Identifying the most Efficient Detailed Trajectories toward Healthy Diets - A Graph-Based Analysis

Elie Perraud, Juhui Wang, Alison Dussiot, Hélène Fouillet, François Mariotti

Identifying the most efficient detailed trajectories toward healthy diets – a graph-based analysis.

Elie Perraud, Juhui Wang, Alison Dussiot, Hélène Fouillet, François Mariotti

Université Paris-Saclay, AgroParisTech, INRAE, UMR PNCA, 91120, Palaiseau, France

All authors report no conflicts of interest.

Address correspondence to FM (email: francois.mariotti@agroparistech.fr).

Abbreviations used: CIQUAL: database of the French Centre d'Information sur la Qualité des Aliments; DALYs: Disability-Adjusted Life Years; GDB: Global Burden of Disease; HR: Health Risk; INCA3: third individual and national study on food consumption survey; TMREL: Theoretical Minimum Risk Exposure Level

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data sharing: Data described in the manuscript, code book, and analytic code will be made available upon request pending application and approval.

Number of tables: 2

Number of figures: 3

Supplemental material: 12

Abstract

Background: Much effort has been devoted to defining healthy diets, which could lower the burden of disease and provide targets for populations. However, these target diets are far removed from current diets, so at best, the population is expected to move slowly along a trajectory.

Objective: Our aim was to characterize the different possible trajectories toward a target diet and identify the most efficient one for health in order to point out the first dietary changes being the most urgent to implement.

Methods: Using graph theory, we have developed a new method to represent in a graph all stepwise change trajectories toward a target healthy diet, with trajectories all avoiding risk of nutrient deficiency. Then, we have identified and characterized the trajectory with the highest value for long-term health. Observed male and female average diets are from the French representative survey INCA3, and target diets were set using multicriteria optimization. The best trajectories were found using the Dijkstra algorithm with the Health risk criteria based on epidemiological data.

Results: Within ~2.6M diets in the graphs, we found optimal trajectories that were rather similar for males and females regarding the most efficient changes in the first phase of the pathways. In particular, we found that a one-step increase in the consumption of whole/semi-refined bread (60 g) was the first step in all healthiest trajectories. In males, the subsequent decrease in red meat was immediately preceded by increases in legumes.

Conclusions: We show simple practical dietary changes that can be prioritized along an integral pathway that is the most efficient overall for health when transiting toward a distant healthy diet. We put forward a new method to analyze dietary strategy for public health transition and highlight the first critical steps to prioritize.

Keywords: graph theory, trajectory optimization, diet optimization, goal programming, long-term morbidity and mortality

Introduction

Epidemiological studies about the associations between food group consumption or nutrient and chronic diseases have shown that inadequate diets, too poor in healthy foods/nutrients and too rich in unhealthy foods/nutrients intakes are one of the major causes of disability-adjusted life years (DALYs) worldwide (1). A dietary transition is required and has been advocated by all international institutes (2,3). From a dietary viewpoint and based on the expected decrease in the burden of chronic diseases, changes to western diets notably include lower intake of red meat, sugar-sweetened beverages, and refined grains, and higher intakes of whole grains, fruits and vegetables, nuts and seeds, and other healthy plant sources (4). At the nutrient level, it implies lower intakes of sodium, sugar, and SFA and higher intake of PUFA, EPA+DHA, and fiber (5).

Much effort has been devoted to defining healthy diets that make the final objective for populations. Food pattern modeling has identified many healthy diets, notably using optimization models (6,7). By construction, such diets are nutritionally adequate and associated with a lower risk of chronic disease. Although often including limitations of the distance to observed diet, to make feasible objectives, these target diets are much departed from current diets. This can be explained partly by current western diets having nutritional adequacy risks for several nutrients. Thus, modifying the diet to reach the adequacy level of each nutrient and adherence to a healthy dietary pattern is indeed a profound modification of the diet.

A sudden complete change of the diet is not attainable, particularly at the population level, and at best, the population is expected to transition along a trajectory toward a target diet at a rather slow pace (8,9). In reality, the trajectory of change that a population is starting should greatly impact its nutritional health status and, thereby, the cost for society (10). However, there has been little research to characterize such trajectories, particularly to identify those most efficient for health during the long and gradual process toward the target diet. Yet, such data are needed to specify the pathways of changes and point out the first steps within the pathways that are likely to be the most critical and urgent to implement.

Showing the first steps to be taken is also probably an important lever to sustained changes. Significant barriers to change are the lack of information (11) and the efforts needed to change eating routines (12). For example, even if most of the population knows the dietary guidelines, there are applied only by a small percentage (13). The very small changes toward healthy diets in the last years (14,15) call for revising strategies in public health nutrition. Focusing on more precise and effective information may be one of the effective options to be developed. However, for this purpose, we need more information about how small changes, such as an increase or a decrease in portion size, must be coordinated and prioritized to make the most efficient trajectory of changes. Such trajectories should be optimal for long-term health (reducing, at best, the risk of chronic disease) along the trajectory, but also, it should not include a diet with an increased risk of nutrient deficiency. Both the former and the later points are crucial as changes at a population level are slow, and a transitory diet toward the targeted one needs to be safe. Some of the population could stay with this diet for a long time. For example, between 2006 and 2014 in France, daily meat consumption decreased by < 3g, and daily vegetable consumption increased by <9g (14,15).

Many works have aimed to analyze dietary changes and transitions. Two types of methods have been used to identify diet transitions. A first type of study used optimization under constraint as a method. This method enables finding the best diets in the context of an increasingly binding constraint. For example, this method has been used to find the best diets while reducing meat intake (16). However, this method does not produce real trajectories but a succession of milestones. Thus, the trajectory that would consist of going through all these stages is not defined as the most efficient one overall. A second type of study used gradual changes (17,18). According to an incremental algorithm, this method leads to the selection of the best change that could be applied to a diet at each step, but this method is agnostic to the overall trajectory. Accordingly, such trajectories include many local optimums and do not reach any target diet. Thus, no methods have considered a series of simple changes and analyzed complete transition pathways. Dissecting all possible changes in a distant dietary transition requires a systematic approach with a conceptual framework and computer-intensive methods. Here, we report on research that used a new method for analyzing diet trajectories using the graph theory conceptualization. In this study, we identified and characterized the best and safe trajectories toward a target diet using a graph theory and highlight critical simple dietary changes to be prioritized.

