
Red meat intake, g/d	39.0±34.1	38.6±37.8	39.6±33.3	40.0±33.8	40.0±32.8	39.7±31.0	<.0001
Processed meat intake, g/d	28.4±25.7	19.4±21.9	23.6±21.8	27.3±22.7	32.1±24.8	37.0±32.6	<.0001
Poultry intake, g/d	24.8±27.6	31.3±34.6	26.0±27.6	24.1±25.3	22.6±24.1	20.1±23.5	<.0001
Fish (including sea product) intake, g/d	40.7±37.6	52.2±45.2	44.8±37.7	40.2±35.3	36.0±32.8	30.5±31.7	<.0001
Dairy intake, g/d	162.8±145.3	217.2±176.1	178.1±145.8	158.8±134.8	142.1±125.2	117.9±117.8	<.0001

[a] P value for the comparison between quintiles of FSA-NPS DI, by χ^2 tests from age-adjusted ordinal polytomous logistic regressions

[b] Occasional smokers smoke less than once a day

[c] Data available for 40,536 women

[d] Among women in peri- or post-menopause (n=22,924)

For peer review only

Table 2 Associations between the FSA-NPS DI and breast cancer risk, from multivariable Cox proportional hazards models, NutriNet-Santé Cohort, France, 2009-2015

		Age-adjusted model			Multivariable-adjusted model[a]		
FSA-NPS DI	N for cases/ non-cases	HR	95%CI	P-trend	HR	95%CI	P-trend
Overall							
Continuous score	555/46,309	1.07	1.03, 1.11	0.001	1.06	1.02, 1.11	0.005
Quintiles[b]				0.0004			0.002
Q1	82/9,267	1.00	(ref)		1.00	(ref)	
Q2	122/9,273	1.44	1.09, 1.90		1.43	1.08, 1.90	
Q3	117/9,270	1.45	1.09, 1.93		1.43	1.07, 1.91	
Q4	138/9,277	1.83	1.39, 2.40		1.79	1.35, 2.38	
Q5	96/9,222	1.56	1.15, 2.10		1.52	1.11, 2.08	
Premenopausal women[c]							
Continuous score	171/23,483	1.09	1.02, 1.18	0.02	1.09	1.01, 1.18	0.03
Quintiles[b]				0.002			0.004
Q1	12/3,667	1.00	(ref)		1.00	(ref)	
Q2	28/3,982	1.96	0.99, 3.85		1.92	0.97, 3.79	
Q3	31/4,558	1.94	0.99, 3.78		1.89	0.96, 3.71	
Q4	52/5,204	2.88	1.53, 5.39		2.76	1.45, 5.26	
Q5	48/6,072	2.52	1.34, 4.76		2.46	1.27, 4.75	
Postmenopausal women[c]							
Continuous score	384/27,188	1.06	1.01, 1.11	0.02	1.05	1.00, 1.11	0.06
Quintiles[b]				0.03			0.09
Q1	70/6,416	1.00	(ref)		1.00	(ref)	
Q2	94/6,173	1.35	0.99, 1.84		1.36	0.99, 1.86	
Q3	86/5,578	1.38	1.01, 1.90		1.37	0.99, 1.89	
Q4	86/5,028	1.60	1.17, 2.20		1.57	1.13, 2.18	
Q5	48/3,993	1.30	0.90, 1.88		1.25	0.85, 1.84	

[a] Models were adjusted for age (time-scale), BMI (kg/m², continuous), height (cm, continuous), physical activity (high, moderate, low), smoking status (never smokers, former smokers, occasional smokers, smokers), numbers of dietary records (continuous), alcohol intake (g/d, continuous), energy intake (without alcohol, g/d, continuous), family history of cancer (yes/no), educational level (<high-school degree, <2 years after high-school degree, ≥2 years after high-school degree), number of biological children (continuous), menopausal status at baseline (pre-menopause, perimenopause, post-menopause), hormonal treatment for menopause (postmenopausal women, yes/no) and oral contraception use (premenopausal women, yes/no).

[b] Cut-offs for quintiles of the FSA-NPS DI were 4.1/5.4/6.5/7.7

[c] P for interaction between the FSA-NPS DI and menopausal status=0.06

For peer review only

Supplemental file 1 FSA NPS score computation at food/beverage level

Points are allocated according to the nutrient content for 100g of foods or beverages.
Points are allocated for 'Negative' nutrients (A points) and can be balanced according to 'Positive' nutrients (C points).

A points
Total A points = (points for energy) + (points for saturated fat) + (points for total sugar) + (points for sodium)

Points	Energy (kJ)	Saturated Fat (g)	Total Sugars (g)	Sodium (mg)
0	≤ 335	≤ 1	≤ 4.5	≤ 90
1	> 335	> 1	> 4.5	> 90
2	> 670	> 2	> 9	> 180
3	> 1005	> 3	> 13.5	> 270
4	> 1340	> 4	> 18	> 360
5	> 1675	> 5	> 22.5	> 450
6	> 2010	> 6	> 27	> 540
7	> 2345	> 7	> 31	> 630
8	> 2680	> 8	> 36	> 720
9	> 3015	> 9	> 40	> 810
10	> 3350	> 10	> 45	> 900

C points
Total C points = (points for fruits and vegetables) + (points for fibers) + (points for proteins)

Points	Fruits, Vegetables (%)	Fiber (g) *	Protein (g)
0	≤ 40	≤ 0.7	≤ 1.6
1	> 40	> 0.7	> 1.6
2	> 60	> 1.4	> 3.2
3	-	> 2.1	> 4.8
4	-	> 2.8	> 6.4
5	> 80	> 3.5	> 8.0

*FSA score allocates different thresholds for fibers, depending on the measurement method used. We used NSP cut-offs to compute fibers score.

Overall score computation

- If Total A points <11, then FSA score =Total A points – Total C points
- If Total A points ≥11,
  - If points for fruits and vegetables =5, then FSA score =Total A points – Total C points
  - Else if points for fruits and vegetables <5, then FSA score = Total A points – (points for fiber + points for fruits and vegetables).

For 100g of a given food, the percentage of fruits and vegetables is obtained by summing up the amount (in grams) of all fruits, legumes and vegetables (including oleaginous fruits, dried fruits and olives) contained in this food.

For peer review only

Exceptions were made for cheese, fat, and drinks to better rank them according to their nutrient profile, consistently with nutritional recommendations:

Score computation for cheese

For cheese, the score takes in account the protein content, whether the A score reaches 11 or not, i.e.: FSA score =Total A points – Total C points

Score computation for fat

For fat, the grid for point attribution is based on the percentage of saturated fat among total lipids and has a six-point homogenous ascending step, as shown thereafter:

Points	**Saturated Fat/Lipids (%)**
0	< 10
1	< 16
2	< 22
3	< 28
4	< 34
5	< 40
6	< 46
7	< 52
8	< 58
9	< 64
10	≥ 64

Score computation for drinks

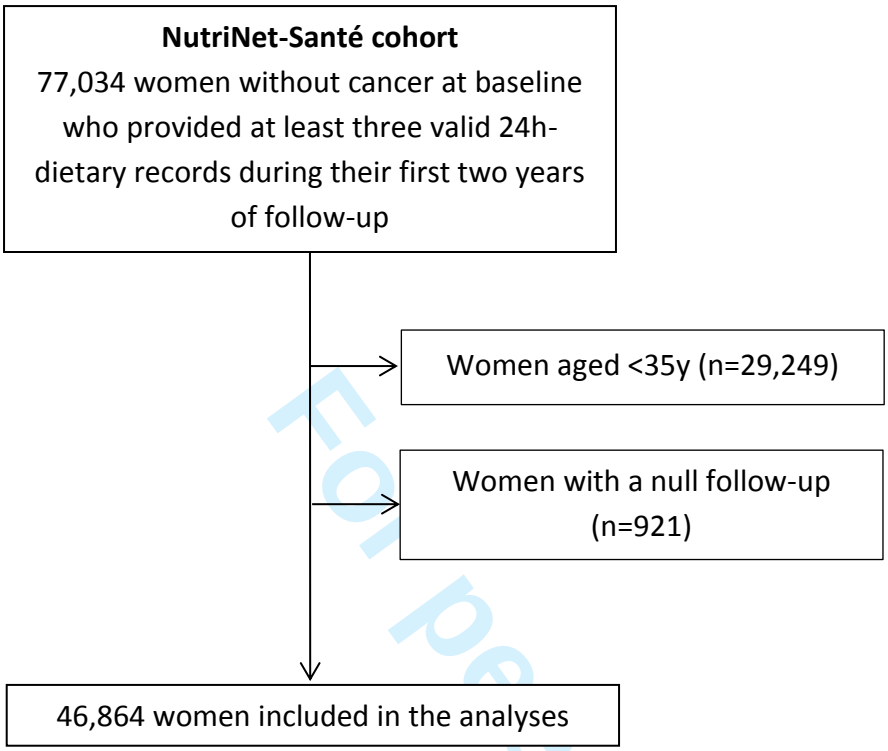
For drinks, the grids for point attribution regarding energy, total sugars and fruits and vegetables (%) were modified. The attribution of points for sugars takes into account the presence of sweeteners, in which case the grid maintains the total sugar score to 1 (instead of 0).

Points	**Energy (kJ)**	**Total Sugar (g)**	**Fruits, Vegetables (%)**
0	≤ 0	≤ 0	< 40
1	≤ 30	≤ 1.5	
2	≤ 60	≤ 3	> 40
3	≤ 90	≤ 4.5	
4	≤ 120	≤ 6	> 60
5	≤ 150	≤ 7.5	
6	≤ 180	≤ 9	
7	≤ 210	≤ 10.5	
8	≤ 240	≤ 12	
9	≤ 270	≤ 13.5	
10	> 270	> 13.5	> 80

Milk and vegetable milk are not concerned by this exception. Their scores are computed using the overall score computation system.

