Conservative to disruptive diets for optimizing nutrition, environmental impacts and cost in French adults from the NutriNet-Santé cohort

Louise Seconda, Hélène Fouillet, Jean-François Huneau, Philippe Pointereau, Julia Baudry, Brigitte Langevin, Denis Lairon, Benjamin Allès, Mathilde Touvier, Serge Hercberg, et al.

Conservative to disruptive diets for optimizing nutrition, environmental impacts and cost in French adults from the NutriNet-Santé cohort

Louise Seconda[1,2], Hélène Fouillet[3], Jean-François Huneau[3], Philippe Pointereau[4], Julia Baudry[1], Brigitte Langevin[4], Denis Lairon[5], Benjamin Allès[1], Mathilde Touvier[1], Serge Hercberg[1, 6], François Mariotti[3], and Emmanuelle Kesse-Guyot[1]

[1] Université Sorbonne Paris Nord, Inserm, Inrae, Cnam, Nutritional Epidemiology Research Team (EREN), Epidemiology and Statistics Research Centre – University of Paris (CRESS), Bobigny, France

[2] ADEME (Agence de l'Environnement et de la Maîtrise de l'Energie), 20 avenue du Grésillé, Angers, France

[3] UMR PNCA, AgroParisTech, INRAE, Université Paris Saclay, 16 rue Claude Bernard, Paris, France

[4] Solagro, 75 Voie Toec, Toulouse, France

[5] Aix Marseille Université, INSERM, INRAE, C2VN. Faculté de Médicine de la Timone, 27 boulevard Jean Moulin, Marseille, France

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

Correspondence: j.baudry@eren.smbh.univ-paris13.fr

Equipe de Recherche en Epidémiologie Nutritionnelle (EREN)

SMBH Université Sorbonne Paris Nord, 74 rue Marcel Cachin,

93017 Bobigny, France

Keywords: optimization, environmental impact, sustainable diet,

PubMed indexing: Seconda, Fouillet, Huneau, Pointereau, Baudry, Langevin, Lairon, Allès, Touvier, Hercberg, Mariotti, Kesse-Guyot

Abstract

Improving sustainability of diets requires the identification of diets that meet nutritional requirements of populations, promote health, are within planetary boundaries, affordable and acceptable. Here, we explore the extent to which dimensions of sustainability could be optimally aligned and identify more sustainable dietary solutions, from the most conservative to the most disruptive among 12,166 participants of the NutriNet-Santé cohort. We aim to concomitantly lower environmental impacts (including greenhouse gas emissions, cumulative energy demand and land occupation), and increase organic food consumption, and study departure from observed diet (considered as a proxy for acceptability).

From the most conservative to the most disruptive scenario, optimized diets were gradually, richer in fruits, vegetables, and soya-based products and markedly poorer in animal-based foods and fatty and sweet foods. The contribution of animal protein to total protein intake gradually decreased by 12% to 70% of the observed value. The greenhouse gas emissions from the food production for the diets gradually decreased across scenarios (% of observed values) by 36% to 86%, land occupation for food production by 32% to 78%, and energy demand by 28% to 72%. Our results offer a benchmark of scenarios of graded diet changes against graded sustainability improvements.

Introduction

In high-income countries, rebalancing diets in favour of plant food is considered a major route to improve the sustainability of food systems [1–5]. The EAT-Lancet commission concluded that a drastic reduction of red meat consumption, to less than 28 $g.d^{-1}$, is required to improve the sustainability of diet [4]. High intakes of animal-based food have been identified as contributing to greenhouse gas emissions, a threat to planetary boundaries [5,6] and a risk factor for chronic diseases [1,2,5].

There have been attempts to model future sustainable diets with mathematical diet optimization techniques, taking environmental, nutritional, economic and food habit metrics into account [7–9].

In almost all previous modelling studies, the environmental impacts of diet have been considered as constraints rather than objectives in the optimization model, *e.g.*, by limiting the final level of greenhouse gas emissions required to maintain global warming below 2°C or gradual reduction of the environmental impacts of diet production [7–9]. We are aware of only one such study that tried to minimize three environmental indicators (greenhouse gas emissions (GHGe), water use and land occupation) and monetary cost [10].

Finally, in these modelling studies, the impact of diets on toxicological and ecotoxicological exposure is rarely considered [11]. Differences in environmental impacts of diets composed of conventionally grown foods compared to organic ones have been shown, in particular with respect to soil quality and biodiversity [12–18] and so, the food production method should be introduced in optimization models as an alternative to intensive farming practices.

In general, previous studies have addressed the optimization problem at the level of average diets for the entire population or some of its subgroups [7]. The inter-individual variability in diet composition, however, would enable greater assessment of the robustness of solutions identified; for example, optimized diets could vary according to the proportion of plant-based foods in the initial diet.

Here, we identify and compare the dietary changes needed to achieve a nutritionally adequate and economically acceptable diet with lower diet-related environmental impacts and higher organic food contributions. We used individual-based multi-criteria optimization in a large sample of adult participants. To explore the departures from usual diets that are required, we consider scenarios offering graded levels of suboptimal values for sustainability criteria encompassing nutritional and environmental (using the pReCiPe, a synthetic indicator summarizing three indicators GHGe, land occupation and energy demand) characteristics. Nutritional characteristics of the diets were described using the PANDiet score (a score reflecting the probability to reach nutritional references) [19]. To better identify the required changes for dietary habits, we presented the optimized diets according to the level of plant-foods in the baseline diet.

Results

Individual characteristics

We performed the optimization process on a sample composed of 12,308 participants from the NutriNet-Santé cohort. No solution was found by the model for 142 participants. The final sample, composed of 12,166 participants, were likely to be older, with a higher income, living with a partner and without obesity compared to the sample of participants included in the NutriNet-Santé cohort in 2014 (**Supplemental Table 1**). This population included more often female and highly educated people than the general population [20]. **Table 1** presents the socio-economic and lifestyle characteristics of participants for which optimization succeeded by tertiles of provegetarian score (a score reflecting the preference for plant-based foods without total exclusion of animal food).

We found significant differences between tertiles for most of the characteristics tested, except for the proportion of women and income categories. Participants with higher provegetarian score were more likely to be more highly educated, physically active, non-smokers, and moderate or non-drinkers.

Intermediate optimization steps and the extent of potential improvements

The characteristics of the observed and optimized diets after the steps 0 (closest diet to the observed diet meeting the nutritional needs), 1 (diet inducing the lowest environmental impact while satisfying nutritional and price constraints) and 2 (diet inducing the highest consumption of organic foods while satisfying nutritional and price constraints, for different scenarios of concomitant reduction in environmental impacts) are shown in **Supplemental Figure 1, Supplemental Figure 2** and **Supplemental Table 2** .

After step 0, the monetary cost of the diet meeting nutritional needs was higher than the monetary costs of the observed diet for 2,711 participants (22.2% of the sample). For these 2,711 participants, the maximum monetary cost imposed during the following steps was set to the price obtained in this step 0.

After step 1 aiming to estimate the maximum improvement of the environmental impacts of the diet production (E, based on the pReCiPe) without any consideration of organic food intake (O) and diet departure (D), we observed that E could be reduced by as much as 90%, regardless of the baseline provegetarian score.

After step 2 aiming to estimate the maximal improvement in O, without any consideration of D, we obtained distinct solutions according to considered scenario of concomitant improvement in E (value of the parameter p%, imposing an E improvement of at least p% of its maximal improvement determined in step 1), but with diets being always composed almost exclusively of organic foods regardless of the p% scenario.