Methods

1. Studied population and dietary data

This study used data from the third individual and national study on food consumption survey (INCA3) performed in France between 2014 and 2015 (19,20).

The participant selection process for this study involved a three-stage random sampling design conducted by the National Institute of Statistics and Economic Studies (INSEE), based on the 2011 annual population census. Ineligible individuals included those who were institutionalized, those planning to move out of their homes within the next 2 months, and those who were unable to be interviewed. Further description of the method is in (19).

Participants identified as under-reporters were excluded using the basal metabolic rate as estimated by the Henry equation (21), using the cut-off values recommended by Black (22). Older adults (above 54 years old for female and 64 years old for male (23)) were excluded as their risk of long-term morbidity and mortality differed from younger adults, and we wanted to keep a homogeneity in our population. The final sample contained 1125 adults (Supplemental Figure 1).

Dietary data were collected by professional investigators assisted by dietary software from three unplanned, non-consecutive, 24h dietary recalls spread over a three-week period (two during the week and one at the weekend). Portion sizes were estimated using validated photographs (19). The nutrient content values of the food were extracted from the 2016 database of the French Centre d'Information sur la Qualité des Aliments (CIQUAL) (24). Mixed foods were broken down into ingredients and then gathered into 33 food groups. These food groups were derived from the 45 food groups established by Dussiot et al. (6). The "Fruits" group encompasses fresh fruit, dried fruits, and processed fruits. The "Whole meal and semi-refined bread and bakery products" group comprises whole grain and semi-refined grain bread and bakery products. The "Other refined starches" group includes rice, pasta, other starches, potatoes, and other tubers. The "Processed starch-based products" group consists of starch-based products, sweet/fat processed products, and salt/fat processed starch products. The "Red meat" group encompasses beef, veal, pork, and other meats, and offal. The "Other fish" group includes other fish, mollusks, and crustaceans. The "Fresh dairy products" group includes fresh natural dairy products and fresh sweetened dairy products. The "Animal fats and assimilated fats" group comprises animal fats, assimilated fats, butter, and low-fat butter. The "Sugar-sweetened beverages, including fruit juices" group includes sweetened soda-type drinks and fruit juices. The "Condiments" group encompasses condiments, aromatic herbs, and spices except salt. For a detailed description of each food group, please refer to the Supplementary Material in Dussiot et al (6). The nutrient content of each food group was calculated as the mean nutrient content of food items constituting the group weighted by their mean intake, as previously described (6).

A sex-specific average diet was calculated using the data from the 564 males and the 561 females (Table 1) and used as the initial diet for defining the trajectories.

2. Target diet

The target diets were identified for each sex, according to a multi-criteria optimization approach previously developed by our group (6,25). Using this approach, by construction, target diets ensure adequate nutrient intakes and are found within current consumption limits (5th and 95th percentiles), and they combine minimal long-term health risks and minimal departure from the observed diet. The method is described in full in Supplemental method 1. Target diets are described in Table 1.

3. Graphs theory and construction

A graph is a pair G = (N, E) of nodes (N) and edges (E) (26). In our model, nodes represent daily consumed diets, and edges correspond to a modification of one portion step. If two nodes are linked with an edge, it is possible to transition from one diet to the other.

As explained above, all foods consumed in the INCA3 survey have been classified into 33 food groups according to their nutritional characteristics. A node (diet) is thus a vector with 33 variables, one for each food group. To simplify the graph, the quantity eaten for each food group was rounded to the number of portions of this group.

For each food group, we decided on portion size; it represents the food quantity of a mean food portion eaten in INCA3. The portion sizes were based on the distribution of food consumption and known portioned food. For example, yogurts are portioned in France and weigh 125g. Therefore, this portion size was used for the group "fresh dairy product."

We also adapted the portion sizes, so that food groups with a similar role in the dietary pattern have the same portion size. For example, “poultry” and “red meat” have the same portion size. Supplemental Table 1 presents the portion size for each of the food groups.

4. Method of graph creation and use

Graphs were constructed from both the initial and objective diet nodes. All nodes were created by proximity; this means adding nodes one portion step away from the previous one; until the graph was connected. A connected graph is a graph with all nodes having a path to all other nodes of the graph. Then another extension was added to ensure the graph had no bottleneck due to the way it was created. It was difficult to increase the graph size again as the number of nodes in the graph, especially the one for male, was above a million nodes. In order to reduce the number of nodes, constraints and limits were added during the extension process. The nodes not following these constraints were not added to the graph.

Each node had to respect the following constraints. The number of steps of portion size for each food group should be between 0 and the maximum step of portion size as observed in the INCA3 population. The energy of the diet of the node should be between 80% and 120% of the energy brought by the initial diet. The SecDiet of the node, a score of nutrient security (27) (Supplemental Method 2), should be >99% of the SecDiet of the initial diet. The percentage of macronutrients in each node should be between the minimum and maximum macronutrient intake in the population. The following constraints were added based on French dietary guidelines. In cases where the daily consumption of nuts falls below 40g or the consumption of fruits and vegetables falls below 400g, it is considered below the minimum threshold and cannot be further reduced. Similarly, if the consumption of processed meat exceeds 150g or the consumption of red meat exceeds 500g per day, it is considered above the maximum threshold and cannot be further increased.

5. Trajectories analysis

The optimized trajectory was found using Dijkstra algorithm (28). The trajectories were optimized to minimize the long-term health risk all along the trajectory using the Health Risk (HR) criterion (Supplemental Method 1). The trajectory is thus optimized, so the sum of the HR of all trajectory nodes reaches its minimum. Suboptimal trajectories were also identified.

The HR criterion was set to target the dietary recommendations from the GBD based on epidemiological studies about the associations between the consumption of different food groups and the risk of chronic diseases (1,29). HR represents the combined and normalized distance to the theoretical minimum-risk exposure levels (TMREL) for three unhealthy food groups (red meat, processed meat and sweetened beverages), six healthy food groups (whole grain products, fruits, vegetables, legumes, nuts and seeds, and milk) food groups, four healthy nutrients (fiber, calcium, EPA/DHA, PUFA) and one unhealthy nutrient (sodium), weighted by the importance of reaching each TMREL value in terms of its relative impact on DALYs in France. HR can range from 0 to1, from the situation where the diet meets all the

TMREL values (minimum risk) to the situation where it is farthest from them (maximal risk). A precise description of the HR criterion can be found in Supplemental Method 1.

