

Nutritional Epidemiology Research Unit
INSERM U557/INRA 1125/CBAM/University of Paris 13

RESEARCH PROTOCOL



THE NUTRINET-SANTE STUDY

*A WEB-BASED PROSPECTIVE COHORT STUDY OF THE
RELATIONSHIP BETWEEN NUTRITION AND HEALTH
AND OF DIETARY PATTERNS AND NUTRITIONAL
STATUS PREDICTORS*

Study coordination:

Principal Investigator: Serge Hercberg

Co-Principal Investigator: Pilar Galan

Co-investigators: Mathilde Touvier, Caroline Méjean, Emmanuelle Kesse, Léopold Fezeu,
Sandrine Péneau, Chantal Julia, Katia Castetbon

UMR INSERM U557/INRA 1125/CBAM/University of Paris 13

In partnership with:

Ministry of Health and Sports, INPES, InVS, FRM, INSERM, INRA, CNAM,
University of Paris 13

EXECUTIVE SUMMARY

Numerous clinical, physiopathological and epidemiological studies have highlighted the important role of nutrition - as a risk or a protective factor – regarding the incidence of chronic conditions and diseases, such as cardiovascular disease (CVD), some types of cancer, type 2 diabetes, obesity, hypertension, cognitive decline, and depression. These pathologies constitute a heavy public health burden in France and in the industrialized world in general. Hence, the identification of risk and protective factors related to nutrition represents an important step in the development of dietary guidelines aimed at health promotion and disease prevention.

Such chronic diseases have multifactorial etiology, including genetic, biological, and environmental factors. In turn, the field of nutritional epidemiology has a paramount role in the study of the relations between nutrition and health because it can provide direct evidence of both dietary exposures and disease incidence in the context of everyday life. In order to determine the role of specific nutritional factors, large-scale population-based cohorts are critical, as are comprehensive data on eating habits, dietary intake, and a slew of potential confounders.

Along these lines, the Internet provides valuable opportunities to reach a sizeable number of volunteer participants, to collect information on a wide range of topics on a regular basis, and to automatically store and treat such data. France is indeed at the top of the list regarding Internet usage across Europe. For example, in May 2008, 32 700 000 French individuals aged 11+ years (ie, 61.9% of the population in that age range) were connected to the Internet at home, work, public places or private locations. The profile of Internet users suggests that it is indeed feasible to reach a volunteer sample that closely resembles the general population in terms of age (and especially senior citizens), socioeconomic status (SES), and geographic distribution. According to OCDE, one out of every four Internet users is older than 55 years and about 29% of Internet users belong to low SES. In 2011, 79.6% of the French had Internet access (source: OCDE).

Overarching aim of the NutriNet-Santé Study

The principal study aim pertains to a comprehensive investigation of the relations between dietary/nutrient intake, eating habits and all-cause and cause-specific (eg, CVD, cancer) mortality. In addition, the study aims to elucidate the link between such nutrition-based predictors and chronic disease incidence (eg, CVD, cancer, type 2 diabetes, obesity, hypertension, cognitive decline, depression, migraine, metabolic syndrome, rheumatoid arthritis), as well as healthy aging and quality of life.

Additional study objectives

- Investigate multiple determinants (eg, socioeconomic, psycho-cognitive, cultural, sensory, biological, genetic) of eating habits, dietary patterns, nutritional status, and overall health status;
- Investigate the relations of dietary/nutrient intake and eating habits with various biomarkers;
- Provide a long-term follow-up and description of population-level trends as regards dietary behaviors and nutritional status;
- Evaluate the impact of public health efforts in terms of perception, effectiveness and acquired knowledge.

Given the quantity and quality of the collected information as well as the size of the study sample, the NutriNet-Santé Study is well positioned to serve as a large databank regarding population-level information on nutrition and health in France, and will also constitute one of the largest epidemiological data sources worldwide.

Materials and methods

The NutriNet-Santé is a large observational prospective cohort study (sample size >500 000 registered individuals, with a target of 300 000 active participants) with a planned long-term follow-up (10 years). All participants (called “Nutrinautes”) are aged 18+ years and a target of least 250 000 participants older than 45 years has been set. Recruitment was launched in April-May 2009 and is carried out via a large, nation-wide, recurrent multimedia campaign (including television, radio broadcasts, print media, and the Internet). The call for volunteers stresses the need to recruit individuals “who will be engaged in furthering the state of research and public health.” Enrollment in the study will remain open for its entire duration.

All “Nutrinautes” will be followed via a website specifically developed for this purpose (www.etude-nutrinet-sante.fr) and via adapted research instruments which were initially developed for the SU.VI.MAX and ENNS studies. All questionnaires have been designed for online self-administration, via a secure HTML interface. All data storage, confidentiality and security measures will be strictly observed.

At enrollment, the participants are asked to complete a baseline set of questionnaires, providing information on dietary practices (via 3 24h dietary records completed over a period of 15 days), physical activity, anthropometrics, lifestyle, and health status. The individuals are considered enrolled and active once this baseline set of questionnaires has been fully filled out. Given the results of the pilot study, we expect a 60% active enrollement (ie, 300 000 active participants from >500 000 individuals who have registered on the website). Over the course of the follow-up, the “Nutrinautes” receive automated email messages regarding the study progress as well as any upcoming questionnaires to be completed via a built-in link. Collected health status data pertain to all-cause and cause-specific mortality, morbidity, and quality of life. Detailed information on incidence and causes of mortality will be obtained from the national vital statistics registry. Regarding the study outcomes, we expect an approximately 20% attrition (ie, 240 000 participants over 5 years).

The NutriNet-Santé study also includes a Biobank where biological specimens from a subsample are collected and stored.

In response to the evolution of research interests and potential collaborations, it is entirely feasible to add *ad hoc* questionnaires and ancillary protocols pertaining to selected subsamples.