Final multi-optimized and acceptable diets

At the end of the optimization process (step 3), we obtained different diet solutions for each of the five considered p% scenarios of combined improvements in E and O, which were constrained to be at least p% of their maximal improvements determined in step 2. The changes in nutrient adequacy (PANDiet), monetary cost, energy density, organic food contribution (O in % of dietary intake), environmental impact (E based on pReCiPe) and animal protein contribution (% of animal in total protein intake) from the observed diet to the optimized diets issued from the five scenarios are presented in **Figure 1** by provegetarian score tertile, and these data are further detailed in **Supplemental Table 3**. The mean population values for the relative variations between optimized and observed diets in the environmental and nutritional indicators are summarized in **Table 2**. From the most conservative (p=25%) to the most disruptive (p=90%) scenario, we observed gradual improvements towards environmentally-friendlier and nutritionally adequate diets, while the monetary cost varied little. Among scenarios of progressive disruption, the environmental impacts (pReCiPe) thus gradually decreased by 33% to 80% while the nutritional adequacy (PANDiet) gradually increased from 16% to 28% of the initial observed values. The adequacy probabilities for the main nutrients (PANDiet subscores) are further detailed in the **Supplemental Table 4**. It is noteworthy that, as expected, most probabilities were close to 1, except for a few whose reference values differed from the constraint being set in **Table 3**. For most indicators, the more conservative (less disruptive) the scenario, the greater the differences between tertiles of provegetarian score.

As for the dietary patterns represented in **Figure 2**, the contributions of fruit, vegetables (in particular orange vegetables), starchy foods and soya progressively increased from the most conservative to the most disruptive scenarios, whereas the consumption of meat, dairy products, eggs, mixed dishes, fatty and sweetened or salted foods progressively decreased. The contribution of nuts and legumes increased in the most conservative scenario (25%) but

decreased thereafter in the more disruptive scenarios (and notably from p=70%) in favour of further increases in fruit, vegetables and soya. The differences in the structure of diets across tertiles gradually decreased as p increased and were only minor in the most disruptive scenario (p=90%). Finally, changes in food group consumption over scenarios were similar across the different tertiles of provegetarian score, except for fish, whose consumption increased over scenarios only for consumers of fish in the third tertile.
Consumption data (g/d) by food groups and scenario for observed, step 2 and step 3 diets per tertile of Provegetarian score are presented in **Supplemental Table 5.**

Tensions between environmental impacts and organic consumption

Figure 3 illustrates the variations in environmental impact (pReCiPe) and organic food contribution (%Org) through the different scenarios. From the observed diet to the most disruptive scenario, as pReCiPe progressively decreased, %Org progressively increased until reaching an inflection point for a pReCiPe of ~0.3, from which %Org stabilized around 95% or even slightly decreased. This inflection point showed a conflict between further reducing pReCiPe (below 0.3) and further increasing %Org (above 95%) at fixed monetary cost once a very low pReCiPe has been achieved.

Discussion

This diet optimization study conducted at individual level in a large French sample of adults has identified affordable and nutritionally adequate diets with reduced environmental impacts (pReCiPe) and increased organic food content (%Org), and we graded those diets against thresholds of improvements in these sustainability criteria.
From conservative to disruptive scenarios, the changes in food group consumption were progressive. By improving the sustainability (encompassing nutritional, environmental, and economic characteristics) of diets, the progressive substitution of animal products by plant products observed in our work was in line with the results of other optimization studies in high income countries [9]. It would be interesting to compare our solutions with regard to the context and evolution of agricultural sectors in countries like France. For example, our solutions to reduce environmental impacts are characterised by an increase in soya-based products. In the case of imported soya, the transportation phase may lead to counterproductive effects. However, this was not assessed in our study as we considered impacts at the production level only. In addition, their nutritional profile could be advantageous for optimization. For instance, the risks or benefits of high isoflavone consumption from soya

have not been assessed. In addition, a number of other legume or nut-based foods were not included (not detailed in the FFQ) in the items considered in the optimization. However, given the ongoing trend to eat more plant protein, the food supply of products rich in plant proteins increases and diversifies so that presumably the place of soya-based foods could be less important in the future than it was at the time when the dietary data were collected in the present study (2014). Along the same line, the solutions obtained revealed a high proportion of organic food. This is in line with the latest French national nutrition and health program, in which it is recommended to increase organic food consumption, if possible. Nevertheless, today the French organic food sector does not meet this societal demand [12,21] and this mode of production would need to be expanded to make our prospective scenarios valid. In addition, for the moment, the use of green manure without use for livestock is still very scarce in France. However, the potential development of organic farming and plant-based diets will induce crucial nitrogen needs and this should be considered in the future as the nitrogen from manure or slurry (allowed in organic farming) will be probably insufficient. Finally, ruminant meat consumption was removed from most solutions of disruptive scenarios while dairy product consumption was only reduced compared to observed diets. It would be relevant to assess whether the livestock required to produce this quantity of dairy products is in line with the consumption levels of ruminant meat in each of the scenarios. Some authors have attempted to consider the co-products in their optimization models [22], and indeed reported a lower reduction of the consumption of ruminant meats, but with an extent that strongly depended on the coefficient used to link milk production to meat production. In addition, such a co-product approach makes sense at the population level but not at the individual level. Indeed, it does not seem necessary that each participant's ruminant consumption complies with her/his dairy product consumption as long as the co-products are balanced at the population level. Finally, we observed that fish and seafood consumption decreased on average for the participants of the first two tertiles, notably because of the introduction of constraints. Our objective for adding constraints on fish was to take into account the depletion of fish stocks and the acknowledged toxicological risks related to over-consumption of seafood products. Nevertheless, we can see that with these solutions the nutritional needs for eicosapentoenoic acid and docosahexaenoic acid were not covered by the diets. It would be necessary to consider the introduction of other foods that are sources of these nutrients, such as marine oils.

The optimized diets generated under our scenarios were in line with observed diets identified as more sustainable or proposed by others scenarios in previous studies [1–4]. Indeed, they are more plant-based with few fat and sweet foods. Previous results showed that this kind of diet is associated with improved health conditions [4,23–26]. However, plant-based diets (100%) may have consequences on nutritional status. On the one hand, the bioavailability of some nutrients (iron, zinc, vitamin A) is jeopardized in plant-based diets due, for instance, to phytic acid. On the other hand, meat and meat products play an important role in bioavailable intakes of protein, iron, zinc and vitamin (A and B12) [27]. Consequently, it is important to consider that a shift towards highly plant-based diets may prevent nutritional requirements from being met, although some food synergies may help the absorption [28]. In our study, nutritional quality was assessed by the PANDiet which considers bioavailable zinc and iron in its calculation. Although for some nutrients, quality may be impaired, overall, the PANDiet score is progressively improved in scenarios with increasing plant food content. It however reaches a plateau in the most disruptive scenarios, illustrating that nutritional gain becomes low. Moreover, given that the nutritional constraints are fixed, meat, especially from ruminants, is not totally eliminated. Of note, we have not been able to conduct a specific analysis on the individual amino acids as they are not available in our database.