The percentage of the Population Reference Intake for several nutrients was calculated to describe the trajectories.

All methods were implemented on python 3.8.3 (30). The networkX package version 2.4 was used to implement the graph and the Dijkstra algorithm (31). The community package version 0.15 was used to implement the Louvain algorithm (32).

Results

Table 2 presents the characteristics of the graphs for male and the ones for female. The number of nodes between the initial and target nodes was 24 steps for male and 20 steps for female. Due to this increased number of steps between the initial and target nodes, the graph for male was bigger than for female. The mean number of degrees (*i.e.*, edges per node) representing the density of the graph was slightly higher for male than for female. Supplemental Table 2 describe the lifestyle and sociodemographic characteristics of the included participants and Supplemental Table 3 describe the daily energy and nutrient intake of the sample.

Figure 1 describes the evolution of the health risk (HR) along the trajectory as a percentage of the difference between the worst HR value (of the observed initial diet) and the best HR value (which is that of the final target diet). As the trajectory was optimized to minimize the health risk (HR), the first steps were the most HR-efficient, i.e. those with the highest HR yield, which then progressively decreased with each step. The HR criteria decreased strongly (around -80% of its initial, maximal value) in the first eight steps for both sexes.

Figure 2 and **Figure 3** describe the optimal diet trajectory for male and female respectively. For both male and female with quite distinct target diets, the first step, the most efficient one, was the increase of whole and semi-refined bread, with an increase of one step for female to reach the number of portions in their target diet and an increase of two steps for male (step 1 and step 2), which also reached the number of portions in their target diet.

After these first increases in whole and semi-refined bread, the fruit category was increased. Fruits were increased for woman in step two and for male in step three. The increase in vegetables occurred at step five for female and step seven for male. For male, between the increases in fruit and vegetable there was a substitution of red meat by legumes in three successive steps (two increases in legumes followed by one decrease in red meat). After these substitutions and increases, bread and refined bakery products were decreased for male. For female, between the increases in fruit and vegetable were the milk increase and the bread and refined bakery products decrease. Supplemental Figures 2 and 3 describe the nutrient intakes along the steps of the trajectory.

For each sex, fifty suboptimal trajectories were analyzed and were found similar to the optimal one for the first 17 steps for male and the first 12 steps for female (i.e., only the following less HR-effective steps differed). Supplemental Figures 4 and 5 and Supplemental Tables 4 and 5 describe the other choices available for the different steps of the trajectories.

Figure 1 Decrease in the HR criterion along the diet trajectories from the observed average diet to the set target diet of male (left panel) or female (right panel).

Changes in the HR criterion is shown as a percentage of the difference between the worst HR value (of the observed initial diet) and the best HR value (which is that of the final target diet).

HR: Health Risk

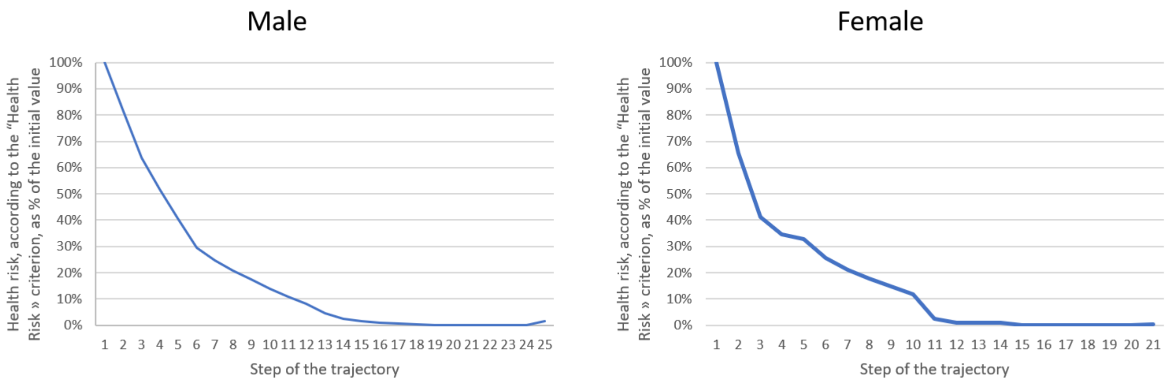

Figure 2 Best trajectory identified in male based on the Health Risk Criterion, which uses estimates from the Global Burden of Disease.

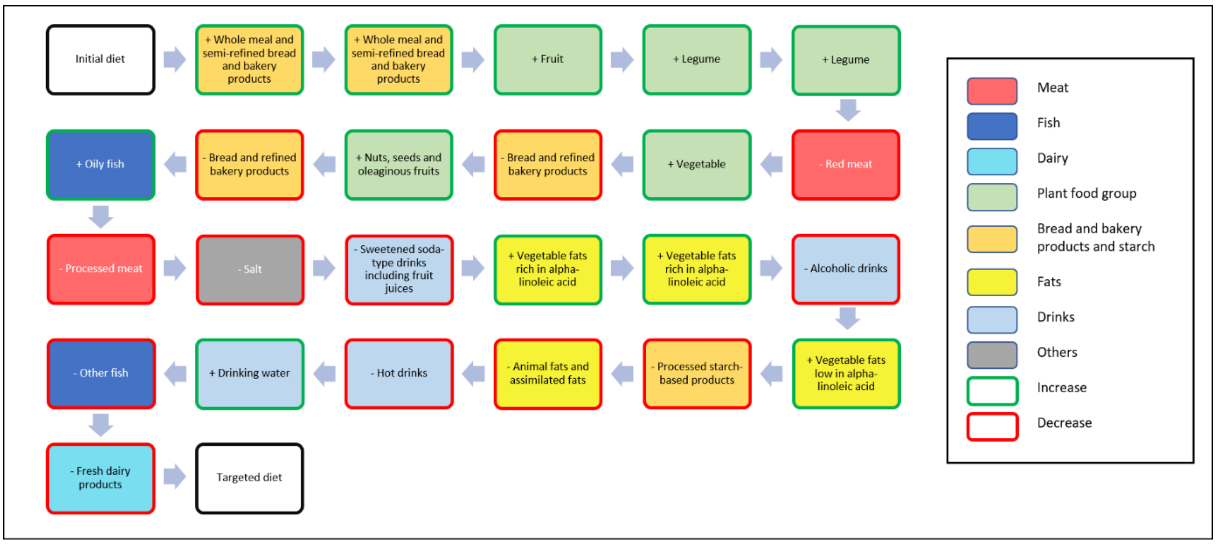

Figure 3 *Best trajectory identified in female based on the Health Risk Criterion, which uses estimates from the Global Burden of Disease.*