Generally, the study website will provide useful information on nutrition, health, and scientific research, thus promoting the feeling of “belonging” among the cohort participants.

The NutriNet-Santé Study is supported by several ministries, research institutions, and science-based foundations. The databank will constitute a valuable and rich source of information available for use by the scientific community, by health surveillance and needs assessment agents, and for the provision of expertise on specific topics.

The NutriNet-Santé Study

Nutritional Epidemiology Research Unit (UREN)
INSERM U557/INRA 1125/CNAM/University of Paris 13

Principal coordinators

- Pr Serge Hercberg (PU-PH, University of Paris 13/Department of Public Health, Avicenne Hospital, Bobigny; Director, UREN), Principal Investigator
- Dr Pilar Galan (DR1, INRA, Deputy director, UREN), Co-Principal Investigator
- Dr Mathilde Touvier (CR2, INSERM/UREN), co-investigator
- Dr Caroline Méjean (CR2, INRA/UREN), co-investigator
- Dr Emmanuelle Kesse (CR1, INRA/UREN), co-investigator
- Dr Sandrine Péneau (MCF University of Paris 13/UREN), co-investigator
- Dr Katia Castetbon (InVS/USEN/UREN), co-investigator
- Dr Léopold Fezeu (MCF-Chaire d'Excellence INRA, University of Paris 13/UREN), co-investigator
- Dr Chantal Julia (AHU, Avicenne Hospital, University of Paris 13/UREN), co-investigator

Technical support

Younes Esseddik (IT engineer, INSERM), Thi Hong Van Duong (IT engineer, University of Paris 13), Laurent Bourhis (statistician, INRA), Nathalie Arnault (statistician, IE INSERM), Véronique Gourlet (statistician, IE INSERM), Charlie Menard (data manager, University of Paris 13), Fabien Szabo (data manager, IR CNAM), Florence Guénard-Charpentier (dietitian, CNAM), Charlotte Voegtlin (dietitian, University of Paris 13), Priscillia Leopold (dietitian, University of Paris 13), Adrien Massau (dietitian, University of Paris 13), Mohand Aït Oufella (IT engineer, IE INSERM), Yasmina Chelghoum (IT engineer, AI INSERM), Paul Flanzky (IT engineer, TR INRA) Nathalie Pecollo (research engineer, INRA).

Protocol approval

- Institutional Review Board, *Comité de Qualification Institutionnelle* INSERM, July 10, 2008, n°IRB0000388 FWA00005831
- CCTIRS, July 11, 2008, n°08.301; July 15, 2010, n°10-367
- CNIL, February 24, 2009, n°908450; February 12, 2010, n° 909216

Support

The Ministry of Health and Sports (DGS), INPES, InVS, INSERM, INRA, CNAM, University of Paris 13, and the Medical Research Foundation (FRM).

Table of contents

1	BACKGROUND	7
1.1	Principal public health domains where nutritional factors play a major role	7
1.1.1	Cancer	7
1.1.2	Cardiovascular disease (CVD)	7
1.1.3	Conditions involved in the etiology of cancer and CVD	8
1.1.3.1	Obesity	8
1.1.3.2	Type 2 diabetes	8
1.1.3.3	Hypertension	9
1.1.4	Cognitive decline, vascular dementia, Alzheimer's disease	9
1.2	Principal methodological considerations in nutritional epidemiology regarding the relations among nutrition, morbidity, and quality of life	9
1.2.1	Determination of the study design and the sample size	9
1.2.2	Dietary intake in relation to health and morbidity	10
1.3	Internet use in France: number and profiles of users	11
1.4	Overarching aims of the NutriNet-Santé Study	13
2	OBJECTIVES OF THE NUTRINET-SANTE STUDY	14
2.1	Main objectives	14
2.2	Additional objectives	14
3	EXPOSURE AND OUTCOME ASSESSMENT	15
3.1	Main exposures	15
3.2	Main outcomes	15
3.3	Hypotheses, sample size, and statistical power	15
4	MATERIALS AND METHODS	17
4.1	Recruitment	17
4.2	Registration	17
4.3	Enrollment in the study	19
4.4	Data collection timeframe	20
4.5	Data collection: modalities and type of information collected	21
4.5.1	Information collected as part of the «Baseline Kit»	21
4.5.1.1	Dietary questionnaire	21
4.5.1.2	Sociodemographic and lifestyle questionnaire	24
4.5.1.3	Health status questionnaire	24

4.5.1.4	Anthropometrics questionnaire.....	24
4.5.1.5	Physical activity and sedentary behavior questionnaire	25
4.5.2	Additional monthly questionnaires	25
4.5.3	Collection of follow-up data	26
4.5.4	Follow-up and validation of health events.....	26
4.5.4.1	Self-reporting of health events.....	27
4.5.4.2	Health event management.....	29
4.5.4.3	Health event validation committees.....	29
4.5.5	Health status follow-up using national medical registries.....	31
4.5.5.1	Follow up for mortality and cause of death.....	31
4.5.5.2	Follow up for health events.....	31
5	ANCILLARY PROTOCOLS.....	32
6	VOLUNTEER RETENTION PROTOCOL.....	32
7	DATA QUALITY CONTROL AND INSTRUMENT QUALITY.....	33
8	IT ASPECTS AND DATA SECURITY.....	33
9	NUTRI^{NET}-SANTE BIOBANK: COLLECTION OF BIOLOGICAL SPECIMENS.....	34
9.1	Population subsample.....	34
9.2	Logistics.....	34
9.2.1	Centralized participant scheduling system.....	34
9.2.2	Clinic-based procedures and centralized specimen collection.....	35
9.2.2.1	Coding of test tubes and urine specimen cups.....	35
9.2.2.2	Types of test tubes.....	35
9.2.2.3	Test tube pre-analysis treatment in local clinics.....	36
9.2.3	Transportation of specimens to central laboratory.....	36
9.2.4	Procedures implemented at the central laboratory	36
9.2.5	Specimen storage	37
9.3	Organization and management of the Biobank.....	37
10	REFERENCES.....	38

1. Background

Numerous clinical, physiopathological and epidemiological studies have highlighted the role of nutrition - as a risk or a protective factor – regarding the incidence of chronic conditions and diseases, such as cardiovascular disease (CVD), some types of cancer, type 2 diabetes, obesity, hypertension, and osteoporosis. These pathologies constitute a heavy public health burden in France and in the industrialized world in general, owing to substantial healthcare costs and increased mortality. Hence, the identification of risk and protective factors related to nutrition represents an important step in the development of dietary guidelines aimed at health promotion and disease prevention.