For peer review only

Supplementary file 2 Participants' flowchart



For peer review only

Supplementary file 3 Associations between the FSA-NPS DI and breast cancer risk by hormonal receptor status of the tumors, from multivariable Cox proportional hazards models, NutriNet-Santé Cohort, France, 2009-2015

		Multivariable-adjusted model[a]		
FSA-NPS DI	N for cases/ non-cases	HR	95%CI	P-trend
ER+	307/46,557			
Continuous score		1.05	0.99, 1.12	0.07
Quintiles[b]				0.07
Q1		1.00	(ref)	
Q2		1.59	1.08, 2.34	
Q3		1.77	1.20, 2.60	
Q4		1.59	1.06, 2.39	
Q5		1.60	1.04, 2.46	
ER-	54/46,810			
Continuous score		1.07	0.94, 1.23	0.3
Quintiles[b]				0.1
Q1		1.00	(ref)	
Q2		1.70	0.66, 4.37	
Q3		0.74	0.23, 2.37	
Q4		3.24	1.32, 7.95	
Q5		1.54	0.54, 4.42	
PR+	260/46,604			
Continuous score		1.04	0.97, 1.11	0.2
Quintiles[b]				0.3
Q1		1.00	(ref)	
Q2		1.73	1.14, 2.62	
Q3		1.64	1.07, 2.52	
Q4		1.62	1.04, 2.51	
Q5		1.46	0.91, 2.35	
PR-	102/46,762			
Continuous score		1.11	1.01, 1.23	0.04
Quintiles[b]				0.01
Q1		1.00	(ref)	
Q2		1.29	0.64, 2.62	
Q3		1.68	0.84, 3.34	
Q4		2.46	1.26, 4.79	
Q5		1.99	0.95, 4.17	
ER+/PR+				
Continuous score	258/46,606	1.04	0.97, 1.11	0.3
Quintiles[b]				0.3
Q1		1.00	(ref)	
Q2		1.72	1.14, 2.62	
Q3		1.58	1.03, 2.42	
Q4		1.61	1.04, 2.49	
Q5		1.45	0.90 ,2.33	
ER-/PR-	53/46,811			
Continuous score		1.07	0.93, 1.23	0.3
Quintiles[b]				0.1
Q1		1.00	(ref)	
Q2		1.68	0.65, 4.32	

For peer review only

Q3		0.58	0.17, 2.01	
Q4		3.16	1.29, 7.77	
Q5		1.50	0.52, 4.32	
ER+/PR-	49/46,815			
Continuous score		1.16	1.00, 1.35	0.047
Quintiles[b]				0.03
Q1		1.00	(ref)	
Q2		0.88	0.29, 2.66	
Q3		2.89	1.18, 7.11	
Q4		1.53	0.53, 4.38	
Q5		2.69	0.96, 7.57	

[a] Models were adjusted for age (time-scale), BMI (kg/m², continuous), height (cm, continuous), physical activity (high, moderate, low), smoking status (never smokers, former smokers, occasional smokers, smokers), numbers of dietary records (continuous), alcohol intake (g/d, continuous), energy intake (without alcohol, g/d, continuous), family history of cancer (yes/no), educational level (<high-school degree, <2 years after high-school degree, ≥2 years after high-school degree), number of biological children (continuous), menopausal status at baseline (pre-menopause, perimenopause, post-menopause), hormonal treatment for menopause (postmenopausal women, yes/no) and oral contraception use (premenopausal women, yes/no).
[b] Cut-offs for quintiles of the FSA-NPS DI were 4.1/5.4/6.5/7.7

For peer review only

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cohort studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1; 2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-7
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	6-8
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	n.a
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-8
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6-8
Bias	9	Describe any efforts to address potential sources of bias	7-9
Study size	10	Explain how the study size was arrived at	8
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8-9
		(*b*) Describe any methods used to examine subgroups and interactions	8-9
		(*c*) Explain how missing data were addressed	9
		(*d*) If applicable, explain how loss to follow-up was addressed	8
		(*e*) Describe any sensitivity analyses	9
Results			

For peer review only

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	8
		(b) Give reasons for non-participation at each stage	8
		(c) Consider use of a flow diagram	Supplementary file 2
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	10; Table 1
		(b) Indicate number of participants with missing data for each variable of interest	9
		(c) Summarise follow-up time (eg, average and total amount)	10
Outcome data	15*	Report numbers of outcome events or summary measures over time	10
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	10; Table 2
		(*b*) Report category boundaries when continuous variables were categorized	8; Table 2
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n.a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	10-11
Discussion			
Key results	18	Summarise key results with reference to study objectives	12;14
Limitations			
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	12-15
Generalisability	21	Discuss the generalisability (external validity) of the study results	12-15
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	16

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

For peer review only

BMJ Open

Are self-reported unhealthy food choices associated with an increased risk of breast cancer: prospective cohort study using the British Food Standards Agency Nutrient Profiling System

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-013718.R2
Article Type:	Research
Date Submitted by the Author:	08-Mar-2017
Complete List of Authors:	Deschasaux, Mélanie; Equipe de Recherche en Epidemiologie Nutritionnelle, Julia, Chantal; Equipe de Recherche en Epidemiologie Nutritionnelle; Hopital Avicenne, Public Health Department Kesse-Guyot, Emmanuelle; Equipe de Recherche en Epidemiologie Nutritionnelle Lécuyer, Lucie; Equipe de Recherche en Epidemiologie Nutritionnelle Adriouch, Solia; Equipe de Recherche en Epidemiologie Nutritionnelle Méjean, Caroline; Equipe de Recherche en Epidemiologie Nutritionnelle Ducrot, Pauline; Equipe de Recherche en Epidemiologie Nutritionnelle Péneau, Sandrine; Equipe de Recherche en Epidemiologie Nutritionnelle Latino-Martel, Paule; Equipe de Recherche en Epidemiologie Nutritionnelle Fezeu, Léopold; Equipe de Recherche en Epidemiologie Nutritionnelle Fassier, Philippine; Equipe de Recherche en Epidemiologie Nutritionnelle Hercberg, Serge; Equipe de Recherche en Epidemiologie Nutritionnelle; Hopital Avicenne, Public Health Department Touvier, Mathilde; Equipe de Recherche en Epidemiologie Nutritionnelle
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Nutrition and metabolism, Oncology
Keywords:	Breast tumours < ONCOLOGY, nutrient profiling system, nutrition policy, prospective study, food labelling

SCHOLARONE™
Manuscripts

Are self-reported unhealthy food choices associated with an increased risk of breast cancer: prospective cohort study using the British Food Standards Agency Nutrient Profiling System

Mélanie Deschasaux[1,2*], Chantal Julia[1,3], Emmanuelle Kesse-Guyot[1,2], Lucie Lécuyer[1,2], Solia Adriouch[1], Caroline Méjean[1], Pauline Ducrot[1], Sandrine Péneau[1], Paule Latino-Martel[1,2], Léopold K Fezeu[1], Philippine Fassier[1,2], Serge Hercberg[1,2,3], Mathilde Touvier[1,2]

[1] Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), U1153 Inserm, U1125, Inra, Cnam, Paris 13 University, Nutritional Epidemiology Research Team (EREN), Bobigny, France

[2] French network for Nutrition And Cancer Research (NACRe network), www.inra.fr/nacre

[3] Public Health Department, Avicenne Hospital, Bobigny, France

Corresponding author: Mélanie Deschasaux, Sorbonne Paris Cité Epidemiology and Statistics Research Center (CRESS), Nutritional Epidemiology Research Team (EREN), Inserm U1153, Inra U1125, Cnam, Paris 13 University, SMBH Paris 13, 74, rue Marcel Cachin, F-93017, Bobigny Cedex, France; e-mail:m.deschasaux@eren.smbh.univ-paris13.fr; Telephone number: +33 1 48 38 89 44.

Word count: 4064

For peer review only

ABSTRACT

Objectives: French authorities are considering the implementation of a simplified nutrition labeling system on food products to help consumers make healthier food choices. One of the most documented candidates (5-CNL/Nutri-score) is based on the British Food Standards Agency Nutrient Profiling System (FSA-NPS), a score calculated for each food/beverage using the 100g-amount of energy, sugar, saturated fatty acid, sodium, fibers, proteins, and fruits and vegetables. To assess its potential public health relevance, studies were conducted on the association between the nutritional quality of the diet, measured at the individual level by an energy-weighted mean of all FSA-NPS scores of foods usually consumed (FSA-NPS DI), and the risk of chronic diseases. The present study aimed at investigating the relationship between the FSA-NPS DI and breast cancer risk.

Design: prospective study

Setting: population-based, NutriNet-Santé cohort, France

Participants: 46,864 women aged ≥35y who completed ≥3 24h-dietary records during their first 2y of follow-up.

Primary outcome measure: Associations between FSA-NPS DI and breast cancer risk (555 incident breast cancers diagnosed between 2009 and 2015) were characterized by multivariable-adjusted Cox proportional hazard models.

Results: A higher FSA-NPS DI (lower nutritional quality of the diet) was associated with an increased breast cancer risk ($HR_{1\text{-point increment}}$=1.06 (1.02-1.11), P=0.005; $HR_{Q5vs.Q1}$=1.52 (1.11-2.08), P-trend=0.002). Similar trends were observed in pre- and post-menopausal women ($HR_{1\text{-point increment}}$=1.09 (1.01-1.18) and 1.05 (1.00-1.11) respectively).

This study was based on an observational cohort using self-reported dietary data thus residual confounding cannot be entirely ruled out. Finally, this holistic approach does not allow investigating which factors in the diet most specifically influence breast cancer risk.

Conclusions: These results suggested that unhealthy food choices, as characterized by the FSA-NPS, may be associated with an increase in breast cancer risk, supporting the potential public health relevance of using this profiling system in the framework of public health nutritional measures.