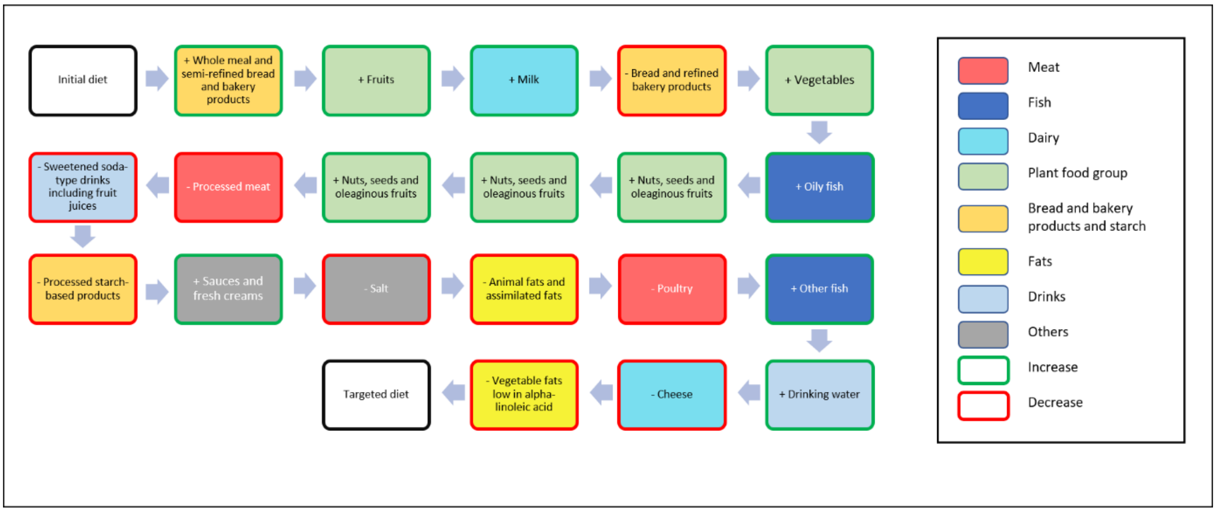

Table 1 Food group consumptions (g/d) by sex in the initial and target diets

	Initial diet		Target diet	
	Male	Female	Male	Woman
Vegetables	176	160	400	368
Fruits	142	124	419	316
Nuts and seeds	3	2	11	13
Bread and refined bakery products	168	115	27	11
Whole meal and semi-refined bread and bakery products	11	15	99	60
Other refined starches	177	121	193	139
Other whole and semi-refined starches	4	4	5	8
Processed starch-based products	26	21	0	13
Legumes	13	6	86	19
Poultry	30	31	46	17
Red Meat	79	42	54	31
Processed meat	50	30	0	0
Oily fish	8	6	26	21
Other fish	27	19	1	24
Eggs and egg-based dishes	14	14	11	12
Milk	84	75	190	126
Fresh dairy products	81	81	28	74
Sweet milky desserts	19	16	0	6
Cheese	49	36	48	13
Animal fats and assimilated fats	11	10	4	0

Vegetable fats rich in α-linoleic acid	0	0	16	5
Vegetable fats low in α-linoleic acid	12	10	15	4
Sauces and fresh creams	35	32	35	100
Sweet products	103	83	103	98
Drinking water	1007	929	878	1108
Sugar-sweetened beverages, including fruit juices	221	208	0	75
Hot drinks	494	507	379	507
Salt	1	1	0	0
Condiments	6	4	13	5
Soups and bouillons	76	79	0	53
Substitutes for animal products	3	5	29	28
Other foods	4	2	4	2
Alcoholic drinks	216	59	0	3

Table 2 Description of the graph and of the initial and target nodes

Sex		Male	Female
Number of nodes in the graph		2 344 253	281 257
Number of nodes between the initial and target nodes		24	20
Number of neighbors per node	mean	18.1	15.7
	SD	3.7	3.0
	min	2	2
	max	24	19

Discussion

Using graph analysis as a new method applied to dietary transitions, we could identify and characterize trajectories of dietary changes that are optimal for long-term health while avoiding nutrient deficiency risk.

This work was based on the premise that identifying detailed trajectories could help decipher the most important steps in a dietary transition, mainly by analyzing the pattern of changes along the trajectory and highlighting the first steps being the most effective in mitigating health risks. In this regard, our findings point out critical changes to transition toward a healthy diet.

1. Trajectories analysis

Firstly, it is important to note that the graph generated for males is nearly ten times larger than the one for females. This discrepancy is mostly related to the different number of steps required to transition from the initial to the target diet. As the gap between observed and target diets increases, the number of nodes in the graph expands exponentially. This high gap is a result of the poorer quality of observed diet in males. However, despite this difference in size, the graph structure remains similar, as the average number of neighbors per node remains similar.

With the first six steps of the trajectories, males and females HR scores are already improved by at least two third of the maximum improvement. It confirms that this method identifies very efficient steps, especially in the first steps of the trajectories.

The first dietary change in all dietary trajectories was an increase in whole/semi-refined bread. This result aligns with the importance of whole grains in health, as shown by the GBD studies (1,33). In France, given that a diet low in whole grain is the highest dietary risk factor in both males and females (34), it is not surprising that an increase in whole grain products is the first step. The importance of the increase in whole grain is clearly shown by the presence of two successive increases in whole bread in the first two steps of the trajectory in males. When whole grains are concerned, attention should be paid to the residual pesticide levels related to agricultural practices. Whole grains products are, on average, higher in pesticides than refined grains products (35), which might be considered as potentially opposing the benefits of whole and semi-refined bread. However, this potential drawback should not be expected to call into question the final benefit of whole and semi-refined bread since, as accounted for in the GBD data, the DALYs gain is based on epidemiological data, with consumption mostly coming from conventional agriculture (36).