Such chronic diseases have multifactorial etiology, including genetic, biological, and environmental factors. In order to determine the role of specific nutritional aspects, large-scale population-based cohorts are critical, as are comprehensive data on eating habits, dietary intake, and a slew of potential confounders. At present, this represents a major area of research in the field of nutritional epidemiology.

1.1 Principal public health domains where nutritional factors play a major role

1.1.1 Cancer

In France, cancer represents the leading cause of mortality among men and the second leading cause of mortality among women, following CVD (1). In 2005, 29.5% of the mortality in the country was attributed to cancer (34% of all mortality in men and 24.7% - in women), which represented 155 407 individuals (92 106 men and 63 301 women). Over the past three decades, the annual incidence of cancer has nearly doubled, from 170 000 new cases in 1980 to 320 000 new cases in 2005 (2). This constitutes an 84% increase among women and 93% - among men. During the same period, mortality has declined by 24%. This trend is due to a decrease in the incidence of the most aggressive types of cancer (eg, esophagus, stomach, upper digestive tract), and to an increase in early detection via screening (eg, prostate, breast) which is associated with better treatment options and a better prognosis.

Cancer incidence has marked geographic variation: the highest rates are observed in the Nord-Pas-de-Calais region (669 per 100 000 among men, 372 per 100 000 among women), whereas the lowest rates are observed in the Midi-Pyrénées region (398 per 100 000 among men, 274 per 100 000 among women) (3).

1.1.2 Cardiovascular disease (CVD)

In France, as in the rest of the industrialized world, CVD constitutes one of the major causes of death. In 2005, 150 000 deaths were attributed to CVD (ie, 28.5% overall: 26% in men and 31% in women) (4). Death following a stroke represented 24% of all CVD-related mortality. This proportion has indeed decreased by 21% from 1990 to 2000, and by 11.7% between 2000 and 2005. Nonetheless, stroke represented the primary cause of death for approximately 34 000 individuals in 2005. Death due to coronary heart disease represented 28% of all CVD-related mortality. This proportion has indeed decreased by 8% and 10%, respectively, during the same period. However, it represented the primary cause of death for approximately 40 500 individuals in 2005.

CVD incidence also has marked geographic variation: the highest rates are observed in the Nord-Pas-de-Calais region (455 per 100 000 among men, 279 per 100 000 among women), whereas the lowest rates are observed in the Île-de-France region (299 per 100 000 among men, 189 per 100 000 among women) (4).

In turn, mortality due to hypertension-related disorders (eg, essential and secondary hypertension, hypertensive nephropathy) showed a slight increasing trend from 1991-1993 to 1998-2000: the rates

per 100 000 individuals have changed from 11.9 to 13.8 among men, and from 11.2 to 12.9 among women (4).

1.1.3 Conditions involved in the etiology of cancer and CVD

1.1.3.1 Obesity

According to the National Nutrition and Health Survey (*Etude Nationale Nutrition Santé*, ENNS, 2006), the prevalence of overweight and obesity among French adults aged 18-74 years was 32.4% and 16.9%, respectively (5). The proportion of obese men and women was nearly identical (6.1% and 17.6%, respectively), whereas overweight ($25 \leq \text{BMI} < 30$) was more common among men than among women (41.0% versus 23.8%). In both sexes, obesity prevalence increases with age and among adults aged 55-74 years it has reached 24.0% among men and 24.1% among women. Overall, among French adults, 49.3% are categorized as overweight or obese, which corresponds to 57.2% of men and 41.4% of women.

ENNS data indicate that the prevalence of class I obesity ($30.0 \leq \text{BMI} \leq 34.9$) among adults is 12.5% (12.9% in men, 12.1% in women), whereas the prevalence of class II (eg, morbid) obesity ($35.0 \leq \text{BMI} \leq 39.9$) is 3.4% (2.5% in men, 4.2% in women). Finally, the prevalence of class III obesity ($\text{BMI} \geq 45$) is estimated at 1.0% (0.7% in men, 1.2% in women). Regarding the prevalence of class I (but not II or III) obesity, there are age-specific differences: it concerns 3.7% of men aged 18-29 years, 13.1% of men aged 30-54 years, and 20.2% of those aged 55-74 years. Among women, the prevalence of class I obesity is 6.1% among those aged 18-29 years, 11.3% in those aged 30-54 years, and 17.8% in those aged 55-74 years.

Several national surveys assess obesity trends in France. However, anthropometric data are often self-reported, likely leading to underestimation of the actual rates. Such surveys are relatively easy to implement and administer on a recurrent basis. The ObEpi survey, for example, is administered every 3 years to a sample of 20 000 individuals aged 15+ years (6). It thus estimated the prevalence of obesity to be 8.2% in 1987, 9.6% in 2000, 11.3% in 2003, and 12.4% in 2006. The same increasing trend was reported by the ongoing “Santé” Survey administered by the Census Bureau (INSEE) to over 15 000 households (7). That survey allowed the collection of data on weight and height of 21 400 individuals aged 18-65 years in 1981, 13 400 individuals in 1992, and 22 600 individuals in 2003. According to the 2003 data, the prevalence of overweight was 34.8% in men and 21.2% in women, whereas the prevalence of obesity was 9.8% in men and 10.2% in women.