Keywords: breast cancer, Nutrient Profiling System, nutrition policy, food labelling, prospective study

For peer review only

ARTICLE SUMMARY

Strengths and limitations of this study

- This study examined the association between an indicator of the overall nutritional quality of the diet based on the Food Standards Agency Nutrient Profiling System (FSA-NPS DI) and the incidence of breast cancer using data from a large prospective cohort study, NutriNet-Santé
- Dietary intakes were assessed using repeated 24h-dietary records based on a recent food composition database with a large choice of items (>3300) allowing a better insight into the food products consumed and their intrinsic nutritional composition
- Unlike other a priori scores, components of the FSA-NPS DI cannot be studied separately since the FSA-NPS DI is first calculated at the food level (FSA-NPS) and then aggregated at the individual level. In addition, the calculation of the FSA-NPS score (Supplemental file 1) is based on thresholds and is conditional. Thus, the specific contribution of each component of the FSA-NPS DI score to breast cancer risk could not be studied.
- This study included volunteers involved in a long-term cohort study investigating the association between nutrition and health, with overall more health-conscious behaviors and higher professional and/or educational level compared to the general population so that unhealthy dietary behaviors may have been underrepresented.
- This study was based on self-declared dietary intakes and on an observational cohort, thus residual confounding cannot be ruled out even though a lot of potential confounders were taken into account.

For peer review only

INTRODUCTION

Breast cancer is the most common female cancer worldwide, with 1.7 million new cases diagnosed in 2012, representing 25% of all cancers [1]. According to the estimations of the World Cancer Research Fund/American Institute for Cancer Research (WCRF/AICR), around one third of breast cancers could be avoided with appropriate diet, body fatness and physical activity [2].

Nutrition has therefore the potential to be a key factor in breast cancer prevention since it can be modified at the individual level and thus can be targeted by public health policies. To help consumers make healthier food choices, several scientific organizations worldwide have recommended the implementation of a simplified nutrition labeling system on the front-of-pack of food products [3-7]. In France, a five-color labeling system (Five-Color Nutrition Label, 5-CNL) based on the British Food Standards Agency Nutrient Profiling System (FSA-NPS) [8;9] has been proposed to summarize the overall nutritional quality of food products [10]. The FSA-NPS attributes a single score to food products based on a limited number of input variables: amount per 100g of energy, total sugars, saturated fatty acids (SFA), sodium, fruits and vegetables, dietary fibers and proteins. This scoring system was initially developed and validated in the UK, where it is used for advertising regulation [8;9;11;12], and it has been adapted and validated in the French context [13-16]. At the individual level, the nutritional quality of the diet can be characterized with a dietary index based on the FSA-NPS (FSA-NPS DI). The FSA-NPS DI has been associated to food and nutrient intakes, nutritional status and adherence to the French nutritional recommendations [17;18].

To evaluate the relevance and potential public health impact of the 5-CNL adoption, it is important to assess whether there is a relationship between the nutritional quality of food choices at the individual level, as graded by the FSA-NPS DI, and the occurrence of nutrition-related chronic diseases. To our knowledge, our group was the first to investigate the associations between the FSA-NPS DI and health outcomes. Using prospective designs, studies were conducted in the SU.VI.MAX cohort (13,017 participants, 1994-2007) on the associations between the FSA-NPS DI and 13-year weight gain/obesity onset [19], metabolic syndrome [20], cardiovascular diseases [21] and cancer [22]. A higher FSA-NPS DI, reflecting a diet of lower nutritional quality, was associated with an increased risk for all the studied outcomes and, in particular, with an increased risk of cancer overall [22]. No significant association with breast cancer risk was detected in this study [22], but the statistical power was limited for site-specific analyses (n=125 breast cancer cases).

For peer review only

Thus, our objective was to study the association between the FSA-NPS DI (an indicator of the nutritional quality of the diet based on a nutrient profiling system) and breast cancer risk, using data from NutriNet-Santé, a large prospective cohort with up-to-date assessment of dietary intakes.

For peer review only

METHODS

Study population

The NutriNet-Santé study is a French ongoing web-based cohort launched in 2009 with the objective to study the associations between nutrition and health as well as the determinants of dietary behaviors and nutritional status. This cohort has been previously described in details [23]. Participants aged ≥18y with access to the Internet are continuously recruited since May 2009 among the general population by means of vast multimedia campaigns. All questionnaires are completed online though a dedicated website (www.etude-nutrinet-sante.fr). The NutriNet-Santé study is conducted according to the Declaration of Helsinki guidelines and was approved by the Institutional Review Board of the French Institute for Health and Medical Research (IRB Inserm n°0000388FWA00005831) and the "Commission Nationale de l'Informatique et des Libertés" (CNIL n°908450/n°909216). Electronic informed consent is obtained from each participant (EudraCT no.2013-000929-31).

Data collection

At inclusion, participants fulfilled a set of five questionnaires on socio-demographic and lifestyle characteristics [24] (e.g. occupation, educational level, smoking status, alcohol consumption, number of children), anthropometrics [25;26] (e.g. height, weight), dietary intakes (see below), physical activity (validated IPAQ questionnaire) [27], and health status (e.g. personal and family history of diseases, medication use including hormonal treatment for menopause and oral contraception, menopausal status). Follow-up of participants began when participants answered their last baseline questionnaire. The date of completion of the last baseline questionnaire is thus used as inclusion date. Participants are then invited to complete these five baseline questionnaires every year as part of the follow-up.

Dietary intakes were assessed at baseline and every six months through series of three non-consecutive validated web-based 24h-dietary records, randomly assigned over a 2-week period (2 weekdays and 1 weekend day) [28-30]. Thus, over the first two years of follow-up, up to five series of three 24h-dietary records could have been completed. To be considered as valid, a series must have included at least two out of three 24h-dietary records. Participants used a dedicated interface of the study website to declare all foods and beverages consumed during a 24h-period: three main meals (breakfast, lunch, dinner) or any other eating occasion. Portion sizes were estimated using validated photographs [31]. Mean daily energy, alcohol and nutrient intakes were estimated using a published French food composition table (>3300 items) [32] and a weighting for week days and week-end days. Amounts consumed from

For peer review only

composite dishes were estimated using French recipes validated by food and nutrition professionals. Dietary underreporting was identified on the basis of the method proposed by Black [33].

FSA-NPS DI computation

As described previously [9;13;34], the FSA-NPS score for all foods (processed and unprocessed) and beverages was computed based on the nutrient content for 100g. FSA-NPS scores for foods and beverages are based on a discrete continuous scale from -15 (most healthy) to +40 (less healthy) (**Supplemental file 1**). FSA-NPS score allocates points (0-10) for the amount of energy (kJ), total sugar (g), SFA (g) and sodium (mg). Points (0-5) are subtracted from the previous sum based on the amount of fruits and vegetables (%, including legumes and nuts), fibers (g) and proteins (g). Specific modifications of the score for particular food groups were made to maintain a high consistency with French nutritional recommendations, as proposed by the French High Council for Public Health (HCSP) [34].
In a second step, the FSA-NPS DI was computed at the individual level using arithmetic energy-weighted means with the following equation [17], in which FS_i represents the food (or beverage) score, and E_i represents the energy intake from this food or beverage (all 24h-dietary records from the first two years of follow-up were averaged to a mean 24-hour energy intake from this food/beverage):

$$\text{FSA} - \text{NPS DI} = \frac{\sum_{i=1}^{n} FS_i E_i}{\sum_{i=1}^{n} E_i}$$

Increasing FSA-NPS DI reflects decreasing nutritional quality of foods consumed.

Case ascertainment

Participants self-declared health events through the yearly health status questionnaire, through a specific check-up questionnaire for health events (every three months) or at any time through a dedicated interface on the study website. Following this declaration, participants are invited to send their medical records (diagnosis, hospitalization, etc.) and, if necessary, the study physicians contact the participants' treating physician or the medical structures to collect additional information. Then, data are reviewed by an independent physician expert committee which validates all major health events (such as cancers). Cancer cases were classified using the International Chronic Diseases Classification, 10th Revision, Clinical Modification (ICD-10) [35]. In this study, all first primary breast cancers diagnosed between the inclusion and August 2015 were considered as cases. Information on death and cause of

For peer review only

death was obtained through linkage to the national database on mortality of the French population [36].

Statistical analyses

So far, 77,034 women without cancer at baseline provided at least three valid 24h-dietary records during their first two years of follow-up. Women aged <35y at baseline (n=29,249) were excluded due to a very low susceptibility to develop breast cancer in these women [37] and a potentially limited influence of nutrition on breast cancers diagnosed in young women. Women with a null follow-up were also excluded from the analyses (i.e. women for whom baseline questionnaires were the last completed questionnaires, n=921), thus leaving 46,864 women included in the analyses (flowchart in **Supplementary file 2**).

For each woman, the FSA-NPS DI and usual dietary intakes were calculated using all 24h-dietary records available in their first two years of follow-up. Associations between the FSA-NPS DI (continuous variable and quintiles) and breast cancer risk were characterized (HR and 95%CI) using multivariable Cox proportional hazards models with age as the primary time variable. We confirmed that the assumptions of proportionality were satisfied through examination of the log–log (survival) vs. log–time plots. Tests for linear trends were performed with the ordinal score on quintiles of FSA-NPS DI. Women contributed person-time to the model until the date of cancer diagnosis, the date of last completed questionnaire, the date of death or August 2015, whichever occurred first. Women who reported a cancer other than breast cancer during the study period were included and censored at the date of diagnosis (except basal cell skin carcinoma, not considered as cancer).

Models were adjusted for classic risk factors for breast cancer: age (time-scale), BMI (kg/m², continuous), height (cm, continuous), physical activity (high, moderate, low, computed following IPAQ recommendations [38]), smoking status (never smokers, former smokers, occasional smokers, smokers), number of dietary records (continuous), alcohol intake (g/d, continuous), energy intake (without alcohol, g/d, continuous), family history of cancer (yes/no), educational level (<high-school degree, <2 years after high-school degree, ≥2 years after high-school degree), number of biological children (continuous), menopausal status at baseline (pre-menopause, perimenopause, post-menopause), hormonal treatment for menopause (postmenopausal women, yes/no) and oral contraception use (premenopausal women, yes/no).