We should acknowledge some limitations of the present study. Firstly, we conducted our analyses on diets from volunteers involved in a long-term cohort focusing on nutrition and health. Indeed, the NutriNet-Santé participants are more often women, highly educated and exhibit healthier behaviours compared to the French population [20,29]. This may have led to an over-representation of sustainable dietary patterns (rich in vegetables, fruits, whole-grains, legumes and nuts) compared to the general population. Thus, our diet solutions are applicable only to diets similar to those of this sample, and it would be necessary to question our results before their generalization or application even if we have worked on a large number of people with different dietary patterns. Then, environmental indicators were available for organic food only at the production stage. Thus, we used production-related impacts for GHGe, energy and land occupation which may have led to an underestimation of overall impact. This limitation is relative since most of the impacts occur during the production phase [30]. Given that data for organic food are scarce, we were able to consider only three environmental indicators. The three indicators included in the pReCiPe can be considered sufficient for an acceptable representativeness of the overall environmental impact [31]. For some usual foods (*e.g.* tea, etc.), no pReCiPe values were available. We therefore excluded these items in the modelling procedure. In addition, the plant-based meat substitutes that young generation of vegetarians are often fond of [32] may be less environmental friendly [3] than crude plant-based foods and may also depend on the farming system. Moreover, our diet solutions were driven by our methodological approach including definition of the objective functions, constraints, scenarios and process. However, we elected to perform 5 scenarios to propose solutions according to the extent of the changes to be made. We also assumed, as in most diet optimization studies [7], that the most acceptable diets are the closest to the observed diets, but this remains simplistic at a time where the eating habits change very quickly. In addition, we worked with ~200 generic items representing sub-categories, which was a modest sample compared to the variety of French food offer in terms of food items, but this was in the range of numbers used in other optimization studies. Food-consumption data were self-reported and the use of 5-points ordinal scale may have probably led to overestimation of the actual organic food consumption. However, these data derived from a validated food frequency questionnaire that had shown relative validity and reproducibility [33]. Finally, due to vast gaps in the field, we did not distinguish the food composition according to the model of production (conventional *vs.* organic), place of purchase or seasonality.

In conclusion, for the first time, this study identified at the individual level the existence of sustainable diets that notably comply with a large set of environmental metrics (including GHGEs, energy demand and land occupation) and economical and nutritional criterions, and with high organic food content in a large French adult sample. This exploratory study also offers five scenarios that are graded according to the underlying disruption of the food system, on the one hand, and, on the other hand, the efficiency for meeting the environmental challenges.

Our study provides important features concerning the composition of sustainable diets, based on a multi-criteria sustainability approach, under nutritional constraints, and at controlled cost. Our work illustrates the compatibility of various dimensions. This work could contribute to the development of recommendations for sustainable diets. Importantly, the more the impacts are reduced, the more the diets deviate from the initial intakes. All possible levers must be used so as to increase food knowledge of the population regarding sustainable issues. It is, however, important to bear in mind that even small changes on a large scale can lead to significant reductions in impacts.

Materials and Methods

Study population

The study population was composed of adult volunteers from the prospective NutriNet-Santé cohort, which was launched in May 2009 in France [34]. At initiation of the cohort and yearly thereafter, participants completed a baseline set of self-administered questionnaires regarding their dietary intake, socio-economic, anthropometric, health status, and lifestyle characteristics. Participants were also regularly invited to complete complementary questionnaires. This study was conducted in accordance with the Declaration of Helsinki, and all procedures were approved by the Institutional Review Board of the French Institute for Health and Medical Research (IRB Inserm 0000388FWA00005831) and the Commission Nationale de l'Informatique et des Libertés (CNIL 908,450 and 909,216). Electronic informed consent was obtained from all participants. The NutriNet-Santé study was registered in ClinicalTrials.gov (NCT03335644).

Assessment of Dietary Data

From June to December 2014, participants were asked to fill in a self-administered semi-quantitative organic food-frequency questionnaire (Org-FFQ) based on a validated FFQ [33]. The development and sensitivity analyses of the Org-FFQ have been published elsewhere [35]. Briefly, the Org-FFQ collected information on consumption frequencies (yearly, monthly, weekly, and daily units) and portion sizes for 264 items over a year. We estimated the total food intake by multiplying the portion size and the consumption frequency for each item. A 5-point ordinal scale (never, rarely, half of the time, often and always) was added to measure the frequency of organic food consumption for 257 food and beverage items produced under the organic label. We obtained the organic share for the 257 food items by attributing the respective percentages: 0, 25, 50, 75 and 1, to the modalities. We evaluated the share of organic food to the diet by dividing the total organic food intake (g/day) by the total food intake (g/day) excluding water.

We used the NutriNet-Santé food composition database [36] to estimate daily nutrient intake from the diets, regardless of the food production method (organic *vs.* conventional) due to gaps in the field limiting the coverage of the whole diet. In addition, much of the scientific literature on the topic has underlined that some factors such as weather conditions, crop species, soil type, location, livestock nutrition could prevail over organic *vs.* conventional practices [37]. Finally, to assess the nutritional quality of diet, we computed the updated version considering the 2016 ANSES (French National Health Security Agency for food, environment and workplace) guidelines of nutrient-based probability of adequate nutrient intake diet, named PANDiet [19,38]. It is composed of two subscores: an adequacy composed of nutrients for which intake should be above a reference value and a moderation score for items for which the usual intake should not exceed a reference value.

The provegetarian score is a dietary index reflecting the proportion of plant-based food consumed in a diet [39]. It has been previously developed and adapted in the NutriNet-Santé cohort [40,41]. We adjusted the consumption (g/d) of 5 animal food groups (eggs, fish, dairy products, meat and added animal fats) and 7 vegetable food groups (fruit, vegetables, nuts, cereals, potatoes, legumes and olive oil) for the total energy intake by using the residual method, separately for men and women. For each plant component, we allocated 1 to 5 points to energy-adjusted sex-specific quintile values. For animal food groups, the quintile values were reversed (from 5 for the first quintile to 1 for the fifth quintile). We obtained the final provegetarian score (range: 12-60 points) by summing the points of vegetable and animal food groups.

Price database and computation of the monetary cost of diet

We assigned a price to each food item considering the mode of food production (organic *vs.* conventional) as well as and the place of purchase using the 2012 Kantar Worldpanel purchase database and a price database obtained through price collections carried out by members of Bioconsom'acteurs for prices in short supply chains [42].
The main place of food purchase was assessed for 12 food groups gathering 264 items using a secondary questionnaire concomitant with the Org-FFQ. This information was used to assess the individual daily monetary cost of the diet by multiplying the quantities consumed (g/d) by the corresponding item prices (€/g), while accounting for the place of food purchase and the mode of food production.

Environmental impact database and computation of the environmental impacts

The method used to assess the environmental impacts related to raw products as well as the sources of data used have been extensively described in Seconda et al.[43]. Briefly, we considered three environmental indicators measured per kg of each item: the GHGEs, including carbon dioxide, methane and nitrous oxide emissions, measured as kg of CO_2 equivalent by the global warming potential for a 100-year time horizon, the cumulative energy demand in MJ and the land occupation expressed in m^2 and defined as the area required to produce raw agricultural products without considering the duration of land use Environmental indicators were estimated using standardized procedures for LCA computation [44–48]. The DIALECTE database, comprising 2,000 French farms, half of which are organic, was used to calculate the environmental impacts of agricultural raw product at the farm gate. When DIALECTE [49] data were too few or lacking, we used other data sources such as Agribalyse [50] (heated greenhouses products, conventional pork, coffee) and literature results (seafood, imported food such as sugarcane or tea). Environmental impacts were computed for

92 agricultural raw products at the farm gate, 62 came from DIALECTE and 30 from other sources.