Important steps in the dietary trajectory that we identified were increases in fruit and, to a lesser extent, vegetables, which is again in line with the relative importance of fruit and vegetables in the HR criteria. Importantly, these increases are early and that of fruit immediately follows the increase in whole/semi-refined bread.

A decrease in red meat in the first phase of the trajectory was only found in male, at step #6, despite the importance of red meat in the HR function. This result can be explained mainly by the importance of the food groups being increased before, which are consumed in very small amounts compared to the recommended value. It should be noted that red meat could have been decreased at each of the first steps of the trajectories because it was not limited by a constraint related to nutritional security. Thus, its reduction at step #6 only comes from the optimization criteria. For females, no decrease was needed since the target diet was not very low in red meat and, when expressed as portion steps, the initial and the target values had similar amounts. On the other hand, since meat is dense in protein, bioavailable iron, and vitamin B12 (37), it has often been proposed that reducing meat consumption may increase the inadequacy of some nutrients. In a recent work, we showed that the critical nutrients when decreasing meat in healthy diets are bioavailable iron, zinc, and vitamin A (16). However, we found no increase in the risk of inadequacy for all those nutrients when reducing meat within this trajectory (Supplemental Figure 2.N, 2.U and 2.A and Supplemental Figure 3.N, 3.U and 3.A). This is in line with previous studies showing that protein quantity is not an important risk in the transition in western countries to a plant-based diet (38,39). The risk of insufficient intake of vitamin B12 also appears to depend on the overall decrease in foods of animal origin rather than meat in isolation (40).

Furthermore, as for other trajectory steps, decreasing red meat was done without increasing the risk of deficiency since the graph was trimmed to avoid diets featuring any increase in the risk of overt deficiency, as assessed using the SecDiet metric. It should be noted that for female all along the trajectory and for male after step#18, bioavailable iron is under the reference value (Supplemental Figure 3.N). However, the intake remains above the lower threshold value ensuring <5% deficiency prevalence. Thus, iron-deficiency anemia in the population would stay around today's level. We did not impose that the iron reference value was met during the trajectory and in the target diet. This choice was made as the current diet does not meet this high reference value, in particular for female, and previous work has demonstrated that the current value regarding bioavailable iron is over constraining when intending to model healthier diets (6). It is striking that increases in legumes preceded the decrease in red meat consumption. As we have previously discussed (41), this shows the importance of the transition between animal to plant protein. The importance of legumes in this pathway may mean that in diet changes consisting in lowering red meat consumption, some pathways not featuring legumes are discarded because of nutritional risk; since legumes are rich in nutrients that may be limiting in this case.

2. Nutrient-secured healthiest trajectories and possible implication for guidelines

In this respect, the trajectory we have identified is the result of a combination of the long-term health value of the portioned changes, the position of the initial and target diets, and the strict avoidance of pathways leading to diets that would be nutritionally unsecured. This latter characteristic is particularly important because while the target diet is fully nutritionally adequate, pathways of changes leading to that diet may include an increase in the risk of nutrient deficiency. Therefore, not all changes were available at all steps and the optimization consisted in choosing best choice at each step (Supplemental Figures 4 and 5 and Supplemental Tables 4 and 5) to make the best trajectory overall, i.e. the most efficient for health.

Our findings of critical steps within the best overall pathways toward a healthier diet may have implications for more precise dietary guidelines. Pointing out simple changes to be prioritized should be easier for the general audience and, therefore, may help food-based dietary guidelines to be more effective. For instance, although increasing whole grain products is included in dietary guidelines, it is not the recommendation that is the most widely known and applied (42) compared to the well-known recommendation regarding fruit and vegetables (13). Furthermore, pointing out a more precise change, such as here focusing on whole-meal bread, should give more applicable advice while maintaining a strong scientific rationale. Thus, prioritizing some dietary changes could change how public health nutrition programs are put forward. For whole grain in general, and whole bread in particular, which can be eaten similarly to white bread, the option may be even easier to apply. In this regard, it should be conceded that because the graph-based method consists of a series of increases/decreases, our results identify isolated increases in whole-meal bread, not swaps for white bread. However, an increase in one portion step of whole-meal bread consists in adding 60g of whole-meal bread, and this change can be conceived and implemented apart from a reduction in white bread. Furthermore, some directly successive changes might be translated as swaps in the recommendations, provided it makes sense from a practical viewpoint. We think it is the case for the decrease in red meat, which is immediately preceded by two successive increases in legumes and can be easily operationalized as replacing 40g red meat with 80g legumes. Of course, our results do not mean that these changes do not require a large effort from the population. Indeed, in western countries, a large part of the population does not know how to cook legumes (43), and culturally speaking, meat is still an important part of the meal. Lastly, insisting on the diet changes that appear first in the identified trajectories could stand as a rationale way to prioritize changes, while keeping those changes integrated within a full trajectory. Defining ordered step-by-step changes could concur with setting graded tasks, which is part of the behavioral changes techniques (44,45).

3. Graph analysis for identifying optimal dietary transition toward target diets

Graph analysis methods are well known for the analysis of road networks (14), social networks (15), or protein structures (16). However, to our knowledge, it has never been used to analyze diet transition. Other methodological approaches have been used to analyze diet transitions. While optimization studies are often used to identify target diets that could back food-based dietary guidelines (46), optimization has also been used to identify intermediate diets. For example, we recently identified a series of diets with progressively proportionally reduced amounts of meat (16) and this could be taken as showing a trajectory of change made of intermediate diets towards the lowest level of meat possible. However, this approach is different since it shows a series of healthy diets after forced scenarios of changes (here, reduction in meat consumption) and not the best overall healthy pathway toward the final target. These diets may not be part of the most efficient pathway. Furthermore, this serial optimization method shows diets that are not directly connected at the finest level, i.e., they are not demonstrated as being connected to one another with changes pathways with no nutritional risks, and diets can be very far from each other. Few methods have considered trajectories as a series of very fine level changes. Verger et al. (18) and De Gavelle et al. (17) have used such an approach made by step-by-step changes at portion size level. However, this method also has some limits. If a local optimum is reached, it is impossible to continue to a better one. Also, the method finds the best next step at each point and not the best path overall. A lesser improvement at one step could unlock a diet path that is in fact much more efficient overall.