1.1.3.2 Type 2 diabetes

Through the end of 2005, a total of 2 325 000 individuals were treated for diabetes in France (excluding its overseas territories) (8). The same year, the combined prevalence of type I (insulin-dependent) and type II (non-insulin dependent) diabetes was 3.8% (versus 2.9% in 2000), which underscores an unexpected annual increase of 5.7%. In fact, given data about the prevalence of diabetes in 1999, the increase in the size of the total population, the growing obesity prevalence, and the aging of the population, the projected prevalence of diabetes by the year 2016 is 4.5% (2.8 million individuals), which corresponds to an annual rate of increase of 2.1% (9).

According to the ENTRED study, type 2 diabetes accounts for 91% of the diabetes cases in France (10, 11), whereas the number of undiagnosed cases was estimated to be anywhere from 200 000 to 500 000. In turn, according to the 2006 ENNS study, the prevalence of fasting hyperglycemia along with the prevalence of diabetes medication use were estimated at 4.7%, suggesting that the prevalence of undiagnosed hyperglycemia might be around 1.3%.

1.1.3.3 Hypertension

According to ENNS, the prevalence of hypertension increases with age and in 2006 it reached 31% among French adults aged 18-74 years (28% in men, 34% in women) (5).

Age group (years)	Men			Women		
	18-29	30-54	55-75	18-29	30-54	55-75
Treated hypertension (%)	0.0	6.6	39.0	1.5	11.2	38.3
Untreated hypertension (%)	2.4	23.9	27.0	4.1	10.9	15.0

1.1.4 Cognitive decline, vascular dementia and Alzheimer's disease

In France, there were 770 000 individuals aged > 75 years suffering from dementia, of whom 563 000 were women, according to 2003 data (12). Based on European-level information, the current prevalence of dementia (any type) is estimated at 6.4%, while that of Alzheimer's disease is estimated at 4.4% among those aged > 65 years (13). Results of meta-analyses suggest that the prevalence doubles with every 5-year increase in age, going from 1.5% in those aged 60-69 years to 40% in nonagenarians. This substantial increase coupled with the overall aging of the population and the relatively long period of nursing care for affected individuals, underscore the critical importance of addressing dementia in the elderly as a major public health priority.

In 2001, a group of international experts estimated that approximately 24.3 million individuals worldwide suffered from dementia, and noted that this prevalence will likely double every 20 years, to reach 81.1 million by the year 2040 (14).

1.2 Principal methodological considerations in nutritional epidemiology regarding the relations among nutrition, morbidity, and quality of life

1.2.1 Determination of the study design and the sample size

The field of nutritional epidemiology has a paramount role in the study of the relations between nutrition and health because it can provide direct evidence of both dietary exposures and disease incidence in the context of everyday life. For several decades, epidemiological studies (ecological, cross-sectional, case-control, and longitudinal) have been underscoring the fact that nutrients, food groups, and eating habits could act as predictors of mortality, chronic disease incidence, healthy aging and quality of life. However, the level of proof varies by the type of study and causality cannot be inferred.

Observational epidemiological studies use population samples and evaluate the relationships between the exposure of interest (ie, food consumption, physical activity, nutritional status, etc.) and risk of chronic disease. Such studies do not involve any intervention on part of the researchers. The following types of observational epidemiological studies exist: ecological, cross-sectional, case-control, and prospective (longitudinal).

- **Ecological studies** are often regarded as the first step in epidemiological pursuit, as they can help uncover potential correlations between food intake and disease risk in a given, geographically-defined population. Individuals in that population often differ along many aspects in addition to food intake, thus the established group-level correlations are considered only as a preliminary evidence of a relation.
- **Descriptive cross-sectional studies** are often intended to be representative of the population of interest. Such studies are useful for hypothesis generation as regards the relationships between nutrition and morbidity. However, the simultaneous assessment of both the exposure (ie, dietary

intake) and outcome (ie, disease/condition of interest) does not allow the establishment of temporality. Strength of such studies pertains to the relative representativeness of the sample, thus reducing selection bias.

- **Case-control studies** are often set up for the comparison of dietary intake (or other exposures of interest) between patients and healthy controls. In general, these studies are less expensive to implement than are cohort studies, given the need for a smaller number of participants. However, there are methodological limitations which could render the results difficult to interpret, especially from the perspective of nutritional epidemiology. For example, recall bias might be present as regards the recall of eating habits over a number of years (and before the expression of clinical symptoms). Another concern is the selection criteria applied to the control participants.
- **Prospective cohort studies** are often set up for the investigation of dietary intake (or other exposures of interest) of disease-free participants who are followed over several years in order to establish risk of disease incidence. At the end of the follow-up, the dietary intake of those who have developed the disease is compared with the dietary intake of those who have remained disease-free. This allows the establishment of temporality (ie, cause-effect). In other words, a notable strength of such type of studies is the possibility to avoid bias by accounting for the temporal ordering of the exposure (eg, dietary intake, physical activity, nutritional status, biomarker status) and the outcome.
- **Controlled intervention trials** constitute the means of establishing whether a causal relationship exists between nutrition and morbidity and/or mortality. In the field of nutritional epidemiology, such studies can be challenging to implement. Specifically, modifications in eating habits can occur and such changes are acceptable either over a short period of time (which automatically precludes the possibility of assessing the role of eating habits on morbidity/mortality) or following intensive, long-term education efforts, which would materially increase the cost/logistic burden of the study. In addition, the interpretation of the results of such studies might be challenging owing to the fact that individuals' dietary habits are subject to numerous influences and determinants which cannot be controlled. In reality, this type of studies is somewhat more useful for assessing the acceptability of dietary interventions rather than for the possible effects on morbidity and/or mortality.