The data were compared to the literature [30,51] for validation purpose. Impacts of food products were calculated from impacts of raw products using economic factors when the transformation of the raw product yielded several valuable co-products [52]. We computed daily diet-related GHGEs, cumulative energy demand and land occupation per person by multiplying the reported intake of each food item by their respective environmental impacts considering the mode of food production (conventional *vs.* organic).

We used the pReCiPe, a synthetic score [31,53] to aggregate these three indicators of diet environmental impact into. The ReCiPe system was established to take into account trade-offs and conflicts between environmental indicators and to consider the alignment of midpoint-oriented and endpoint-oriented indicators, using weighing values, as defined by a panel based on European data [54]. Kramer et al. [31] documented that the three indicators, namely GHGEs, primary energy consumption and land occupation, included in the partial ReCiPe (pReCiPe) allow a satisfactory representativeness (about 90%) of the total environmental impact. However, many other relevant indicators [55] also exist. We focused on these three indicators due to lack of data concerning LCA for organic food.

It is defined as:

$$p\text{ReCiPe} = [0.0459 \times \text{GHGEs} + 0.0025 \times \text{CED} + 0.0439 \times \text{LO}]$$

where GHGEs is greenhouse gas emissions, in kgCO2 eq/kg, CED is cumulative energy demand, in MJ/kg and LO is land occupation, in m^2/kg.

We obtained the pReCiPe per day of each individual diet by multiplying the pReCiPe of each food item accounting for the food production method by the daily quantity of food consumed and by summing them up.

Sociodemographic and lifestyle characteristics

Participants filled in validated web-questionnaires collecting data on sociodemographic and lifestyle characteristics [56,57]. We used the data closest to the Org-FFQ completion date for each participant. Sociodemographic and lifestyle characteristics included sex, age (over 18 years), last scholar qualification (<high school diploma, high school diploma, and post-secondary graduate), marital status, household size, monthly income per household unit (<1,200€, between 1,200 and 1,800€, between 1,800 and 2,700€, and >2,700€ per household unit) obtained using the household income per month and the household composition, smoking status (former, occasional, current, or non-smoker), level of physical activity measured by the International Physical Activity questionnaire (IPAQ) [58], and alcohol consumption status (abstainers, moderate drinkers (<14 g alcohol/day), and heavy drinkers).

Optimization process

Optimization functions and constraints

We used individual data about food consumption, place of food purchase, nutritional composition, environmental impacts and prices of items to build a model aiming at optimizing diet according to the three following objectives, while ensuring coverage of the nutritional needs and controlling the monetary cost, 1) to minimize the environmental impact of diet production, 2) to maximize the organic food contribution to the diet and 3) to minimize the total departure from the observed diet (initial condition) and the corresponding modelled diet for maximizing its acceptability.

Two types of variables composed the optimization model: the quantities consumed and the proportions in the organic form for each item. We removed the items for which environmental impacts were missing (N=25, listed in **Supplemental Method**), so that a maximum of 239 items were included. We distinguished three types of items: the initially consumed items, the non-consumed items that can be added to the diet and those that cannot be added to the diet for health or cultural reasons (as meat or sweet food). The first two types of items were included in the optimization model; thus, the number of items in the model depended on each participant.

Mathematically, the objective functions for the environmental impact (E), organic intake (O) and diet departure (D) were defined for each participant as follows:

$$E = \sum_{i=1}^{239} \left[pReCiPe_{org}(i)*intake(i)*\%_{org}(i) + pReCiPe_{conv}(i)*intake(i)*(1-\%_{org}(i)) \right]$$

$$O = \sum_{i=1}^{235} \left[intake\ (i)*\%_{org}(i) \right]$$

$$D = \sum_{i=1}^{235} \left[\frac{Moy_{obs}\ (i) - Moy_{opt}\ (i)}{SD(i)} \right]^2$$

with i denoted the item (food or beverage), org and conv denoted organic and conventional, respectively, intake(i) and $\%_{org}(i)$ represented the consumed quantity (g) and proportion of organic for the considered item, and $Moy_{obs}(i)$ and $Moy_{opt}(i)$ represented the mean daily ingested quantities of item i in the observed and optimized diets, respectively.

To ensure that optimized diets belong to a conceivable range, we introduced an upper limit for each item, each food category and each food group. The upper limits of the intake are set at the 95th percentile of the distribution of items intakes, food categories or food group by participant categories (men, menopausal women, and non-menopausal women).

For each participant, we set as a constraint that the energy intake in the optimized diet was comprised between 92% and 108% of the individual energy requirement (as assessed with estimates of physical activity levels and basal metabolic rate, using Schofield equations). Moreover, to ensure the nutritional adequacy of the optimized diets, we imposed a set of nutritional constraints pertaining to 26 nutriments as presented in **Table 3**. Alcohol intake in the optimized diet had to be below the minimum of the observed intake and the World Health Organization recommendation of 14 g/d. Finally, in order to take into account exposures to harmful substances through fish consumption, we added two additional constraints, according to the French Agency for Food, Environmental and Occupational Health Safety (ANSES) guidelines. We imposed a total fish consumption of less than 28 g/d and the consumption of fatty fish of less than 14 g/d.

Finally, we imposed an upper limit of the diet monetary cost. For this purpose, we identified the minimal price required to meet nutritional requirements, by minimizing deviations from the observed diet (function D) under nutritional and food constraints to ensure coverage of the nutritional needs. Thereafter, for the resolution of the optimization, the upper limit of the diet cost was set to the maximum between the observed cost and the cost required to ensure nutritional needs.

<u>Hierarchical method to solve the multi-objective problem</u>

The optimization was multi-objective including the three objective functions E, O and D. To solve this multi-objective problem, we applied a hierarchical method, as described by Mausser [59]. This method consists in ranking the objective functions in descending order of importance and beginning with linear functions. Each function is then optimized individually, under the constraints of concomitant improvements in the higher-ranked functions of at least a specified fraction (p%) of their previously identified maximum potential improvements.

The different steps and deliverables of the hierarchical optimization process as set here are presented in **Figure 4**. After a preliminary step to identify the diet monetary price required to meet nutritional requirements (step 0), we first assessed the maximum potential improvement of environmental impact, by minimizing the E function under the aforementioned constraints (step 1). Then (step 2), the potential improvement in organic food consumption was identified by maximizing the O function, under the usual constraints and an additional constraint corresponding to an improvement in E of at least p% of its maximum improvement assessed in the step 1. Then (step 3), we optimized diet to be as close as possible to the observed diet (minimization of the D function), under the usual constraints and additional constraints corresponding to E and O improvements by at least p% of their previously estimated potential combined improvements assessed in step 2. We conducted 5 scenarios of increasing disruption, where p% was set at 25%, 50%, 70%, 80% and 90%. The p% means that in the final step (step 3), the process achieves p% of the maximum improvement. For instance, in the case of p=25%, the scenario allows to achieve 25% to the possible improvement in pReCiPe, 25% of the improvement in organic food consumption while minimizing the deviations from the observed diet.