In contrast, with the present graph-based method, we can optimize all along the trajectory, i.e., find the best series of diets taken together considering their health value. Another central advantage of the method is that we can apply any set of constraints on all nodes included in the graph. Here, the dietary and nutritional constraints that we chose were embedded in the graph, so they applied to all trajectories. For example, in all the graph during its construction, the SecDiet could not be lower than that of the observed diet, and this precluded that the identified trajectories could include a point diet with a nutritional security risk.

One major feature of this method is that it identifies a trajectory from the observed diet to a target diet, but it requires that the target diet should be defined with another method. This study used multicriteria optimization under constraints to define the target diet. Another parametrization of the model (in particular for less conservative weighting of diet deviation) or another method that would change the target diet may lead to a different graph and a different optimal trajectory than the one we identified here. One of the limits of this method is the restraint on the target diet in terms of the total distance from the observed diet because the target diet should be close enough to the initial one, given the fine level of resolution of the graph if the target is not close enough, its dimension will be too large. Being close enough means having less than 25 number of portions steps (i.e., nodes) between the node representing the observed diet and the node representing the target diet. Indeed, because of how the graph is created and even with the constraints added during the construction, the number of nodes in the graph is exponential to the number of portions steps between the initial and target diets, and the graph for male is already composed of 2.3 million nodes. Thus, if choosing a target diet that is farther from the initial diet, then the resolution of the representation of the diet with the graph should be lower, which means choosing a lower number of larger food groups or higher steps in portion sizes of each food. It should also be noted that this methodological approach strongly depends on the optimization criteria. In this study, the HR criteria we chose represent the diet proximity to reference values (TMREL) that are those of the lower risk for chronic diseases, weighted by the attributable fraction of the risk factor. Thus, optimizing this score aims to identify the trajectory that would lead to the lower risk of long-term morbidity and mortality, therefore being the integral trajectory that is the more efficient for long-term health. The improvement of health risk criteria along the trajectory is not uniform, and the best trajectory that has been identified is the best overall, accounting together for the 14 components of the criteria and taking into account the whole trajectory till the end. Therefore, some components that align less with the others in the optimization may be only reduced in the second part of the trajectory and may even be transiently deteriorated. This was the case for sodium. Reaching sodium reference value has been found constraining in modelling studies (6), and changes within food categories, which were not addressed here, are helpful to lower sodium intake (18,47). The advantage of the method is to weight all components in the health criteria to find the overall best change for the general population, but it may be less appropriate for some specific populations with different risk factors. However, the method can be adapted to account for different criteria in a more specific population. For instance, in people with hypertension, sodium intake could be the optimization criteria either alone or combined with others. The method presented here is indeed versatile, with the possibility to add constraints directly into the nodes of the graph. In this example, a maximum level of sodium could be added if the goal is to identify trajectories strictly avoiding transient increase in sodium intake. The same method can be used with other target diets and optimization criteria. For example, another study could find optimal pathways to adherence to the dietary guidelines or optimal pathways for reducing an environmental pressure score or a combination of a nutritional and environmental score. This method could also be used to identify individuals' trajectories of diet transition. Indeed, the only change necessary is the modification of the initial diet. Other changes, like adapting the target diet, could also be made depending on the question. However, it should be noted that generating graphs for each individual and identifying the optimal trajectories requires considerable computational time and can, therefore, only be carried out for a restricted number of individuals. Moreover, as they progress, the trajectories of different individuals will probably rapidly converge to a similar path. Therefore, rather than at the individual level, a set of trajectories may easily be drawn for subgroups of the population and give enough information about the variability of the first steps according to the characteristics of these groups, such as their initial dietary pattern, as identified in a population survey. Defining the individual trajectory of a given individual to provide tailored information is also possible on this basis but remains a very different field of application.

It should also be noted that this method does not include all the factors that need to be taken into account if the results are to be readily translated into dietary recommendations. Indeed, this work focuses on nutrition/health impact of diet trajectories and does not take into account to what extent changes in the consumption of a food group are accessible to the people, depending on their characteristics. For example, the affordability, accessibility, or individual preferences could be factors modifying the implementation of the changes that we have identified here.

Moreover, this work only considered nutrient intake coming from food groups. In a population with a high use of dietary supplement, it should be useful to identify whether specific solutions could be found while considering nutrient intakes from supplements. However, the percentage of individuals taking supplements in the study population was relatively low (28.5%), and supplements were not taken regularly (48). Importantly, using of supplements has been found to have no impact on the risk of insufficient intakes in the population (49). The accessibility of supplements is not the same for all the population, and supplements may be more beneficial in lower socioeconomic status individuals who eat diets of lower quality and have higher risks of insufficient nutrient intakes (27). However, in future studies, it could be interesting to evaluate what could be the effect of supplementation on the trajectories and if better solutions could be found if an effective supplementation strategy was set. Likewise, the effect of fortifying specific foods or food staples may also be the subject of further studies.

One of the other advantages of this method is the possible use of multiple variables to characterize the nodes (diets). These variables can be used as constraints to invalidate some pathways. These variables can also be used to describe the trajectory with respect to other dimensions of the diets.

Conclusion

Methods based on graph theory can be used to draw all possible nutrient-secure pathways of changes to reach a reference diet and identify the detailed integral pathway that is optimal overall for improving long-term health. Here, we could notably show that an increase in the portion size of whole/semi-refined bread is a critical first step in this pathway before increases in fruit and vegetables. It can show how small and realistic changes should be prioritized while keeping with the final objective. Such analysis models may be useful for refining food-based dietary guidelines on practical grounds.

Acknowledgements

The authors' responsibilities were as follows – EP, JW and FM designed research; EP was primarily responsible for preprocessing the data, conducting trajectory optimization and interpreting the results, with advisory inputs from JW, AD, HF and FM; AD and HF provided some data and tools; EP drafted the manuscript; EP and FM wrote the manuscript with critical comments from JW, AD, and HF. All authors approved the final manuscript. The authors report no conflicts of interest.