Interaction analysis was conducted between BMI and the FSA-NPS DI and stratified analyses were performed by overweight status (BMI < vs. ≥25kg/m^2).

For peer review only

Sensitivity analyses were performed including only women that provided at least six 24h-dietary records during their first two years of follow-up or excluding cases diagnosed during their first year of follow-up. Analyses were also performed on invasive breast cancer cases only, by hormonal receptor status of the tumors and by menopausal status. For the latter, women contributed person-time to the "pre-menopause model" until their age of menopause and to the "post-menopause model" from their age of menopause. Age at menopause was determined using the yearly health status questionnaires available during the follow-up.

For all covariates except physical activity, ≤ 5% of values were missing and were imputed to the modal value. For physical activity (N=6,328 missing values), a "missing class" was introduced into the models.

All tests were two-sided, and $P<0.05$ was considered statistically significant. SAS version 9.4 (SAS Institute) was used for the analyses.

For peer review only

RESULTS

Between May 2009 and August 2015 (median follow-up time: 4.0y; 174,491 person-years), 555 incident breast cancer cases were diagnosed: 171 premenopausal and 384 postmenopausal; 71.4% ER+/PR+, 14.7% ER-/PR-, 13.6% ER+/PR-, 0.3% ER-/PR+ (data available for 361 cases); 83.6% invasive and 16.4% *in situ* (data available for 463 cases). Mean age at diagnosis was 56.6y (SD=9.2) and mean baseline-to-diagnosis time was 2.4y (SD=1.6). Mean number of dietary records per participant over their first two years of follow-up was 5.9 (SD=2.8).

In **Table 1,** the characteristics of participants at baseline are described overall and by quintiles of the FSA-NPS DI. Mean FSA-NPS DI was 5.9 (SD=2.2; min=-5.8; max=18.1). Women with a higher FSA-NPS DI (diet of lower nutritional quality), were more likely to be young, to smoke, to have a higher educational level and to have higher energy or alcohol intakes. As expected, women in the lowest quintiles of FSA-NPS DI (diet of higher nutritional quality) had overall healthier food intakes: higher intakes of fiber, fruits, vegetables, legume, fish and lower intakes of red and processed meat and lipids.

Compared to women that provided at least three 24h-dietary records over their first two years of follow-up, women that did not (15,918 women with a non-null follow-up) were younger, pre-menopause, were more likely to be overweight/obese, to smoke, to practice physical activity and were less likely to have a family history of cancer or to take a hormonal treatment for menopause [data not tabulated].

Associations between the FSA-NPS DI and breast cancer risk overall and by menopausal status are shown in **Table 2**. A direct association was observed between the FSA-NPS DI and breast cancer risk: $HR_{Q5vs.Q1}$=1.52 (95%CI 1.11-2.08), P-trend=0.002; $HR_{per\ 1\text{-}unit\ increment}$=1.06 (1.02-1.11), P=0.005. These associations were similarly observed in premenopausal women ($HR_{Q5vs.Q1}$=2.46 (1.27-4.75), P-trend=0.004; $HR_{per\ 1\text{-}unit\ increment}$=1.09 (1.01-1.18), P=0.03) and in postmenopausal women ($HR_{Q5vs.Q1}$=1.25 (0.85-1.84), P-trend=0.09; $HR_{per\ 1\text{-}unit\ increment}$=1.05 (1.00-1.11), P=0.06), although the associations seemed stronger for premenopausal women and only trends were observed for postmenopausal women (P-interaction=0.06).

Analyses performed by overweight status showed that associations tended to be stronger in non-overweight women (368 cases/ 31,401 non-cases, $HR_{Q5vs.Q1}$=1.97 (95%CI 1.31-2.96), P-trend=0.0007; $HR_{per\ 1\text{-}unit\ increment}$=1.09 (1.03-1.15), P=0.003) compared to overweight/obese women (187 cases/14,908 non-cases, $HR_{Q5vs.Q1}$=1.02 (95%CI 0.61-1.73), P-trend=0.6; $HR_{per\ 1\text{-}}$

For peer review only

$_{\text{unit increment}}$=1.03 (0.95-1.11), P=0.5), but the interaction was not statistically significant (P=0.07).

Information regarding hormone receptor status was not available for all cases (ER status: 361 cases, PR status: 362 cases, ER/PR status: 361 cases). Significant direct associations between the FSA-NPS DI and breast cancer risk were observed for breast cancer types PR- (102 cases/ 46,762 non-cases) and ER+/PR- (49 cases/ 46,815 non-cases). For ER+ tumours, the linear trend was not statistically significant (P=0.07, 307 cases/46,557 non-cases) but compared to women in the lowest quintile of FSA-NPS DI, those with higher scores had an increased breast cancer risk (e.g. $HR_{Q5vs.Q1}$=1.60 (1.04-1.46)). Associations were non-significant for the other hormone receptor status (**Supplementary file 3**). However, these exploratory findings should be considered with caution due to limited statistical power for analyses by cancer sub-types.

Similar results were observed when analyses excluded cases diagnosed during their first year of follow-up (425 cases/46,309 non-cases included; $HR_{Q5vs.Q1}$=1.54 (1.08-2.19), P-trend=0.007; $HR_{\text{per 1-unit increment}}$=1.07 (1.02-1.12), P=0.01) or when analyses were restricted to invasive breast cancers (387 cases/46,309 non-cases; $HR_{Q5vs.Q1}$=1.51 (1.03-2.22), P-trend=0.01; $HR_{\text{per 1-unit increment}}$=1.06 (1.01-1.12), P=0.03).

Results were also similar when analyses were restricted to women that provided at least 6 24h-dietary records during their first two years of follow-up (399 cases/25,439 non-cases; $HR_{Q5vs.Q1}$=1.63 (1.11-2.38), P-trend=0.006; $HR_{\text{per 1-unit increment}}$=1.08 (1.02-1.14), P=0.01) [data not tabulated].

Finally, similar but weaker trends were observed when women aged <35y at baseline were included in the analyses (585 cases/ 74,617 non-cases, $HR_{Q5vs.Q1}$=1.17 (95%CI 0.83-1.64), P-trend=0.1; $HR_{\text{per 1-unit increment}}$=1.05 (1.01-1.10), P=0.02).

For peer review only

DISCUSSION

In this prospective study conducted in a large sample of women from the French general population, a higher FSA-NPS DI, which reflects a diet composed of food products of lower nutritional quality, was associated with a 52% increase in breast cancer risk (highest vs. lowest quintile of the FSA-NPS DI score).

In a previous study performed in the SU.VI.MAX cohort [22], we observed a direct association between the FSA-NPS DI and cancer risk overall but did not detect a significant association for breast cancer risk, probably due to limited power in site-specific analyses (n=125 breast cancer cases, 13y-follow-up). To our knowledge, no other study investigated the relationship between breast cancer risk and a score that characterizes the nutritional quality of an individual's diet based on a nutrient profiling system at the level of foods/beverages consumed.

However a few studies have been conducted on the association between NPS-based dietary scores and other health outcomes. While in this study, we used the FSA-NPS as a continuous score at the food/beverage level as a basis for the construction of the FSA-NPS DI at the individual level, the FSA-NPS was also recently used to define a variety score of "healthier" and "less healthy" foods/beverages (Ofcom binary cut-off used for advertising regulation in the UK [12]). This binary indicator was then studied in relation to mortality in the Whitehall II cohort [39]. The authors observed that a greater variety of healthier foods, as defined with the FSA-NPS Ofcom binary cut-off, was associated with a reduced all-cause and cancer mortality while a greater variety of less healthy food was not associated with the studied outcomes. No association was observed when another nutrient profiling system, the SAIN, LIM [40;41], was used [39].

To our knowledge, the Overall Nutritional Quality Index (ONQI-f) is the only other dietary score based on a nutrient profiling system that has been studied in relation to health outcomes [42]. It was tested in association with chronic diseases and mortality within the Nurses' Health Study and the Health Professionals Follow-up Study [42]. A higher ONQI-f, reflecting a higher nutritional quality of the diet, was associated with a decreased risk of cardiovascular diseases, diabetes and mortality but was not associated with cancer. Some arguments may explain this lack of association: 1) the ONQI-f is based on 30 nutrients among which few have shown a consistent association with cancer risk, thus, its relevance regarding the cancer outcome may be lower than for other outcomes; 2) dietary intakes were assessed with an

For peer review only

aggregated food frequency questionnaire (135-138 items), which provides less precise estimates than 24h-dietary records (as used in our study).

These studies are, to our knowledge, the only ones that investigated the associations between health outcomes and individual dietary indexes derived from nutrient profiling systems at the food level. Other a priori scores have been designed based on the intake of specific food groups or nutrients and/or other information (e.g. body fatness, physical activity), but not based on a nutrient profiling system at the food/beverage level. These scores were studied prospectively in relation to breast cancer risk and provided relatively contrasted results: 1) scores measuring the adherence to a specific type of diet such as the Mediterranean diet score (no association in prospective cohorts, inverse association in case-control studies [43-45]) or the Healthy Nordic Food Index (HNFI, no association [46]), 2) scores reflecting the adherence to general nutritional recommendations for the population such as the World Health Organization Healthy Diet Index, WHO HDI [47], the Alternate Healthy Eating Index, AHEI [45;48], the Recommended Food Score, RFS [48], the Diet Quality Index revised, DQI-R, [48] or the Dietary Approaches to Stop Hypertension (DASH) [45] (no association overall), and 3) scores measuring the adherence to cancer-specific nutritional recommendations such as the WCRF/AICR adherence score (inverse associations [49;50]) or the American Cancer Society (ACS) cancer prevention guidelines score (inverse association [51]). In these studies, differences according to hormonal receptor status of the tumors have been suggested, with inconsistent results. Indeed, inverse associations between a "healthier" diet and breast cancer risk were particularly observed in ER- type (AHEI, RFS, aMed) [48], ER-/PR+ type (Mediterranean diet score) [43], and ER-/PR-/HER2+ type (DASH) [45], but also with ER+/PR+ type (WCRF/AICR adherence score) [49] and ER+/PR- type ("healthy/Mediterranean" pattern) [52]. In our study, information regarding hormonal receptor status of the tumors was only partially available and the statistical power was limited in the analyses (**Supplementary file 3**), thus preventing to derive firm conclusions.