Optimization Tool

The optimization process was performed using the procedure SAS/OR ® optmodel (version 9.4; SAS Institute, Inc.), with the Activeset algorithm for non-linear optimizations and the option multistart to avoid solutions being only local optimums. The number of starts and iterations for each step were fixed as a compromise to converge towards a solution within a reasonable calculation time. We repeated the steps for failures once by increasing the number of starts and iterations. we removed from the sample the few participants for whom we still had no solution

Data analysis

Sample Selection

We estimated the energy requirement by accounting for the physical activity level and basal metabolic rate computed by Schofield equations [60]. In this study, we selected participants who completed the Org-FFQ, with available data regarding the place of purchase for the monetary cost of the diet assessment, and with no missing covariates. We also removed from the sample the participants whose energy intake/energy requirement ratio was < 0.80 or > 1.20. Finally, the sample is composed of 12,308 participants.

Statistical Analyses

We ranked the participants in three categories according to the tertile-values of the provegetarian score based on observed data. We reported findings globally and across tertiles of provegetarian score, as mean difference in % of the observed values or means and standard error (SE). We performed all statistical analyses using SAS (version 9.4; SAS Institute, Inc.).

Acknowledgements

We thank Oualid Hamza, Christine Boizot-Santai, Louis-Georges Soler and Bioconsom'acteurs' members for price collection and data management. The authors thank Cédric Agaesse (dietitian); Younes Esseddik, Thi Hong Van Duong, Paul Flanzy, Régis

Gatibelza, Jagatjit Mohinder and Aladi Timera (computer scientists); Dr Fabien Szabo de Edelenyi, PhD, Nathalie Arnault, Julien Allegre, and Laurent Bourhis (data-manager/statisticians), and Dr Nathalie Druesne Pecollo, PhD (operational coordinator) for their technical contribution to the NutriNet-Santé study. We thank all the volunteers of the NutriNet-Santé cohort.

Conflict of interests

The authors declare that they have no competing interests.

Authors' contributions

The authors' responsibilities were as follows. JB, DL, BA, MT, SH and EK-G conducted the study; LS, HF, J-F H, FM and EK-G: designed and conducted the research; LS, PP, JB, BL, DL, BA, MT, SH and EK-G: provided essential materials; LS, HF, J-F H, FM, and EK-G: analyzed the data; LS: wrote the paper; EK-G: had primary responsibility for the final content and supervised the research; and all authors: were involved in interpreting the results and editing the manuscript and read and approved the final manuscript.

Transparency statement

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

Data availability statement:

Data can be retrieved from the corresponding author upon reasonable request

Code availability statement:

Code and programs can be retrieved from the corresponding author upon reasonable request.

Funding

The NutriNet-Santé study is funded by French Ministry of Health and Social Affairs, Santé Publique France, Institut National de la Santé et de la Recherche Médicale, Institut National de la Recherche Agronomique, Conservatoire National des Arts et Métiers, and Paris 13 University. The BioNutriNet project was supported by the French National Research Agency (Agence Nationale de la Recherche) in the context of the 2013 Programme de Recherche

Systèmes Alimentaires Durables (ANR-13-ALID-0001). The funders had no role in the study design, data collection, analysis, interpretation of data, preparation of the manuscript, and decision to submit the paper.

Number of tables: 3/**Number of figures:** 4/**Supplemental information:** 5 SI Tables, 2 SI Figure, 1 SI Method

SI Table 1: Comparison between sociodemographic and lifestyle characteristics of NutriNet-Santé population, included participants and the final study sample, 2014

SI Table 2: Characteristics of the observed and optimized diets issued from intermediary steps (0, 1 and 2) of the optimization process according to different scenarios for the total population (N=12,166), NutriNet-Santé 2014

SI Table 3: Characteristics of the observed and optimized diets issued from the 5 scenarios of increasing improvements in pReCiPe and %Org (final step 3 of the optimization process) across tertiles of the provegetarian score (N=12,166), NutriNet-Santé 2014

SI Table 4: Probability of adequacy of nutrient intakes for optimized diets issued from the 5 scenarios of increasing improvements in pReCiPe and %Org (final step 3 of the optimization process), NutriNet-Santé 2014

SI Table 5: Consumption (g/d) by food groups and scenario for observed, step 2 and step 3 diets across tertiles of the provegetarian score

SI Method 1: Observed consumption (g/d) of items not included in the optimization due to missing data for pReCiPe computation

SI Figure 1: Composition of the observed and optimized diets issued from intermediary steps 0 and 1 of the optimization process, NutriNet-Santé Study 2014

SI Figure 2: Structure of the observed and optimized diets issued from the five scenarios of increasing improvements in pReCiPe (step 2 of the optimization process) across tertiles of the provegetarian score, NutriNet-Santé 2014

Figure 1: Sustainable characteristics of observed and optimized diets

Diet quality (PANDiet), price, energy density, organic food contribution (%Org), environmental impact (pReCiPe) and animal protein contribution are presented for the observed and optimized diets issued from the 5 scenarios of increasing improvements in pReCiPe and %Org across tertiles of the provegetarian score in the observed diets (N=12,166), NutriNet-Santé 2014

Figure 2: Structure of observed and optimized diets

Diet composition as share of the diet (in weight) is presented for the observed and optimized diets issued from the 5 scenarios of increasing improvements in pReCiPe and %Org across tertiles of the provegetarian score (N=12,166), NutriNet-Santé 2014

Figure 3: Organic food and environmental impact of observed and optimized diets

Organic food contribution (%Org) according to environmental impact (pReCiPe) is presented for the observed and optimized diets issued from the 5 scenarios of increasing improvements in pReCiPe and %Org (N=12,166), NutriNet-Santé 2014

Figure 4: Optimization process

References

1. Aleksandrowicz, L., Green, R., Joy, E. J. M., Smith, P. & Haines, A. The Impacts of Dietary Change on Greenhouse Gas Emissions, Land Use, Water Use, and Health: A Systematic Review. *PLOS ONE* **11**, e0165797 (2016).
2. Perignon, M., Vieux, F., Soler, L.-G., Masset, G. & Darmon, N. Improving diet sustainability through evolution of food choices: review of epidemiological studies on the environmental impact of diets. *Nutr. Rev.* **75**, 2–17 (2017).

3. Hallström, E., Carlsson-Kanyama, A. & Börjesson, P. Environmental impact of dietary change: a systematic review. *J. Clean. Prod.* **91**, 1–11 (2015).
4. Willett, W. *et al.* Food in the Anthropocene: the EAT–Lancet Commission on healthy diets from sustainable food systems. *The Lancet* (2019) doi:10.1016/S0140-6736(18)31788-4.
5. Clark, M. A., Springmann, M., Hill, J. & Tilman, D. Multiple health and environmental impacts of foods. *Proc. Natl. Acad. Sci.* 201906908 (2019) doi:10.1073/pnas.1906908116.
6. Tilman, D. & Clark, M. Global diets link environmental sustainability and human health. *Nature* **515**, 518–522 (2014).
7. van Dooren, C. A Review of the Use of Linear Programming to Optimize Diets, Nutritiously, Economically and Environmentally. *Front. Nutr.* **5**, (2018).
8. Gazan, R. *et al.* Mathematical Optimization to Explore Tomorrow's Sustainable Diets: A Narrative Review. *Adv. Nutr.* **9**, 602–616 (2018).
9. Wilson, N., Clegharn, C. L., Cobiac, L. J., Mizdrak, A. & Nghiem, N. Achieving Healthy and Sustainable Diets: A Review of the Results of Recent Studies Using Mathematical Optimization. *Adv Nutr* in press.
10. Donati, M. *et al.* Towards a sustainable diet combining economic, environmental and nutritional objectives. *Appetite* **106**, 48–57 (2016).
11. Barré, T. *et al.* Reaching Nutritional Adequacy Does Not Necessarily Increase Exposure to Food Contaminants: Evidence from a Whole-Diet Modeling Approach. *J. Nutr.* **146**, 2149–2157 (2016).
12. Reganold, J. P. & Wachter, J. M. Organic agriculture in the twenty-first century. *Nat. Plants* **2**, 15221 (2016).