References

1. Murray CJL, Aravkin AY, Zheng P, Abbafati C, Abbas KM, Abbasi-Kangevari M, Abd-Allah F, Abdelalim A, Abdollahi M, Abdollahpour I, et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet 2020;396:1223–49.

2. Springmann M, Clark M, Mason-D'Croz D, Wiebe K, Bodirsky BL, Lassaletta L, de Vries W, Vermeulen SJ, Herrero M, Carlson KM, et al. Options for keeping the food system within environmental limits. Nature 2018;562:519–25.

3. Sustainable healthy diets: guiding principles [Internet]. World Health Organization (WHO), Food and Agriculture Organization of the United Nations (FAO); 2019 Oct p. 37. Available from: https://www.who.int/publications-detail-redirect/9789241516648

4. The US Burden of Disease Collaborators. The State of US Health, 1990-2016: Burden of Diseases, Injuries, and Risk Factors Among US States. JAMA 2018;319:1444–72.

5. The French Agency for Food Environmental and Occupational Health & Safety (ANSES). Avis de l'Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail relatif à « Actualisation des références nutritionnelles françaises en vitamines et minéraux » (saisine n°2018-SA-0238, saisine liée n°2012-SA-0103, 2021). 2021; Available from: http://www.anses.fr/fr/system/files/NUT2018SA0238Ra.pdf

6. Dussiot A, Fouillet H, Wang J, Salomé M, Huneau J-F, Kesse-Guyot E, Mariotti F. Modeled healthy eating patterns are largely constrained by currently estimated requirements for bioavailable iron and zinc—a diet optimization study in French adults. Am J Clin Nutr 2022;115:958–69.

7. Gazan R, Brouzes CMC, Vieux F, Maillot M, Lluch A, Darmon N. Mathematical Optimization to Explore Tomorrow's Sustainable Diets: A Narrative Review. Adv Nutr Oxford Academic; 2018;9:602–16.

8. Stewart C, Piernas C, Cook B, Jebb SA. Trends in UK meat consumption: analysis of data from years 1–11 (2008–09 to 2018–19) of the National Diet and Nutrition Survey rolling programme. Lancet Planet Health 2021;5:e699–708.

9. Deliens T, Mullie P, Clarys P. Plant-based dietary patterns in Flemish adults: a 10-year trend analysis. Eur J Nutr 2022;61:561–5.

10. World Health Organization. Regional Office for Europe, Policies EO on HS and, Candari CJ, Cylus J, Nolte E. Assessing the economic costs of unhealthy diets and low physical acitivity: an evidence review and proposed framework

[Internet]. World Health Organization. Regional Office for Europe; 2017 [cited 2022 Jul 27]. 90 p. Available from: https://apps.who.int/iris/handle/10665/326302

11. Lea EJ, Crawford D, Worsley A. Public views of the benefits and barriers to the consumption of a plant-based diet. Eur J Clin Nutr Nature Publishing Group; 2006;60:828–37.

12. Pohjolainen P, Vinnari M, Jokinen P. Consumers' perceived barriers to following a plant-based diet. Br Food J Emerald Group Publishing Limited; 2015;117:1150–67.

13. Escalon H, Beck F, Bossard C. Associations entre la connaissance des recommandations du Programme national nutrition santé et les comportements en matière d'alimentation et d'activité physique. Rev DÉpidémiologie Santé Publique 2013;61:37–47.

14. ANSES. the Third Individual and National Survey on Food Consumption (INCA3 Survey) [Internet]. 2017 [cited 2022 Jan 24]. Available from: https://www.anses.fr/en/system/files/NUT2014SA0234EN.pdf

15. Agence Française de Sécurité Sanitaire des Aliments. Synthèse de l'étude individuelle nationale des consommations alimentaires 2:(INCA 2) 2006-2007 [Internet]. 2009 [cited 2022 Jul 27]. Available from: https://www.anses.fr/fr/system/files/PASER-Ra-INCA2.pdf

16. Dussiot A, Fouillet H, Perraud E, Salomé M, Huneau J-F, Kesse-Guyot E, Mariotti F. Nutritional issues and dietary levers during gradual meat reduction – A sequential diet optimization study to achieve progressively healthier diets. Clin Nutr 2022;41:2597–606.

17. de Gavelle E, Leroy P, Perrimon M, Huneau J-F, Sirot V, Orset C, Fouillet H, Soler L-G, Mariotti F. Modeled gradual changes in protein intake to increase nutrient adequacy lead to greater sustainability when systematically targeting an increase in the share of plant protein. Clim Change 2020;161:129–49.

18. Verger EO, Holmes BA, Huneau JF, Mariotti F. Simple Changes within Dietary Subgroups Can Rapidly Improve the Nutrient Adequacy of the Diet of French Adults. J Nutr 2014;144:929–36.

19. Dubuisson C, Dufour A, Carrillo S, Drouillet-Pinard P, Havard S, Volatier J-L. The Third French Individual and National Food Consumption (INCA3) Survey 2014–2015: method, design and participation rate in the framework of a European harmonization process. Public Health Nutr 2019;22:584–600.

20. French Agency for Food, Environmental and Occupational Health Safety (ANSES). Opinion of the French Agency for Food, Environmentaland Occupational Health & Safety on "the Third Individual and National Survey on

Food Consumption (INCA3 survey)" [Internet]. 2017 [cited 2021 Jul 8]. Available from: https://www.anses.fr/en/system/files/NUT2014SA0234EN.pdf

21. Henry C. Basal metabolic rate studies in humans: measurement and development of new equations. Public Health Nutr 2005;8:1133–52.

22. 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 2000;24:1119–30.

23. ANSES. Actualisation des repères du PNNS : élaboration des références nutritionnelles. Maisons-Alfort, France: French Agency for Food, Environmental and Occupational Health Safety (Anses). [Internet]. 2016 [cited 2022 May 13]. Available from: https://www.anses.fr/fr/system/files/NUT2012SA0103Ra-2.pdf

24. ANSES. Table de composition nutritionnelle des aliments CIQUAL [Internet]. 2016 [cited 2021 Jun 22]. Available from: https://ciqual.anses.fr/#

25. Fouillet H, Dussiot A, Perraud E, Wang J, Huneau J-F, Kesse-Guyot E, Mariotti F. Plant to animal protein ratio in the diet: nutrient adequacy, long-term health and environmental pressure [Internet]. medRxiv; 2022 [cited 2022 Sep 14]. p. 2022.05.20.22275349. Available from: https://www.medrxiv.org/content/10.1101/2022.05.20.22275349v2

26. Diestel R. Graph Theory [Internet]. 5th ed. Heidelberg: Springer-Verlag; 2017 [cited 2022 Jul 11]. Available from: https://diestel-graph-theory.com/eBook.html

27. Salomé M, Kesse-Guyot E, Fouillet H, Touvier M, Hercberg S, Huneau J-F, Mariotti F. Development and evaluation of a new dietary index assessing nutrient security by aggregating probabilistic estimates of the risk of nutrient deficiency in two French adult populations. Br J Nutr Cambridge University Press; 2021;126:1225–36.