Thus, longitudinal studies are considered as reference in nutritional epidemiology because they can provide the most favorable methodological tradeoff as regards investigation of the relationships between nutrition and chronic disease morbidity/mortality. Drawing firm conclusions from such studies is challenging due to potentially insufficient statistical power, especially when the main outcome is mortality.

Given the fact that chronic diseases (eg, cancer, CVD) are multifactorial in origin, it is necessary to work with large samples in order to have sufficient number of cases and to be able to assess the relative contribution of nutritional factors over the period of follow up. Dietary habits are known to be associated with other individual-level characteristics, such as socio-cultural status, educational level, and lifestyle. Such information needs to be taken into account when interpreting results. Thus, working with large population samples with sufficient heterogeneity in dietary habits allows the control for many potential confounders which, in turn, helps identify the specific contribution of different nutritional factors.

1.2.2 Dietary intake in relation to health and morbidity

The study of the role of nutrition in health often pertains to estimating the effect of a single nutrient or a single food group on a certain disease. However, interest in nutrition as a whole has considerably grown over the past 15 years, even though the investigation of some health domains has not benefitted from such approaches. Existing epidemiological data suggest an association between intake of certain foods (eg, protective effect of fruit/vegetables and fish intake; harmful effect of intake of salt, meat products, alcohol, saturated fat), physical activity/sedentary behavior and risk of

chronic diseases (cancer, CVD, diabetes, hypertension, obesity). Studies have also suggested the presence of relationships between these diseases and certain nutrients (eg, vitamins E, B6, B12, folate, polyunsaturated omega-3 fatty acids (PUFA) and CVD; intake of calcium and vitamin D and osteoporosis; intake of protein and energy-dense food and sarcopenia; intake of antioxidants and B vitamins and hyperhomocysteinemia; PUFA intake and neurodegenerative diseases, cancer, age-related macular degeneration) and decreased immune system function. Other lines of research suggest that these relations could be explained by genetic factors, however, there are methodological and conceptual limitations present in such findings (15-17).

Given the fact that food is composed of numerous nutrients, there are complex interactions and synergies among the nutrients consumed which, in turn, renders challenging the isolation of the role of a specific nutrient or type of food. Such limitations have led to the development of novel methods of analysis which simultaneously account for a number of factors or which permit the conduct of stratified analyses, thus reducing the variability in the construct of interest. This, in turn, can render the results difficult to interpret. The general approach using “dietary patterns” or “dietary profiles” allows for a better understanding of the relationship between nutrition and health. Along these lines, the various available methods of analysis have been grouped into 2 principal categories: “a priori” and “a posteriori” (15-19).

The “**a priori**” approach is useful for establishing indices for dietary quality and/or variety as well as for adherence to dietary recommendations (eg, Healthy Eating Index, Dietary Diversity Score, Diet Quality Index, etc.). In turn, the “**a posteriori**” approach relies on cluster and/or factor analysis in order to compute latent factors from the original measured variables, thus accounting for existing correlations and maximizing the explained variance. These analysis methods are promising because they help overcome a large number of limitations entailed in the traditional approaches. The global approach to diet was initially implemented in research on chronic diseases such as cancer, CVD, and diabetes (19-36), where the availability of very large population samples is necessary.

Hence, it is important to implement large prospective studies with sizeable and diverse samples of participants, and to be able to measure and subsequently control for a wide array of potential confounders of the relationships between nutrition and health and/or mortality. The performance of stratified analysis according to prevalent conditions (eg, BMI categories) is also useful in identifying disparities in the relationship between nutrition and health. It is equally important to enroll participants aged 45+ years, given the outcomes of interest (ie, morbidity, mortality).

In that context, the Internet is a useful tool for reaching vast and diverse samples of volunteers, including those who are middle-aged, and for collecting on a regular basis ample information which is then treated automatically.

1.3 Internet use in France: number and profiles of users

According to recent data from the European Interactive Advertising Association (EIAA), France is at the top of the list regarding Internet usage across Europe. For example, in May 2008, 32 700 000 French individuals aged 11+ years (ie., 61.9% of the population in that age range) were connected to the Internet at home, work, public places or private locations. Half of all French households dispose of an Internet connection, and Internet users spend, on average, 13 hours/week online (versus about 9 hours/week in 2004). This exceeds the European average which is 10h15/week and represents an increase of approximately 17% in comparison with the time spent in 2004. Just over the past 2 years, Internet usage has increased by 43%.

According to data from June 2006, the profile of Internet users suggested that it is indeed feasible to reach a volunteer sample that closely resembles the general population in terms of age (and especially senior citizens), SES, and geographic distribution.

France: Profiles of Internet users (June 2006)	
Gender	% of all Internet users
Men	53 %
Women	47 %
Age categories	% of all Internet users
15-24 y	24 %
25-34 y	25 %
35-49 y	29 %
50+ y	22 %
Region	% of all Internet users
Ile-de France	22 %
Countryside	78 %
SES	% of all Internet users
High SES	38 %
Low SES	29 %
Unemployed	5 %
Students	19 %
Retired	10 %

Source: Ipsos Média

Update: 29/08/2006

In France, one out of every four Internet users is older than 55 years and French seniors have the most Internet usage across Europe, relying on novel technologies as reported by the EIAA and SRI. They on average spend 9.8 hrs/week online, whereas the European average is 8.8 hrs/week. In addition, 83% of senior Internet users rely on high-speed ADSL connections (the corresponding European average is 75%). Finally, there is increased popularity of online forum use by French senior Internet users (>200% increase since 2005), with 21% using such forums at least on a monthly basis.