Overall, these studies involving a priori scores provided interesting insights into the relationships between nutrition and breast cancer risk. Although these a priori scores and the FSA-NPS DI included similar nutritional components, the approaches differed, making the comparison between our study and previous findings not straightforward (even though our results were in line with those obtained with scores measuring the adherence to cancer-specific nutritional recommendations [49-51]). The FSA-NPS DI is not primarily built at the individual level but is rather derived from a nutrient profiling system at the food level (FSA-

For peer review only

NPS) thus taking into account the nutritional quality of each food/beverage consumed and not only of the overall diet or overall consumption of food groups. In addition, the objective behind the FSA-NPS DI construction was not to obtain the best predictive score for breast cancer but to specifically test its association with breast cancer risk, as the FSA-NPS is envisioned to serve as a basis for food labelling in the framework of public health policies in several countries such as France and Australia. The FSA-NPS displays several key advantages in a public health context: 1) it grades the nutritional quality of each food/beverage and thus reflects the variation of nutritional quality between but also within food groups, 2) it has been designed in a perspective of prevention of a large range of chronic diseases (not only breast cancer), and 3) it is easy-to-compute for industrials and public health stakeholders.

Our results are consistent with current evidence from epidemiological and mechanistic studies regarding the association between nutrition and breast cancer. Most of the input variables for the FSA-NPS are indeed parameters for which associations with breast cancer have been established either directly (e.g. dietary fibers [53]) or indirectly, through an association with body fatness, a major risk factor for postmenopausal breast cancer [53-55] (e.g. energy content, total sugars and SFA as components of energy-dense foods; fruits and vegetables as components of low-energy foods).

In our study, although similar trends were observed in pre- and post-menopausal women for the association between the FSA-NPS DI and breast cancer risk, this association was nonetheless stronger in pre-menopausal women. This may be explained by the fact that women pre-menopause were more likely to score high on the FSA-NPS DI, thus resulting in a clearer/stronger association: mean±SD FSA-NPS DI was 6.3±2.3 in women pre-menopause (median:6.4, 25th-75th percentiles: 4.9-7.8) and 5.5±2.1 in women post-menopause (median:5.5, 25th-75th percentiles: 4.1-6.9).

Strengths of this study pertained to its prospective design, its large sample size, and the assessment of usual dietary intakes using repeated 24h-dietary records based on a recent food composition database with a large choice of items (>3300). The latter allowed a better insight into the food products consumed and their intrinsic nutritional quality compared to studies that used a food frequency questionnaire (more aggregated food items). However, some limitations should be acknowledged. First, caution is needed regarding the extrapolation of these results to the entire French population since this study included volunteers involved in a long-term cohort study investigating the association between nutrition and health, with overall more health-conscious behaviors and higher professional and/or educational level compared

For peer review only

to the general population. Thus, unhealthy dietary behaviors may have been underrepresented in this study, which may have weakened the observed associations. Second, information regarding cancer stage was not available. Third, unlike other a priori scores, components of the FSA-NPS DI cannot be studied separately since 1) the FSA-NPS DI is first calculated at the food level (FSA-NPS) and then aggregated at the individual level and 2) the calculation of the FSA-NPS score (Supplemental file 1) is based on thresholds and conditions that are inter-related between the different score components. Fourth, as usually done in nutritional epidemiology, dietary intakes were estimated based on averaged intakes from all 24h-dietary records collected over the first two years of follow-up. Although diet may change over time, it is usually hypothesized that this estimation reflects general eating behavior throughout the adult life [56]. This very classical method allowed us to obtain a reliable estimation of usual dietary intakes, while respecting the prospective design (i.e. estimation of usual dietary intakes prior to cancer diagnosis). Indeed, breast cancer is a disease with relatively long latency so that the involvement of nutritional factors is supposed to be based on long-term processes. Thus, it is important to guarantee sufficient delay between nutritional exposure and cancer outcome. This is why we tested a model (sensitivity analysis) where cancer cases diagnosed during the first year of follow-up were excluded (similar results). In our study, although the follow-up time was appropriate to perform etiological analyses, it did not necessarily guarantee this sufficient delay. Hence, our estimation of usual dietary intakes may reflect dietary protective and risk factors that may have played a role in the first steps of carcinogenesis (initiation) but also later in the carcinogenic process (progression). Nonetheless, previous studies with longer follow-up observed associations between diet and breast cancer risk, suggesting that nutritional factors could play a role in cancer initiation and not only in cancer progression [45;48-52]. Finally, this study was based on an observational cohort and thus residual confounding cannot be entirely ruled out even though a wide range of confounding factors were taken into account.

In conclusion, the FSA-NPS has been designed to characterize the nutritional quality of foodstuffs and to highlight products with a good nutritional profile that should be promoted and products with a lower nutritional quality that should not. The results of this observational study suggest that the self-declared consumption of food products of lower nutritional quality (as characterized by a higher FSA-NPS) may be associated with an increased risk of breast cancer. Along with other etiological observational studies [19-22], these findings suggest that this nutrient profiling system might be of interest in the framework of public health nutritional measures such as front-of-pack nutrition labeling or taxes.

For peer review only

ACKNOWLEDGEMENTS

The authors thank all the volunteers of the NutriNet-Santé cohort. We extend special thanks to Véronique Gourlet for the statistical analyses and to Nathalie Arnault, Stephen Besseau, Laurent Bourhis, Yasmina Chelghoum, Than Duong Van, Younes Esseddik, Paul Flanzy, Charlie Ménard, Mac Rakotondrazafy, Fabien Szabo, Roland Andrianasolo, Fatoumata Diallo, Cédric Agaesse, Claudia Chahine, Marion Genest and Ludivine Ursule for their technical contribution to the NutriNet-Santé study.

Authors' contribution: The authors' contributions were as follow – MD and MT: designed the research; SH, MT, CJ, EKG: conducted the research; MD and MT: supervised statistical analysis; MD and MT: wrote the paper; CJ, EKG, LL, SA, CM, PD, SP, PLM, LF, PF, SH: contributed to the data interpretation and revised each draft for important intellectual content. All authors read and approved the final manuscript. MD and MT had primary responsibility for the final content.

Funding: Mélanie Deschasaux, Philippine Fassier and Solia Adriouch were funded by PhD grants from the Région Ile-de-France (public funding: Cancéropôle Ile-de-France and CORDDIM). The NutriNet-Santé study was supported by the following public institutions: Ministère de la Santé, Institut de Veille Sanitaire (InVS), Institut National de la Prévention et de l'Education pour la Santé (INPES), Région Ile-de-France (CORDDIM), Institut National de la Santé et de la Recherche Médicale (INSERM), Institut National de la Recherche Agronomique (INRA), Conservatoire National des Arts et Métiers (CNAM) and Université Paris 13. The funders had no role in the design, implementation, analysis, or interpretation of the data.

Competing interests: The authors have no conflict of interest to disclose.

Data sharing: All relevant data are in the manuscript and its supporting files. No additional data available.

For peer review only

REFERENCES

(1) WHO/IARC. All Cancers: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012. GLOBOCAN 2012 [2016 Available from: URL:http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx

(2) WCRF/AICR. Cancer preventability estimates for food, nutrition, body fatness, and physical activity. WCRF/AICR [2016 Available from: URL:http://www.wcrf.org/int/cancer-facts-figures/preventability-estimates/cancer-preventability-estimates-diet-nutrition

(3) Food Standards Agency, Welsh Government, Scottish Government, Department of Health. Front of Package Nutrition Labelling: Joint response to consultation. 2013.

(4) Gill T, King L, Vita P, Caterson I, Colagiuri S, Colagiuri R et al. A 'state of the knowledge' assessment of comprehensive interventions that address the drivers of obesity. A Rapid Assessment Prepared for the National Health and Medical Research Council (NHMRC). 2010. Sydney, The Boden Institute of Obesity, Nutrition, Exercise and Eating Disorders, University of Sydney.

(5) Institute of Medecine. Front-of-Package Nutrition Rating Systems and Symbols: Promoting Healthier Choices. 2012. Washington, D.C., The National Academies Press.

(6) WHO. 2008-2013 action plan for the global strategy for the prevention and control of non-communicable diseases. 2009. Geneva, World Health Organization.

(7) WHO Europe. Action Plan for implementation of the European Strategy for the Prevention and Control of Noncommunicable Diseases 2012–2016. 2011. Copenhagen, World Health Organization Regional Office for Europe.

(8) Arambepola C, Scarborough P, Rayner M. Validating a nutrient profile model. *Public Health Nutr* 2008; 11(4):371-378.

(9) Rayner M, Scarborough P, Stockley P, Boxer A. Nutrient profiles: Development of Final Model. Final Report [online]. London, FSA . 2005.

(10) Hercberg S. Propositions pour un nouvel élan de la politique nutritionnelle française de santé publique dans le cadre de la stratégie nationale de santé. 1ère partie: mesures concernant la prévention nutritionnelle [in French]. 2013.

(11) Rayner M, Scarborough P, Stockley L. Nutrient profiles: Applicability of currently proposed model for uses in relation to promotion of food to children aged 5-10 and adults. [online]. London, Foods Standard Agency. 2005.

(12) Rayner M, Scarborough P, Lobstein T. The UK Ofcom Nutrient Profiling Model: Defining 'healthy' and 'unhealthy' foods and drinks for TV advertising to children. London, OfCom . 2009.

(13) Julia C, Kesse-Guyot E, Touvier M, Mejean C, Fezeu L, Hercberg S. Application of the British Food Standards Agency nutrient profiling system in a French food composition database. *Br J Nutr* 2014; 112(10):1699-1705.

(14) Julia C, Kesse-Guyot E, Ducrot P, Peneau S, Touvier M, Mejean C et al. Performance of a five category front-of-pack labelling system - the 5-colour nutrition label - to differentiate nutritional quality of breakfast cereals in France. *BMC Public Health* 2015; 15:179.