13. Gomiero, T., Pimentel, D. & Paoletti, M. G. Environmental Impact of Different Agricultural Management Practices: Conventional vs. Organic Agriculture. *Crit. Rev. Plant Sci.* **30**, 95–124 (2011).
14. Tuomisto, H. L., Hodge, I. D., Riordan, P. & Macdonald, D. W. Does organic farming reduce environmental impacts?--a meta-analysis of European research. *J. Environ. Manage.* **112**, 309–320 (2012).
15. Lynch, D. Environmental impacts of organic agriculture in temperate regions. *CAB Rev. Perspect. Agric. Vet. Sci. Nutr. Nat. Resour.* **7**, (2012).
16. Tuck, S. L. *et al.* Land-use intensity and the effects of organic farming on biodiversity: a hierarchical meta-analysis. *J. Appl. Ecol.* **51**, 746–755 (2014).
17. Henckel, L., Börger, L., Meiss, H., Gaba, S. & Bretagnolle, V. Organic fields sustain weed metacommunity dynamics in farmland landscapes. *Proc. R. Soc. B Biol. Sci.* **282**, 20150002 (2015).
18. Lori, M., Symnaczik, S., Mäder, P., De Deyn, G. & Gattinger, A. Organic farming enhances soil microbial abundance and activity—A meta-analysis and meta-regression. *PLOS ONE* **12**, e0180442 (2017).
19. de Gavelle, E., Huneau, J.-F., Bianchi, C., Verger, E. & Mariotti, F. Protein Adequacy Is Primarily a Matter of Protein Quantity, Not Quality: Modeling an Increase in Plant:Animal Protein Ratio in French Adults. *Nutrients* **9**, 1333 (2017).
20. Andreeva, V. A. *et al.* Comparison of the sociodemographic characteristics of the large NutriNet-Santé e-cohort with French Census data: the issue of volunteer bias revisited. *J. Epidemiol. Community Health* **69**, 893–898 (2015).
21. Muller, A. *et al.* Strategies for feeding the world more sustainably with organic agriculture. *Nat. Commun.* **8**, (2017).

22. Barré, T. *et al.* Integrating nutrient bioavailability and co-production links when identifying sustainable diets: How low should we reduce meat consumption? *PLOS ONE* **13**, e0191767 (2018).
23. Seconda, L. *et al.* Association between sustainable dietary patterns and body weight, overweight, and obesity risk in the NutriNet-Santé prospective cohort. *Am. J. Clin. Nutr.* (2019) doi:10.1093/ajcn/nqz259.
24. Soret, S. *et al.* Climate change mitigation and health effects of varied dietary patterns in real-life settings throughout North America. *Am. J. Clin. Nutr.* **100**, 490S-495S (2014).
25. Cobiac, L. J. & Scarborough, P. Modelling the health co-benefits of sustainable diets in the UK, France, Finland, Italy and Sweden. *Eur. J. Clin. Nutr.* **73**, 624–633 (2019).
26. Milner, J. *et al.* Health effects of adopting low greenhouse gas emission diets in the UK. *BMJ Open* **5**, e007364–e007364 (2015).
27. Platel, K. & Srinivasan, K. Bioavailability of Micronutrients from Plant Foods: An Update. *Crit. Rev. Food Sci. Nutr.* **56**, 1608–1619 (2016).
28. Nair, K. M. & Augustine, L. F. Food synergies for improving bioavailability of micronutrients from plant foods. *Food Chem.* **238**, 180–185 (2018).
29. Andreeva, V. A. *et al.* Comparison of Dietary Intakes Between a Large Online Cohort Study (Etude NutriNet-Santé) and a Nationally Representative Cross-Sectional Study (Etude Nationale Nutrition Santé) in France: Addressing the Issue of Generalizability in E-Epidemiology. *Am. J. Epidemiol.* **184**, 660–669 (2016).
30. Clune, S., Crossi,E. & Verghese, K. Systematic review of greenhouse gas emissions for different fresh food categories. *J Clean Prod* **140 (Part 2)**, 766–783 (2017).
31. Kramer, G. F., Tyszler, M., Veer, P. van't & Blonk, H. Decreasing the overall environmental impact of the Dutch diet: how to find healthy and sustainable diets with limited changes. *Public Health Nutr.* **20**, 1699–1709 (2017).

32. Gehring, J. *et al.* Consumption of Ultra-Processed Foods by Pesco-Vegetarians, Vegetarians, and Vegans: Associations with Duration and Age at Diet Initiation. *J. Nutr.* doi:10.1093/jn/nxaa196.
33. Kesse-Guyot, E., Castetbon, K., Touvier, M., Hercberg, S. & Galan, P. Relative Validity and Reproducibility of a Food Frequency Questionnaire Designed for French Adults. *Ann. Nutr. Metab.* **57**, 153–162 (2010).
34. Hercberg, S. *et al.* The Nutrinet-Santé Study: a web-based prospective study on the relationship between nutrition and health and determinants of dietary patterns and nutritional status. *BMC Public Health* **10**, 242 (2010).
35. Baudry, J. *et al.* Contribution of Organic Food to the Diet in a Large Sample of French Adults (the NutriNet-Santé Cohort Study). *Nutrients* **7**, 8615–8632 (2015).
36. Nutrinet-Santé, E. Table de composition des aliments de l'étude Nutrinet-Santé. *Paris Econ.* (2013).
37. Gomiero, T. Food quality assessment in organic vs. conventional agricultural produce: Findings and issues. *Appl. Soil Ecol.* **123**, 714–728 (2018).
38. Verger, E. O., Mariotti, F., Holmes, B. A., Paineau, D. & Huneau, J.-F. Evaluation of a Diet Quality Index Based on the Probability of Adequate Nutrient Intake (PANDiet) Using National French and US Dietary Surveys. *PLOS ONE* **7**, e42155 (2012).
39. Martinez-Gonzalez, M. A. *et al.* A provegetarian food pattern and reduction in total mortality in the Prevencion con Dieta Mediterranea (PREDIMED) study. *Am. J. Clin. Nutr.* **100**, 320S-328S (2014).
40. Colombet, Z. *et al.* Individual characteristics associated with changes in the contribution of plant foods to dietary intake in a French prospective cohort. *Eur. J. Nutr.* (2018) doi:10.1007/s00394-018-1752-8.