Profile of French Internet users aged 55+ years	
Indicators	%
Proportion of Internet users aged >55 y	25
Internet users with ADSL connections	83
Internet users regularly sending e-mail	87
Internet users who read online forums \geq once/mo	21
Internet users who use instant messaging	37
Internet users who use IP phone calls (voIP)	21

November 2009 update: 34.7 million French individuals aged 11+ years (ie, 65% of the population in that age range) were connected to the Internet at home, work, public places or private locations.

2011 update: According to OCDE, 79.6% of the French adult population has access to the Internet.

The recruitment of volunteer research participants via the Internet has many advantages as regards progress in the field of nutritional epidemiology in France. It allows a regular follow-up of very large

and diverse samples and entails a relatively low cost in comparison to the amount and quality of the collected information. Large national, nutrition-focused cohorts are currently lacking in France, while prior cohorts have been launched for the assessment of narrowly-defined topics. Thus, the NutriNet-Santé Study offers the opportunity to broaden and advance the field of nutritional epidemiologic research.

1.4. Overarching aims of the NutriNet-Santé study

The overarching aims of the NutriNet-Santé study are the following:

1) Identify the role of nutritional factors (dietary habits, intake of food groups and specific nutrients) in cancer and CVD morbidity/mortality (and conditions such as obesity, diabetes, hypertension, dislipidemia), aging and quality of life.

Numerous hypotheses pertain to the role (ie, protective or detrimental) of nutrient and/or food intake in the risk for chronic diseases (CVD, cancer) and in health status in general. Existing prospective epidemiological studies are few, rely on relatively small samples, utilize a relatively short follow-up, and do not employ statistical adjustment for a sufficiently large number of potential confounders which could increase confidence in the observed relations. In many cases, such relations are only suggested, which does not allow advancing new public health recommendations. As regards the NutriNet-Santé study, its methodology, sample size, and feasible control for major confounding factors (obesity, physical activity, etc.) will permit addressing the plausibility of associations between food/nutrient consumption and health, given existing mechanistic hypotheses. The study will also permit the assessment of the role of certain nutrients, food groups and eating habits (and those related to lifestyle), which are not yet well known or recognized as modulators of disease risk, health status, and/or mortality. Particular attention will be given to the role of nutrition in the risk profile regarding various pathologies.

In turn, the application of “a priori” approaches to assessing the global quality of the diet via specific indices will allow validating and/or ameliorating the basis of existing nutritional recommendations. In addition, the “a posteriori” methods will allow the identification of dietary patterns corresponding to lower or higher disease risk, and will also serve to advance national nutritional recommendations.

2) Create a comprehensive database covering food and nutrient consumption and eating habits (and their determinants), physical activity and the nutritional status of the population, as well as the surveillance of the evolution of food consumption and dietary behaviors.

The implementation of a large cohort, the long follow-up, and the collection on a regular basis of data (eg, on nutrition, physical activity, anthropometrics, SES, psychocultural and sensory aspects) will eventually lead to the development of a unique and comprehensive database on the determinants and the evolution of dietary behaviors on the population-level. Enrollment in the cohort will be open for several years, allowing the ongoing recruitment of volunteers as well as the identification of unique dietary habits and the emergence of new nutrition-related challenges.

2. Objectives of the NutriNet-Santé Study

2.1 Main objectives

The main objectives pertain to the investigation of the relationship between dietary behaviors, food and nutrient intake and CVD/cancer mortality. The role of dietary behaviors, food and nutrient intake in the incidence of the following diseases/conditions will also be investigated:

- Cancer
- CVD
- Overweight and obesity
- Type 2 diabetes
- Hypertension
- Metabolic syndrome
- Depression
- Migraine
- Rheumatoid arthritis
- Cognitive decline
- Quality of life
- Aging
- ...

2.2 Additional objectives

- Investigate multiple determinants (eg, socioeconomic, psycho-cognitive, cultural, sensory, biological, genetic) of eating habits, dietary patterns, nutritional status, and overall health status;
- Investigate the relations of dietary/nutrient intake and eating habits with various biomarkers, using selected subsamples of participants;
- Provide a long-term follow-up and description of population-level trends as regards dietary behaviors and nutritional status;
- Evaluate the impact of public health efforts in terms of perception, effectiveness and acquired knowledge.

Given the quantity and quality of the collected information as well as the size of the study sample, the NutriNet-Santé study is well positioned to serve as a large databank regarding population-level information on nutrition and health in France, and will also represent one of the largest epidemiological data sources worldwide.

3. Exposure and outcome assessment

3.1 Main exposures

- Consumption of nutrients (macro- and micro-nutrients)
- Consumption of food (food groups)
- Eating habits and dietary patterns (based on “a priori” indices and “a posteriori”-defined profiles)
- Physical activity

3.2 Main outcomes

- Mortality from cancer and CVD over a scheduled follow-up of 10 years. National registries will be used for ascertaining mortality cases and causes of death
- Morbidity: 10-year disease incidence

3.3 Hypotheses, sample size, and statistical power

In order to calculate the needed sample size, information must be available on the expected effect size of dietary behaviors (consumption of nutrients, foods, dietary patterns) on the 5- and 10-year morbidity and mortality rates. Other useful information includes cancer and CVD incidence and mortality rates as well as current prevalence of different dietary practices in France. The sample size calculation is based on the population aged 45-74 years, even though the cohort will be open to younger participants as well, in order to assess long-term disease risk. Thus, a large and comprehensive database will be created. From a statistical point of view, the possibility of type 1 error will be set at 5% and the statistical power will be set at 95%.