(15) Julia C, Ducrot P, Peneau S, Deschamps V, Mejean C, Fezeu L et al. Discriminating nutritional quality of foods using the 5-Color nutrition label in the French food market: consistency with nutritional recommendations. *Nutr J* 2015; 14:100.

(16) Haut Conseil de la Santé Publique. Avis relatif à l'information sur la qualité nutritionnelle des produits alimentaires. 2015. Paris, HCSP.

For peer review only

(17) Julia C, Touvier M, Mejean C, Ducrot P, Peneau S, Hercberg S et al. Development and validation of an individual dietary index based on the British Food Standard Agency nutrient profiling system in a French context. *J Nutr* 2014; 144(12):2009-2017.

(18) Julia C, Mejean C, Touvier M, Peneau S, Lassale C, Ducrot P et al. Validation of the FSA nutrient profiling system dietary index in French adults-findings from SUVIMAX study. *Eur J Nutr* 2015.

(19) Julia C, Ducrot P, Lassale C, Fezeu L, Mejean C, Peneau S et al. Prospective associations between a dietary index based on the British Food Standard Agency nutrient profiling system and 13-year weight gain in the SU.VI.MAX cohort. *Prev Med* 2015; 81:189-194.

(20) Julia C, Fezeu LK, Ducrot P, Mejean C, Peneau S, Touvier M et al. The Nutrient Profile of Foods Consumed Using the British Food Standards Agency Nutrient Profiling System Is Associated with Metabolic Syndrome in the SU.VI.MAX Cohort. *J Nutr* 2015; 145(10):2355-2361.

(21) Adriouch S, Julia C, Kesse-Guyot E, Mejean C, Ducrot P, Peneau S et al. Prospective association between a dietary quality index based on a nutrient profiling system and cardiovascular disease risk. *Eur J Prev Cardiol* 2016.

(22) Donnenfeld M, Julia C, Kesse-Guyot E, Mejean C, Ducrot P, Peneau S et al. Prospective association between cancer risk and an individual dietary index based on the British Food Standards Agency Nutrient Profiling System. *Br J Nutr* 2015; 114(10):1702-1710.

(23) Hercberg S, Castetbon K, Czernichow S, Malon A, Mejean C, Kesse E 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* 2010; 10:242.

(24) Vergnaud AC, Touvier M, Mejean C, Kesse-Guyot E, Pollet C, Malon A et al. Agreement between web-based and paper versions of a socio-demographic questionnaire in the NutriNet-Sante study. *Int J Public Health* 2011; 56(4):407-417.

(25) Lassale C, Peneau S, Touvier M, Julia C, Galan P, Hercberg S et al. Validity of web-based self-reported weight and height: results of the Nutrinet-Sante study. *J Med Internet Res* 2013; 15(8):e152.

(26) Touvier M, Mejean C, Kesse-Guyot E, Pollet C, Malon A, Castetbon K et al. Comparison between web-based and paper versions of a self-administered anthropometric questionnaire. *Eur J Epidemiol* 2010; 25(5):287-296.

(27) Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003; 35(8):1381-1395.

(28) Lassale C, Castetbon K, Laporte F, Deschamps V, Vernay M, Camilleri GM et al. Correlations between Fruit, Vegetables, Fish, Vitamins, and Fatty Acids Estimated by Web-Based Nonconsecutive Dietary Records and Respective Biomarkers of Nutritional Status. *J Acad Nutr Diet* 2016; 116(3):427-438.

(29) Lassale C, Castetbon K, Laporte F, Camilleri GM, Deschamps V, Vernay M et al. Validation of a Web-based, self-administered, non-consecutive-day dietary record tool against urinary biomarkers. *Br J Nutr* 2015; 113(6):953-962.

(30) Touvier M, Kesse-Guyot E, Mejean C, Pollet C, Malon A, Castetbon K et al. Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. *Br J Nutr* 2011; 105(7):1055-1064.

(31) Le Moullec N, Deheeger M, Preziosi P, Montero P, Valeix P, Rolland-Cachera M et al. Validation du manuel photo utilisé pour l'enquête alimentaire de l'étude SU.VI.MAX. [Validation of the food portion size booklet used in the SU.VI.MAX study] (in French). *Cah Nutr Diet* 1996; 31:158-164.

For peer review only

(32) Arnault N, Caillot L, Castetbon K, Coronel S, Deschamps V, Fezeu L et al. Table de composition des aliments, étude NutriNet-Santé. [Food composition table, NutriNet-Santé study] (in French). Paris: Les éditions INSERM/Economica; 2013.

(33) Black AE. Critical evaluation of energy intake using the Goldberg cut-off for energy intake:basal metabolic rate. A practical guide to its calculation, use and limitations. *Int J Obes Relat Metab Disord* 2000; 24(9):1119-1130.

(34) Haut Conseil de la Santé Publique. Opinion on information regarding the nutritional quality of foodstuffs. 2015. Paris, HCSP.

(35) WHO. ICD-10, International classification of diseases and related health problems. 10th revision. 2010. Geneva, Switzerland: World Health Organization.

(36) Inserm. CépiDC. Interrogation des données sur les causes de décès de 1979 à 2013. Inserm [2016 Available from: URL:http://www.cepidc.inserm.fr/inserm/html/index2.htm

(37) INCa. Incidence et mortalité estimées par classe d'âge et par localisation cancéreuse en 2012. 2016.

(38) IPAQ Group. Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ). 2005.

(39) Masset G, Scarborough P, Rayner M, Mishra G, Brunner EJ. Can nutrient profiling help to identify foods which diet variety should be encouraged? Results from the Whitehall II cohort. *Br J Nutr* 2015; 113(11):1800-1809.

(40) Afssa. Définition de profils nutritionnels pour l'accès aux allégations nutritionnelles et de santé: propositions et arguments [The setting of nutrient profiles for access to nutrition an health claims: proposals and arguments] (in French). 2008. Paris, Agence française de sécurité des aliments.

(41) Darmon N, Vieux F, Maillot M, Volatier JL, Martin A. Nutrient profiles discriminate between foods according to their contribution to nutritionally adequate diets: a validation study using linear programming and the SAIN,LIM system. *Am J Clin Nutr* 2009; 89(4):1227-1236.

(42) Chiuve SE, Sampson L, Willett WC. The association between a nutritional quality index and risk of chronic disease. *Am J Prev Med* 2011; 40(5):505-513.

(43) Schwingshackl L, Hoffmann G. Adherence to Mediterranean diet and risk of cancer: an updated systematic review and meta-analysis of observational studies. *Cancer Med* 2015; 4(12):1933-1947.

(44) Schwingshackl L, Hoffmann G. Does a Mediterranean-Type Diet Reduce Cancer Risk? *Curr Nutr Rep* 2016; 5:9-17.

(45) Hirko KA, Willett WC, Hankinson SE, Rosner BA, Beck AH, Tamimi RM et al. Healthy dietary patterns and risk of breast cancer by molecular subtype. *Breast Cancer Res Treat* 2016; 155(3):579-588.

(46) Li Y, Roswall N, Sandin S, Strom P, Adami HO, Weiderpass E. Adherence to a healthy Nordic food index and breast cancer risk: results from a Swedish cohort study. *Cancer Causes Control* 2015; 26(6):893-902.

(47) Cade JE, Taylor EF, Burley VJ, Greenwood DC. Does the Mediterranean dietary pattern or the Healthy Diet Index influence the risk of breast cancer in a large British cohort of women? *Eur J Clin Nutr* 2011; 65(8):920-928.

(48) Fung TT, Hu FB, McCullough ML, Newby PK, Willett WC, Holmes MD. Diet quality is associated with the risk of estrogen receptor-negative breast cancer in postmenopausal women. *J Nutr* 2006; 136(2):466-472.

For peer review only

(49) Harris HR, Bergkvist L, Wolk A. Adherence to the World Cancer Research Fund/American Institute for Cancer Research recommendations and breast cancer risk. *Int J Cancer* 2016; 138(11):2657-2664.

(50) Nomura SJ, Inoue-Choi M, Lazovich D, Robien K. WCRF/AICR recommendation adherence and breast cancer incidence among postmenopausal women with and without non-modifiable risk factors. *Int J Cancer* 2016; 138(11):2602-2615.

(51) Thomson CA, McCullough ML, Wertheim BC, Chlebowski RT, Martinez ME, Stefanick ML et al. Nutrition and physical activity cancer prevention guidelines, cancer risk, and mortality in the women's health initiative. *Cancer Prev Res (Phila)* 2014; 7(1):42-53.

(52) Cottet V, Touvier M, Fournier A, Touillaud MS, Lafay L, Clavel-Chapelon F et al. Postmenopausal breast cancer risk and dietary patterns in the E3N-EPIC prospective cohort study. *Am J Epidemiol* 2009; 170(10):1257-1267.

(53) Latino-Martel P, Cottet V, Druesne-Pecollo N, Pierre FH, Touillaud M, Touvier M et al. Alcoholic beverages, obesity, physical activity and other nutritional factors, and cancer risk: a review of the evidence. Crit Rev Oncol Hematol 99[308], 323. 2016.

(54) WCRF/AICR. Food, Nutrition, Physical Activity and the Prevention of Cancer: a Global Perspective. 2007. Washington, DC: AICR.

(55) WCRF/AICR. Systematic Literature Review - Continuous Update Project Report: The Associations between Food, Nutrition and Physical Activity and the Risk of Breast Cancer. 2010. Washington, DC: AICR.

(56) Willett WC. Nutritional Epidemiology, 2nd ed. New York: Oxford University Press; 1998.