41. Lacour, C. *et al.* Environmental Impacts of Plant-Based Diets: How Does Organic Food Consumption Contribute to Environmental Sustainability? *Front. Nutr.* **5**, (2018).
42. Seconda, L. *et al.* Assessment of the Sustainability of the Mediterranean Diet Combined with Organic Food Consumption: An Individual Behaviour Approach. *Nutrients* **9**, 61 (2017).
43. Seconda, L. *et al.* Comparing nutritional, economic, and environmental performances of diets according to their levels of greenhouse gas emissions. *Clim. Change* (2018) doi:10.1007/s10584-018-2195-1.
44. ISO 14040:2006. *ISO* https://www.iso.org/cms/render/live/fr/sites/isoorg/contents/data/standard/03/74/37456.html.
45. Audsley, E., European Commission & Directorate-General for Agriculture. *Harmonisation of environmental life cycle assessment for agriculture: final report, Concerted Action AIR3-CT94-2028*. (CE DG VI-Centre de documentation, 2003).
46. Cowell, S. J. & Clift, R. Impact assessment for LCAs involving agricultural production. *Int. J. Life Cycle Assess.* **2**, 99–103 (1997).
47. Nemecek, T. & Kägi, T. Life Cycle Inventories of Agricultural Production Systems. **Ecoinvent v2. 0**, 361 (2007).
48. Joint Research Center, F. International Reference Life Cycle Data System (ILCD) Handbook - General guide for Life Cycle Assessment - Provisions and Action Steps. *EU Science Hub - European Commission* https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/international-reference-life-cycle-data-system-ilcd-handbook-general-guide-life-cycle (2010).

49. Pointereau, P., Langevin, B., Gimaret, M. & others. DIALECTE, a comprehensive and quick tool to assess the agro-environmental performance of farms. in *10th European IFSA Symposium, July* 1–4 (2012).
50. Colomb, V. *et al.* AGRIBALYSE: the French public LCI database for agricultural products. (2015).
51. Koch, P. *et al.* Methodological guidelines for LCA of French agricultural products. *Life Cycle Management Conference (LCM2011), Berlin, Allemagne, August 28-31, 2011* https://agritrop.cirad.fr/561176/ (2011).
52. European Commission & Joint Research Centre. *ILCD handbook: general guide for life cycle assessment : detailed guidance*. (Publications Office of the European Union, 2010).
53. van Dooren, C., Aiking, H. & Vellinga, P. In search of indicators to assess the environmental impact of diets. *Int. J. Life Cycle Assess.* **23**, 1297–1314 (2018).
54. Goedkoop, M. *et al.* ReCiPe 2008: A Life Cycle Impact Assessment Method Which Comprises Harmonised Category Indicators at the Midpoint and the Endpoint Level. 132 (2013).
55. Vanham, D. *et al.* Environmental footprint family to address local to planetary sustainability and deliver on the SDGs. *Sci. Total Environ.* **693**, 133642 (2019).
56. Vergnaud, A.-C. *et al.* Agreement between web-based and paper versions of a socio-demographic questionnaire in the NutriNet-Santé study. *Int. J. Public Health* **56**, 407–417 (2011).
57. Touvier, M. *et al.* Comparison between web-based and paper versions of a self-administered anthropometric questionnaire. *Eur.J.Epidemiol.* **25**, 287–296 (2010).
58. Hagströmer, M., Oja, P. & Sjöström, M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. *Public Health Nutr.* **9**, (2006).

59. Mausser, H. Normalization and other topics in multi-objective optimization. in (2006).

60. Schofield, W. N. Predicting basal metabolic rate, new standards and review of previous work. *Hum. Nutr. Clin. Nutr.* **39 Suppl 1**, 5–41 (1985).

Figure 1: Sustainable characteristics of observed and optimized diets

Diet quality (PANDiet), price, energy density, organic food contribution (%Org), environmental impact (pReCiPe) and animal protein contribution are presented for the observed and optimized diets issued from the 5 scenarios of increasing improvements in pReCiPe and %Org across tertiles of the provegetarian score in the observed diets (N=12,166), NutriNet-Santé 2014

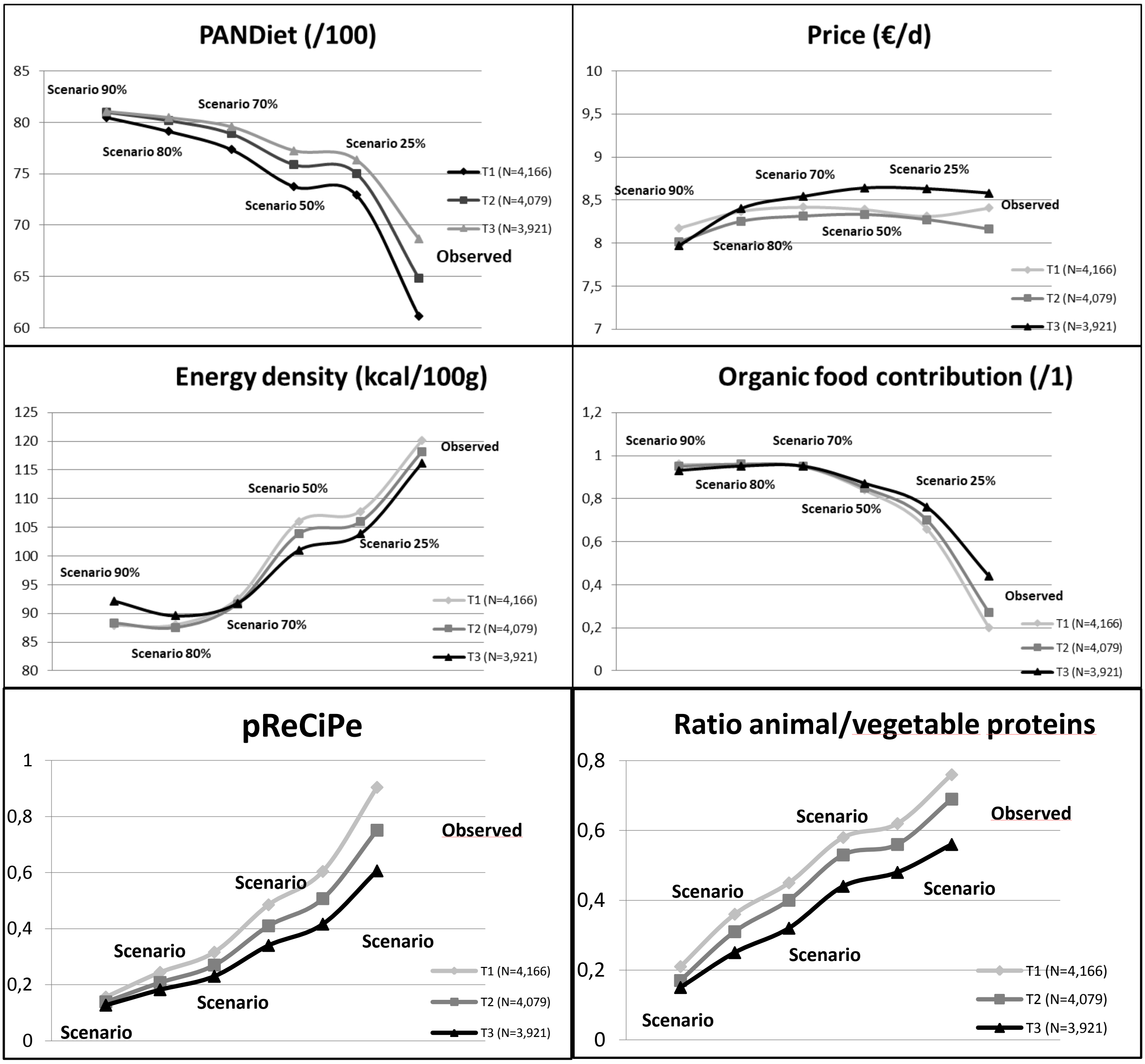

Figure 2: Structure of observed and optimized diets

Diet composition as share of the diet (in weight) is presented for the observed and optimized diets issued from the 5 scenarios of increasing improvements in pReCiPe and %Org across tertiles of the provegetarian score (N=12,166), NutriNet-Santé 2014