It has been established that tens of thousands of participants would be needed in order to properly assess the protective effect of certain foods (eg, fish, fruit and vegetables) which have been linked to a 15-35% reduction in CVD and cancer mortality. Given the epidemiological evidence, it is reasonable to expect a 20% reduction in the 5-year mortality from cancer and CVD in those following a balanced diet. For example, the regular consumption of fish (>2 times/week versus <2 times/week) has been associated with a 17% reduction in coronary heart disease mortality (37). In a sample where 60% of the participants consume fish <2 times/week (as observed in the SU.VI.MAX study), a total sample of 153 697 individuals would be needed in order to observe the same risk reduction. In turn, a 36% reduction in mortality from cancer has been observed in those consuming at least 5 fruits and vegetables a day in comparison with those consuming fewer than 5 (plasma vitamin C =50 mmol/L) (38). Thus, in a population where 40% the participants consume <5 fruits and vegetables a day (as observed in the SU.VI.MAX study), a total sample 13 319 individuals would be needed in order to observe the same risk reduction. Apart from these two examples based on results found in the literature, it is theoretically possible to calculate a necessary sample size by varying the percentage of those exposed or the percent increase or decrease in disease risk. For example, for an expected number of cancer cases of 4 720 occurring over a 5-year period, the needed number of participants would be as follows:

		% risk increase		
		10	15	30
% exposed in the population	10	790 045	360 417	97 090
	40	306 110	141 892	40 035
	50	296 999	138 357	39 580

Likewise, for an expected number of CVD cases of 10 085 occurring over a 10-year period, the needed number of participants would be as follows:

		% risk increase		
		10	15	30
% exposed in the population	10	361 364	164 781	44 329
	40	140 101	64 934	18 314
	50	135 958	63 335	18 116

Thus, a sample of 250 000 participants (half of whom are aged 45+) would provide the opportunity to demonstrate potential disparities in the nutrition-health relationship with sufficient statistical power, while taking into account incomplete data and loss to follow-up. Such a sample size (with participants aged 18+ years) would indeed well position the NutriNet-Santé study for meeting its research objectives. Given the results of our online pilot study, we expect a 60% enrollment rate (ie, sample size >500 000 registered individuals, with a target of 300 000 active participants). Detailed information on incidence and causes of mortality will be obtained from the national vital statistics registry. Regarding the other study outcomes, we expect a 20% attrition (ie, 240 000 participants over 5 years).

4. Materials and methods

The NutriNet-Santé is a large observational prospective cohort study with a planned follow-up of at least 10 years.

4.1 Recruitment

The study's recruitment target is set at 500 000 adult Internet users (called "Nutrinautes"). The participants are aged 18+ years (with a target of at least 250 000 individuals older than 45 years) and are recruited via a vast, nation-wide, recurrent multimedia campaign, as follows:

- In April-May 2009, a vast multimedia campaign was launched (including television, radio, national and regional newspapers, posters, Internet). Its objective was to let the public know about the study, to launch the call for volunteers, and to provide information on the modalities of enrollment (the study website: www.etude-nutrinet-sante.fr was provided).
- The call for volunteers is maintained on the websites of all institutional partners of the study.
- Billboard and bulletin board advertising campaign is undertaken via various professional channels (physicians, pharmacists, dentists, town hall administrators, etc.).

Enrollment in the study will remain open for its entire duration. Progress reports will be disseminated on a regular basis via national and local media in order to boost enrollment of certain subgroups (ie, those residing in underrepresented regions).

4.2 Registration

Interested individuals can register on the study website (www.etude-nutrinet-sante.fr) by clicking on the "Enroll now" button provided on the home page.

The screenshot shows the homepage of the NutriNet-Santé study website. At the top, there is a green banner with the study logo on the left, the text "500 000 nutrinautes pour étudier les relations entre la nutrition et la santé" in the center, and a login section on the right titled "ACCÈS MEMBRE" with fields for "Identifiant" and "Mot de passe" and a "Je m'inscris" button. Below the banner is a navigation menu on the left with links such as "Accueil", "Pourquoi l'étude NutriNet-Santé?", "Objectifs de l'étude", "L'étude NutriNet-Santé en bref", "Qui peut participer?", "Pourquoi participer?", "Comment s'inscrire?", "Mode d'emploi", "Qui coordonne?", "Partenaires ? Qui finance?", "Actualités de l'étude", "NutriNet-Santé dans la presse", and "Foire aux questions". The main content area features a large heading "Bienvenue sur le site de l'ÉTUDE NUTRINET-SANTÉ" followed by a sub-heading "Une cohorte de 500 000 nutrinautes pour faire progresser la recherche publique sur les comportements alimentaires et les relations Nutrition-Santé". Below this is a video player showing a woman speaking, with a "Je m'inscris" button and a link "Afficher le texte de la vidéo". At the bottom, there is a list of funding partners including the Ministère de la Santé et des Sports, l'INPES, l'InVS, l'Université Paris 13, l'INSERM, l'INRA, le Cnam, and la Fondation pour la Recherche Médicale. The footer also includes logos for "powered by @ onematis med4sport" and "Avertissement légal".

- *In order to enhance motivation and to provide support to the participants when filling out the questionnaires, the website includes messages of encouragement as well as video clips, animated demonstrations and helpful text.*

At the beginning, individuals will be presented with some text regarding the study's strengths and limitations and an informed consent form to be signed electronically. Next, the study enrollment page will appear containing the following fields:

- Family name (surname)
- Maiden name (for married women)
- First (given) name
- Sex
- Date of birth
- Country of birth
- Place of birth (town, city)
- Permanent home address
- Primary and secondary e-mail addresses
- Family physician (name and address, if available)
- Other family members participating or wishing to participate in the NutriNet-Santé study (name and identification number, if available)
- Participation in other medical/public health research studies (name(s), if available)

The registration process is considered complete once the informed consent form has been signed and the registration form has been filled in.