For peer review only

Table 1 Baseline characteristics of the study population overall and according to quintiles of the FSA-NPS DI, NutriNet-Santé Cohort, France, 2009-2015

		Quintiles of the FSA-NPS DI					
	All women (n=46,864)	Q1 (n=9,349)	Q2 (n=9,395)	Q3 (n=9,387)	Q4 (n=9,415)	Q5 (n=9,318)	
	N (%) Mean±SD	N (%) Mean±SD	N (%) Mean±SD	N (%) Mean±SD	N (%) Mean±SD	N (%) Mean±SD	P-trend[a]
FSA-NPS DI	5.9±2.2	2.7±1.2	4.8±0.4	6.0±0.3	7.1±0.4	9.0±1.1	<.0001
Age, years	50.8±9.7	53.4±9.6	52.6±9.4	51.2±9.7	49.6±9.4	47.1±8.9	<.0001
Educational level							<.0001
< high-school degree	11269 (24.1)	2658 (28.4)	2345 (25.0)	2172 (23.1)	2083 (22.1)	2011 (21.6)	
≥high-school degree to < 2y after high-school degree	7834 (16.7)	1567 (16.8)	1579 (16.8)	1570 (16.7)	1556 (16.5)	1562 (16.8)	
≥ 2y after high-school degree	27761 (59.2)	5124 (54.8)	5471 (58.2)	5645 (60.1)	5776 (61.3)	5745 (61.6)	
Smoking status							<.0001
Non-smokers	22528 (48.1)	4630 (49.5)	4706 (50.1)	4615 (49.2)	4504 (47.8)	4073 (43.7)	
Former smokers	17904 (38.2)	3744 (40.0)	3640 (38.7)	3561 (37.9)	3527 (37.5)	3432 (36.8)	
Occasional smokers[b]	1622 (3.5)	257 (2.7)	302 (3.2)	336 (3.6)	350 (3.7)	377 (4.0)	
Smokers	4810 (10.3)	718 (7.7)	747 (7.9)	875 (9.3)	1034 (11.0)	1436 (15.4)	
Physical activity[c]							0.08
Low	13955 (34.4)	3312 (41.1)	2979 (36.4)	2800 (34.5)	2569 (31.5)	2295 (28.6)	
Moderate	17062 (42.1)	3224 (40.0)	3462 (42.4)	3487 (43.0)	3522 (43.2)	3367 (41.9)	
High	9519 (23.5)	1521 (18.9)	1732 (21.2)	1829 (22.5)	2066 (25.3)	2371 (29.5)	
BMI, kg/m²	24.1±4.8	24.5±4.9	24.1±4.7	23.9±4.5	23.9±4.6	24.3±5.2	<.0001
Weight status							<.0001
Normal-weight (BMI<25kg/m^2)	31,769 (67.8)	5929 (63.4)	6406 (68.2)	6558 (69.9)	6550 (69.6)	6326 (67.9)	
Overweight (25≤BMI<30kg/m^2)	9975 (21.3)	2270 (24.3)	2002 (21.3)	1971 (21.0)	1924 (20.4)	1808 (19.4)	
Obese (BMI≥30kg/m^2)	5120 (10.9)	1150 (12.3)	987 (10.5)	858 (9.1)	941 (10.0)	1184 (12.7)	
Height, cm	163.4±6.1	162.8±6.0	163.0±6.0	163.4±6.1	163.7±6.0	164.2±6.1	<.0001
Number of biological children	1.8±1.2	1.8±1.2	1.8±1.1	1.8±1.1	1.8±1.1	1.8±1.2	<.0001
Family history of cancer (yes)	21158 (45.2)	4446 (47.6)	4393 (46.8)	4288 (45.7)	4185 (44.4)	3846 (41.3)	0.9
Menopausal status							0.5
Pre-menopause	23940 (51.1)	3767 (40.3)	4078 (43.4)	4637 (49.4)	5296 (56.2)	6162 (66.1)	
Perimenopause	3997 (8.5)	807 (8.6)	871 (9.3)	807 (8.6)	795 (8.4)	717 (7.7)	
Post-menopause	18927 (40.4)	4775 (51.1)	4446 (47.3)	3943 (42.0)	3324 (35.3)	2439 (26.2)	
Hormonal treatment for menopause use (yes)[d]	4068 (17.7)	1025(18.4)	978 (18.4)	806 (17.0)	732 (17.8)	527 (16.7)	0.04
Energy intake without alcohol, kcal/d	1710±385	1510±331	1648±334	1721±344	1792±370	1882±429	<.0001
Alcohol intake, g/d	6.5±9.1	4.5±7.7	5.9±8.4	6.8±9.0	7.4±9.5	7.9±10.5	<.0001

For peer review only

Lipid intake, g/d	76.2±22.6	58.5±17.4	70.2±17.4	76.9±18.2	83.0±19.7	92.4±23.9	<.0001
Protein intake, g/d	76.0±18.3	78.1±20.6	76.3±17.7	75.8±17.1	75.6±17.4	74.4±18.4	<.0001
Carbohydrate intake, g/d	94.9±31.9	88.0±34.5	94.6±31.5	96.3±30.4	98.1±30.9	97.3±31.0	<.0001
Fiber intake, g/d	19.4±6.5	22.4±7.9	20.5±6.2	19.4±5.7	18.4±5.4	16.6±5.2	<.0001
Fruit intake, g/d	247.8±152.3	303.9±185.3	271.1±145.8	249.6±138.2	226.7±130.9	187.4±128.0	<.0001
Vegetable intake, g/d	236.6±113.3	295.8±138.5	255.6±105.7	234.9±98.7	215.2±92.1	181.4±91.0	<.0001
Legume intake, g/d	11.6±21.4	16.8±29.4	12.7±21.3	11.0±19.1	9.7±17.9	7.6±15.7	<.0001
Red meat intake, g/d	39.0±34.1	38.6±37.8	39.6±33.3	40.0±33.8	40.0±32.8	39.7±31.0	<.0001
Processed meat intake, g/d	28.4±25.7	19.4±21.9	23.6±21.8	27.3±22.7	32.1±24.8	37.0±32.6	<.0001
Poultry intake, g/d	24.8±27.6	31.3±34.6	26.0±27.6	24.1±25.3	22.6±24.1	20.1±23.5	<.0001
Fish (including sea product) intake, g/d	40.7±37.6	52.2±45.2	44.8±37.7	40.2±35.3	36.0±32.8	30.5±31.7	<.0001
Dairy intake, g/d	162.8±145.3	217.2±176.1	178.1±145.8	158.8±134.8	142.1±125.2	117.9±117.8	<.0001

[a] P value for the comparison between quintiles of FSA-NPS DI, by χ^2 tests from age-adjusted ordinal polytomous logistic regressions

[b] Occasional smokers smoke less than once a day

[c] Data available for 40,536 women

[d] Among women in peri- or post-menopause (n=22,924)

For peer review only

Table 2 Associations between the FSA-NPS DI and breast cancer risk, from multivariable Cox proportional hazards models, NutriNet-Santé Cohort, France, 2009-2015

		Age-adjusted model			Multivariable-adjusted model[a]		
FSA-NPS DI	N for cases/ non-cases	HR	95%CI	P-trend	HR	95%CI	P-trend
Overall							
Continuous score	555/46,309	1.07	1.03, 1.11	0.001	1.06	1.02, 1.11	0.005
Quintiles[b]				0.0004			0.002
Q1	82/9,267	1.00	(ref)		1.00	(ref)	
Q2	122/9,273	1.44	1.09, 1.90		1.43	1.08, 1.90	
Q3	117/9,270	1.45	1.09, 1.93		1.43	1.07, 1.91	
Q4	138/9,277	1.83	1.39, 2.40		1.79	1.35, 2.38	
Q5	96/9,222	1.56	1.15, 2.10		1.52	1.11, 2.08	
Premenopausal women[c]							
Continuous score	171/23,483	1.09	1.02, 1.18	0.02	1.09	1.01, 1.18	0.03
Quintiles[b]				0.002			0.004
Q1	12/3,667	1.00	(ref)		1.00	(ref)	
Q2	28/3,982	1.96	0.99, 3.85		1.92	0.97, 3.79	
Q3	31/4,558	1.94	0.99, 3.78		1.89	0.96, 3.71	
Q4	52/5,204	2.88	1.53, 5.39		2.76	1.45, 5.26	
Q5	48/6,072	2.52	1.34, 4.76		2.46	1.27, 4.75	
Postmenopausal women[c]							
Continuous score	384/27,188	1.06	1.01, 1.11	0.02	1.05	1.00, 1.11	0.06
Quintiles[b]				0.03			0.09
Q1	70/6,416	1.00	(ref)		1.00	(ref)	
Q2	94/6,173	1.35	0.99, 1.84		1.36	0.99, 1.86	
Q3	86/5,578	1.38	1.01, 1.90		1.37	0.99, 1.89	
Q4	86/5,028	1.60	1.17, 2.20		1.57	1.13, 2.18	
Q5	48/3,993	1.30	0.90, 1.88		1.25	0.85, 1.84	

[a] Models were adjusted for age (time-scale), BMI (kg/m², continuous), height (cm, continuous), physical activity (high, moderate, low), smoking status (never smokers, former smokers, occasional smokers, smokers), numbers of dietary records (continuous), alcohol intake (g/d, continuous), energy intake (without alcohol, g/d, continuous), family history of cancer (yes/no), educational level (<high-school degree, <2 years after high-school degree, ≥2 years after high-school degree), number of biological children (continuous), menopausal status at baseline (pre-menopause, perimenopause, post-menopause), hormonal treatment for menopause (postmenopausal women, yes/no) and oral contraception use (premenopausal women, yes/no).

[b] Cut-offs for quintiles of the FSA-NPS DI were 4.1/5.4/6.5/7.7

[c] P for interaction between the FSA-NPS DI and menopausal status=0.06

For peer review only

Supplemental file 1 FSA NPS score computation at food/beverage level

Points are allocated according to the nutrient content for 100g of foods or beverages.
Points are allocated for 'Negative' nutrients (A points) and can be balanced according to 'Positive' nutrients (C points).