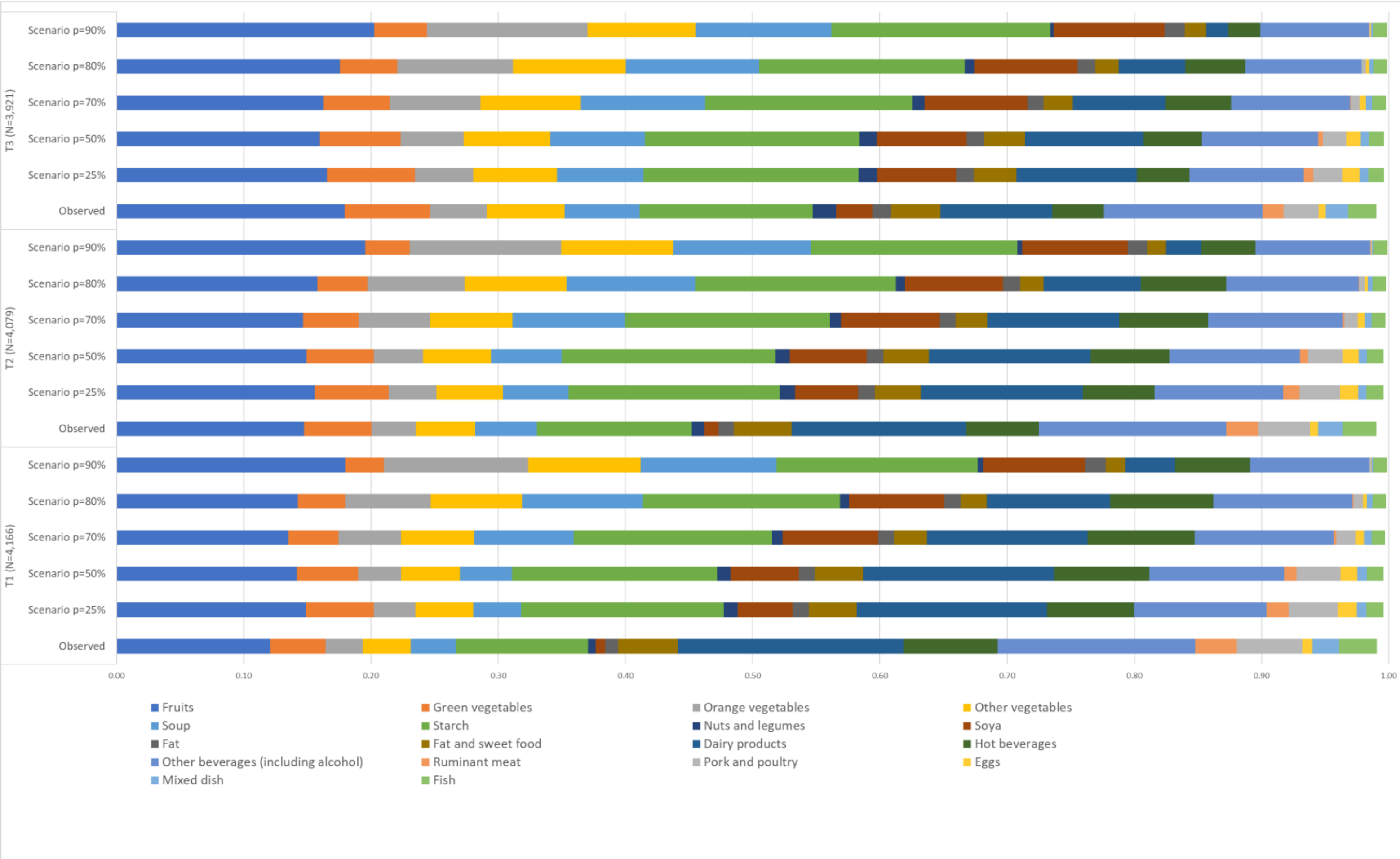

Figure 3: Organic food and environmental impact of observed and optimized diets

Organic food contribution (%Org) according to environmental impact (pReCiPe) is presented for the observed and optimized diets issued from the 5 scenarios of increasing improvements in pReCiPe and %Org (N=12,166), NutriNet-Santé 2014

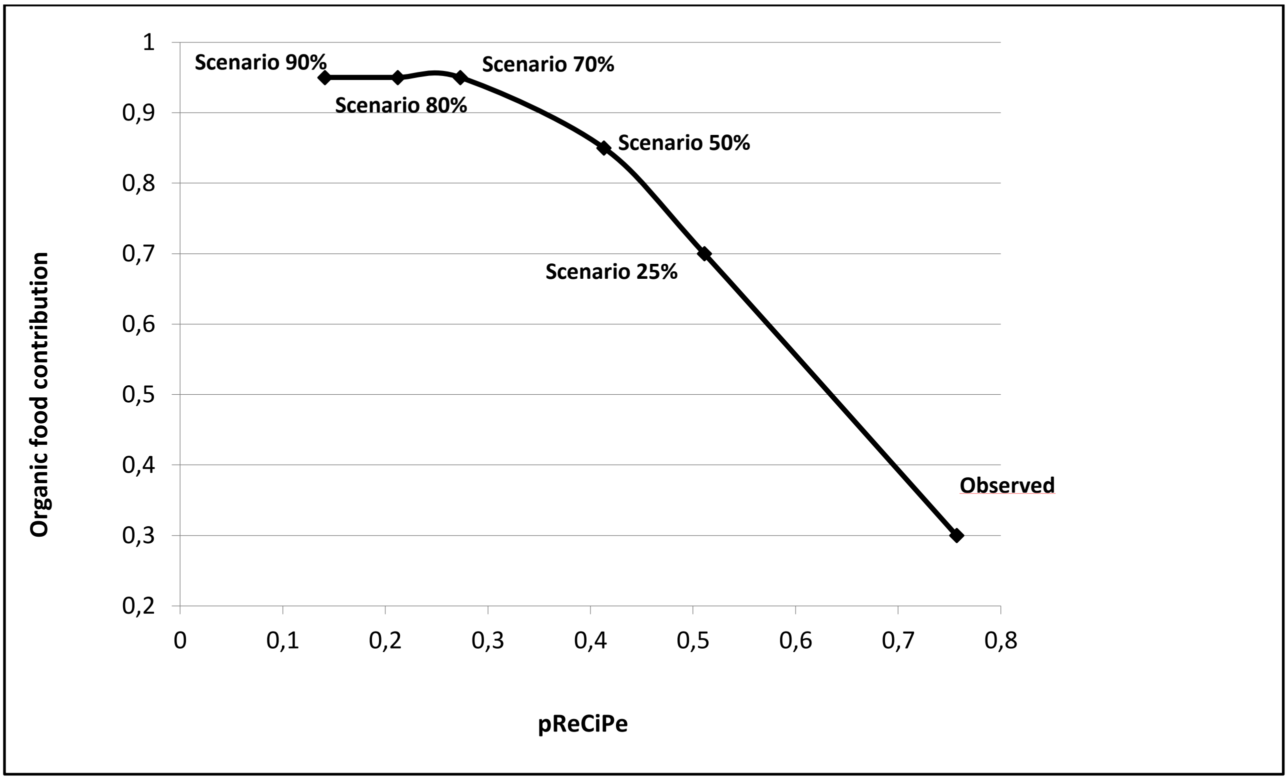

Figure 4: Optimization process