28. Dijkstra EW. A Short Introduction to the Art of Programming [Internet]. Technische Hogeschool Eindhoven; 1971. Available from: https://books.google.fr/books?id=JE8ZAQAAIAAJ

29. Fouillet H, Dussiot A, Perraud E, Wang J, Huneau J-F, Kesse-Guyot E, Mariotti F. Plant to animal protein ratio in the diet: nutrient adequacy, long-term health and environmental pressure. Front Nutr [Internet] 2023 [cited 2023 Jun 20];10. Available from: https://www.frontiersin.org/articles/10.3389/fnut.2023.1178121

30. Rossum G van, Drake FL. The Python language reference. Release 3.0.1 [Repr.]. Hampton, NH: Python Software Foundation; 2010. 109 p.

31. Hagberg A, Swart P, S Chult D. Exploring network structure, dynamics, and function using networkx [Internet]. Los Alamos National Lab. (LANL), Los Alamos, NM (United States); 2008 Jan. Report No.: LA-UR-08-05495; LA-UR-

08-5495. Available from: https://www.osti.gov/biblio/960616-exploring-network-structure-dynamics-function-using-networkx

32. Aynaud T. python-louvain x.y: Louvain algorithm for community detection [Internet]. 2020. Available from: https://github.com/taynaud/python-louvain

33. GBD 2013 Mortality and Causes of Death Collaborators. Global, regional, and national age–sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet 2015;385:117–71.

34. THE LANCET. Global Burden of Disease: GBD cause and risk summaries [Internet]. 2023 [cited 2023 Jun 16]. Available from: https://www.thelancet.com/gbd/summaries

35. Wang J, Hasanalieva G, Wood L, Anagnostopoulos C, Ampadogiannis G, Bempelou E, Kiousi M, Markellou E, Iversen PO, Seal C, et al. Effect of wheat species (Triticum aestivum vs T. spelta), farming system (organic vs conventional) and flour type (wholegrain vs white) on composition of wheat flour – Results of a retail survey in the UK and Germany – 3. Pesticide residue content. Food Chem X 2020;7:100089.

36. eurostat. Organic farming statistics [Internet]. 2022 [cited 2022 Sep 15]. Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Organic_farming_statistics

37. Pereira PM de CC, Vicente AF dos RB. Meat nutritional composition and nutritive role in the human diet. Meat Sci 2013;93:586–92.

38. Richter CK, Skulas-Ray AC, Champagne CM, Kris-Etherton PM. Plant Protein and Animal Proteins: Do They Differentially Affect Cardiovascular Disease Risk? Adv Nutr 2015;6:712–28.

39. Mariotti F. Animal and Plant Protein Sources and Cardiometabolic Health. Adv Nutr 2019;10:S351–66.

40. Ankar A, Kumar A. Vitamin B12 Deficiency. StatPearls. [Internet] Treasure Island (FL): StatPearls Publishing; 2022 [cited 2022 Nov 29]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK441923/

41. Perraud E, Wang J, Salomé M, Huneau J-F, Lapidus N, Mariotti F. Plant and Animal Protein Intakes Largely Explain the Nutritional Quality and Health Value of Diets Higher in Plants: A Path Analysis in French Adults. Front Nutr [Internet] 2022 [cited 2022 Aug 23];9. Available from: https://www.frontiersin.org/articles/10.3389/fnut.2022.924526

42. Kesse-Guyot E, Touvier M, Henegar A, Czernichow S, Galan P, Hercberg S, Castetbon K. Higher adherence to French dietary guidelines and chronic diseases

in the prospective SU.VI.MAX cohort. Eur J Clin Nutr Nature Publishing Group; 2011;65:887–94.

43. Magrini M-B, Fernandez-Inigo H, Doré A, Pauly O. How institutional food services can contribute to sustainable agrifood systems? Investigating legume-serving, legume-cooking and legume-sourcing through France in 2019. Rev Agric Food Environ Stud 2021;102:297–318.

44. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: The CALO-RE taxonomy. Psychol Health Routledge; 2011;26:1479–98.

45. Samdal GB, Eide GE, Barth T, Williams G, Meland E. Effective behaviour change techniques for physical activity and healthy eating in overweight and obese adults; systematic review and meta-regression analyses. Int J Behav Nutr Phys Act 2017;14:42.

46. Mariotti F, Havard S, Morise A, Nadaud P, Sirot V, Wetzler S, Margaritis I. Perspective: Modeling Healthy Eating Patterns for Food-Based Dietary Guidelines—Scientific Concepts, Methodological Processes, Limitations, and Lessons. Adv Nutr 2021;12:590–9.

47. Britten P, Cleveland LE, Koegel KL, Kuczynski KJ, Nickols-Richardson SM. Impact of Typical Rather than Nutrient-Dense Food Choices in the US Department of Agriculture Food Patterns. J Acad Nutr Diet 2012;112:1560–9.

48. ANSES. Étude individuelle nationale des consommations alimentaires 3 (INCA 3) [Internet]. 2017 Jun. Available from: https://www.anses.fr/fr/system/files/NUT2014SA0234Ra.pdf

49. ANSES. AVIS de l'Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail relatif à l'évaluation des apports en vitamines et minéraux issus de l'alimentation non enrichie, de l'alimentation enrichie et des compléments alimentaires dans la population française : estimation des apports usuels, des prévalences d'inadéquation et des risques de dépassement des limites de sécurité [Internet]. 2015 [cited 2023 Jun 22]. Available from: https://www.anses.fr/fr/system/files/NUT2012sa0142.pdf