A points
Total A points = (points for energy) + (points for saturated fat) + (points for total sugar) + (points for sodium)

Points	Energy (kJ)	Saturated Fat (g)	Total Sugars (g)	Sodium (mg)
0	≤ 335	≤ 1	≤ 4.5	≤ 90
1	> 335	> 1	> 4.5	> 90
2	> 670	> 2	> 9	> 180
3	> 1005	> 3	> 13.5	> 270
4	> 1340	> 4	> 18	> 360
5	> 1675	> 5	> 22.5	> 450
6	> 2010	> 6	> 27	> 540
7	> 2345	> 7	> 31	> 630
8	> 2680	> 8	> 36	> 720
9	> 3015	> 9	> 40	> 810
10	> 3350	> 10	> 45	> 900

C points
Total C points = (points for fruits and vegetables) + (points for fibers) + (points for proteins)

Points	Fruits, Vegetables (%)	Fiber (g) *	Protein (g)
0	≤ 40	≤ 0.7	≤ 1.6
1	> 40	> 0.7	> 1.6
2	> 60	> 1.4	> 3.2
3	-	> 2.1	> 4.8
4	-	> 2.8	> 6.4
5	> 80	> 3.5	> 8.0

*FSA score allocates different thresholds for fibers, depending on the measurement method used. We used NSP cut-offs to compute fibers score.

Overall score computation

- If Total A points <11, then FSA score =Total A points – Total C points
- If Total A points ≥11,
  - If points for fruits and vegetables =5, then FSA score =Total A points – Total C points
  - Else if points for fruits and vegetables <5, then FSA score = Total A points – (points for fiber + points for fruits and vegetables).

For 100g of a given food, the percentage of fruits and vegetables is obtained by summing up the amount (in grams) of all fruits, legumes and vegetables (including oleaginous fruits, dried fruits and olives) contained in this food.

For peer review only

Exceptions were made for cheese, fat, and drinks to better rank them according to their nutrient profile, consistently with nutritional recommendations:

Score computation for cheese

For cheese, the score takes in account the protein content, whether the A score reaches 11 or not, i.e.: FSA score =Total A points – Total C points

Score computation for fat

For fat, the grid for point attribution is based on the percentage of saturated fat among total lipids and has a six-point homogenous ascending step, as shown thereafter:

Points	**Saturated Fat/Lipids (%)**
0	< 10
1	< 16
2	< 22
3	< 28
4	< 34
5	< 40
6	< 46
7	< 52
8	< 58
9	< 64
10	≥ 64

Score computation for drinks

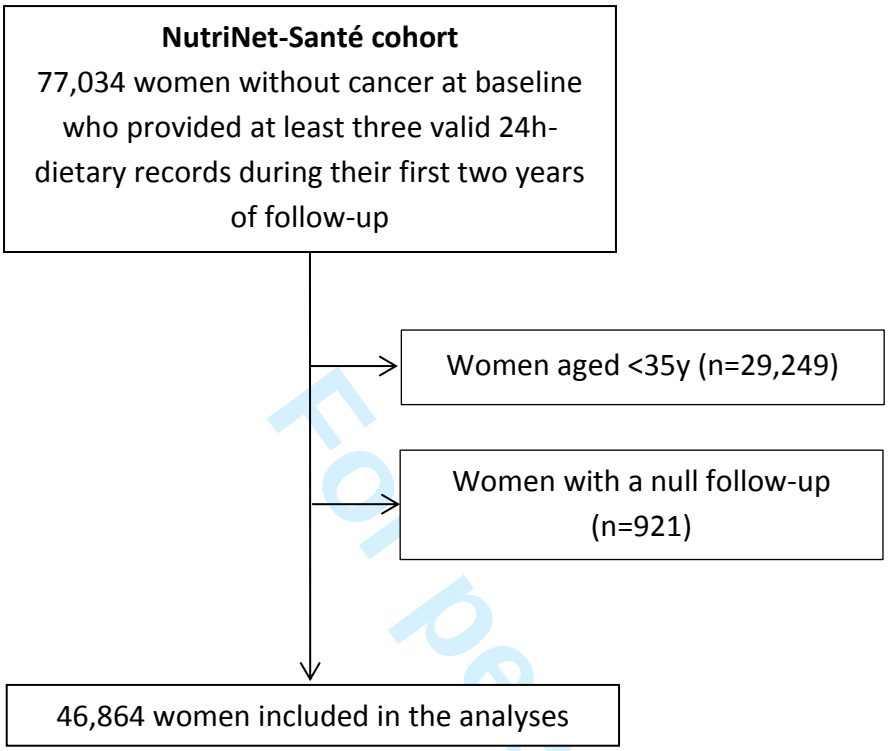
For drinks, the grids for point attribution regarding energy, total sugars and fruits and vegetables (%) were modified. The attribution of points for sugars takes into account the presence of sweeteners, in which case the grid maintains the total sugar score to 1 (instead of 0).

Points	**Energy (kJ)**	**Total Sugar (g)**	**Fruits, Vegetables (%)**
0	≤ 0	≤ 0	< 40
1	≤ 30	≤ 1.5	
2	≤ 60	≤ 3	> 40
3	≤ 90	≤ 4.5	
4	≤ 120	≤ 6	> 60
5	≤ 150	≤ 7.5	
6	≤ 180	≤ 9	
7	≤ 210	≤ 10.5	
8	≤ 240	≤ 12	
9	≤ 270	≤ 13.5	
10	> 270	> 13.5	> 80

Milk and vegetable milk are not concerned by this exception. Their scores are computed using the overall score computation system.

For peer review only

Supplementary file 2 Participants' flowchart



For peer review only

Supplementary file 3 Associations between the FSA-NPS DI and breast cancer risk by hormonal receptor status of the tumors, from multivariable Cox proportional hazards models, NutriNet-Santé Cohort, France, 2009-2015

		Multivariable-adjusted model[a]		
FSA-NPS DI	N for cases/ non-cases	HR	95%CI	P-trend
ER+	307/46,557			
Continuous score		1.05	0.99, 1.12	0.07
Quintiles[b]				0.07
Q1		1.00	(ref)	
Q2		1.59	1.08, 2.34	
Q3		1.77	1.20, 2.60	
Q4		1.59	1.06, 2.39	
Q5		1.60	1.04, 2.46	
ER-	54/46,810			
Continuous score		1.07	0.94, 1.23	0.3
Quintiles[b]				0.1
Q1		1.00	(ref)	
Q2		1.70	0.66, 4.37	
Q3		0.74	0.23, 2.37	
Q4		3.24	1.32, 7.95	
Q5		1.54	0.54, 4.42	
PR+	260/46,604			
Continuous score		1.04	0.97, 1.11	0.2
Quintiles[b]				0.3
Q1		1.00	(ref)	
Q2		1.73	1.14, 2.62	
Q3		1.64	1.07, 2.52	
Q4		1.62	1.04, 2.51	
Q5		1.46	0.91, 2.35	
PR-	102/46,762			
Continuous score		1.11	1.01, 1.23	0.04
Quintiles[b]				0.01
Q1		1.00	(ref)	
Q2		1.29	0.64, 2.62	
Q3		1.68	0.84, 3.34	
Q4		2.46	1.26, 4.79	
Q5		1.99	0.95, 4.17	
ER+/PR+				
Continuous score	258/46,606	1.04	0.97, 1.11	0.3
Quintiles[b]				0.3
Q1		1.00	(ref)	
Q2		1.72	1.14, 2.62	
Q3		1.58	1.03, 2.42	
Q4		1.61	1.04, 2.49	
Q5		1.45	0.90 ,2.33	
ER-/PR-	53/46,811			
Continuous score		1.07	0.93, 1.23	0.3
Quintiles[b]				0.1
Q1		1.00	(ref)	
Q2		1.68	0.65, 4.32	

For peer review only

Q3		0.58	0.17, 2.01	
Q4		3.16	1.29, 7.77	
Q5		1.50	0.52, 4.32	
ER+/PR-	49/46,815			
Continuous score		1.16	1.00, 1.35	0.047
Quintiles[b]				0.03
Q1		1.00	(ref)	
Q2		0.88	0.29, 2.66	
Q3		2.89	1.18, 7.11	
Q4		1.53	0.53, 4.38	
Q5		2.69	0.96, 7.57	

[a] Models were adjusted for age (time-scale), BMI (kg/m², continuous), height (cm, continuous), physical activity (high, moderate, low), smoking status (never smokers, former smokers, occasional smokers, smokers), numbers of dietary records (continuous), alcohol intake (g/d, continuous), energy intake (without alcohol, g/d, continuous), family history of cancer (yes/no), educational level (<high-school degree, <2 years after high-school degree, ≥2 years after high-school degree), number of biological children (continuous), menopausal status at baseline (pre-menopause, perimenopause, post-menopause), hormonal treatment for menopause (postmenopausal women, yes/no) and oral contraception use (premenopausal women, yes/no).
[b] Cut-offs for quintiles of the FSA-NPS DI were 4.1/5.4/6.5/7.7

For peer review only

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cohort studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1; 2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-7
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	6-8
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	n.a
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-8
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6-8
Bias	9	Describe any efforts to address potential sources of bias	7-9
Study size	10	Explain how the study size was arrived at	8
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8-9
		(*b*) Describe any methods used to examine subgroups and interactions	8-9
		(*c*) Explain how missing data were addressed	9
		(*d*) If applicable, explain how loss to follow-up was addressed	8
		(*e*) Describe any sensitivity analyses	9
Results			

For peer review only

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	8
		(b) Give reasons for non-participation at each stage	8
		(c) Consider use of a flow diagram	Supplementary file 2
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	10; Table 1
		(b) Indicate number of participants with missing data for each variable of interest	9
		(c) Summarise follow-up time (eg, average and total amount)	10
Outcome data	15*	Report numbers of outcome events or summary measures over time	10
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	10; Table 2
		(*b*) Report category boundaries when continuous variables were categorized	8; Table 2
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n.a
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	10-11
Discussion			
Key results	18	Summarise key results with reference to study objectives	12;14
Limitations			
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	12-15
Generalisability	21	Discuss the generalisability (external validity) of the study results	12-15
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	16

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

For peer review only