In order to prevent/limit potential e-mail entry errors, a second box will appear where individuals can confirm their e-mail address while becoming aware of the importance of this information. Then, a welcome e-mail message is sent, including their login/password information (along with the possibility to choose a different, personally-meaningful password). The individual is then given 21 days (an e-mail reminder is sent after 15 days) in order to log-in and complete the baseline enrollment kit.

4.3. Enrollment in the study

Inclusion is considered complete once the volunteer has filled out the baseline set of 5 questionnaires (eg, "Baseline kit"). Volunteers are given 21 days to complete this step and counting begins once the 3 days for the 24h dietary records have been randomly assigned. The "Baseline kit" includes the following:

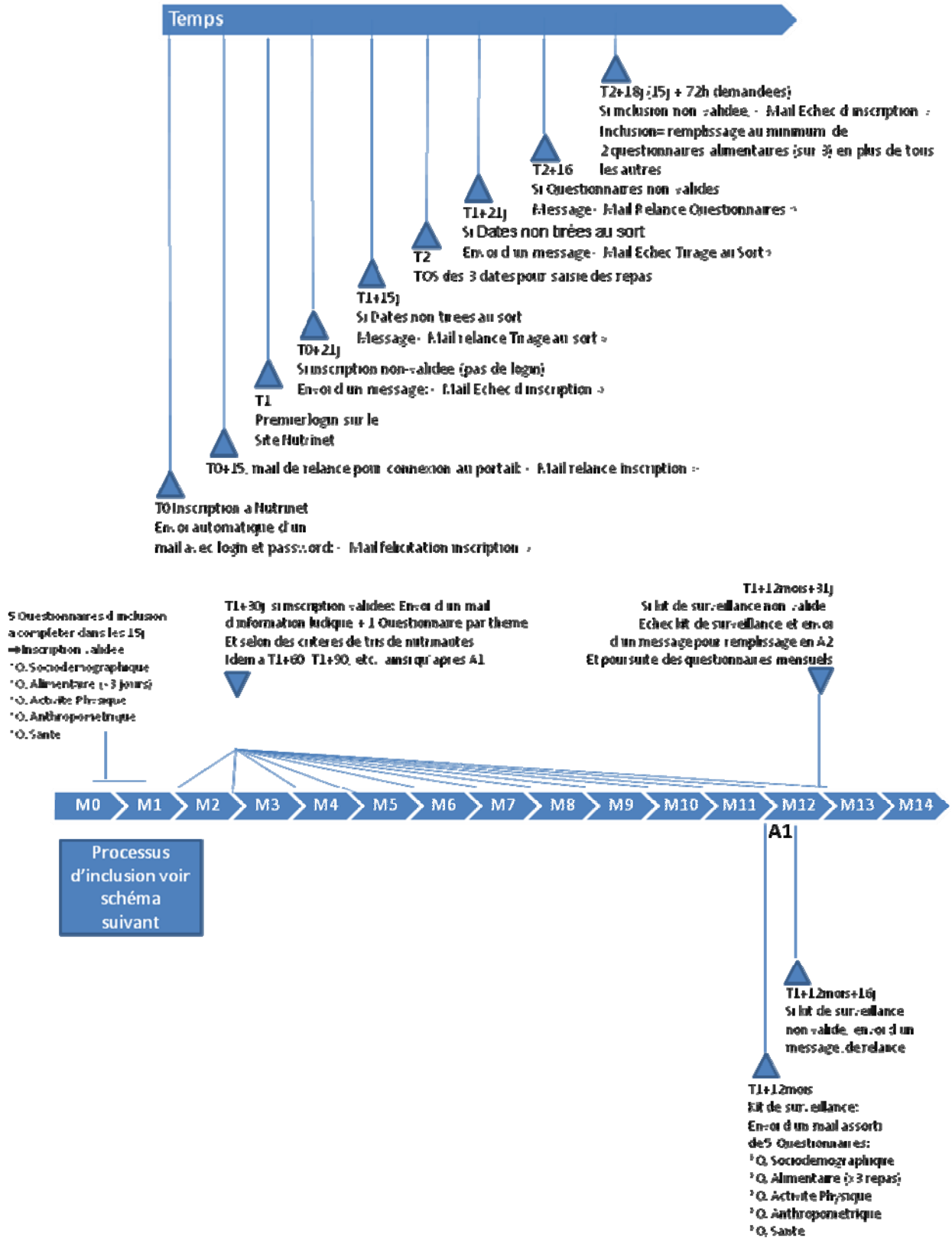
- 3 dietary questionnaires (24h dietary records of which at least 2 should be completed in full);
- Socio-demographic and lifestyle questionnaire;
- Health status questionnaire;
- Anthropometrics questionnaire;
- Physical activity questionnaire

Except for the dietary records, the order in which these questionnaires are completed is left up to each volunteer. As to the dietary records, the randomly assigned dates pertain to 2 weekdays and 1 weekend day, and participants are given the possibility to change one of these three dates.

The registration and enrollment steps, along with the response timeframe and the reminder e-mails are illustrated below.

Given the results of our pilot study, we expect a 60% enrollment rate (ie, sample size >500 000 registered individuals, with a target of 300 000 active participants).

4.4. Data collection timeframe



Over the course of the follow-up, the “Nutrinautes” receive automated e-mail messages regarding any upcoming questionnaires to be completed via a built-in link. Such questionnaires are designed to be completed in ≤ 20 min. In addition, participants will receive newsletters on the study’s progress, updated cohort information, as well as news in the field of nutrition and health that has been added to the website www.etude-nutrinet-sante.fr.

4.5 Data collection: modalities and type of information collected

4.5.1 Information collected as part of the «Baseline Kit»

4.5.1.1 Dietary questionnaire

At inclusion, three 24h dietary records are to be filled by the participants. These records will detail their food consumption (type and quantity of the consumed foods and beverages) along with the corresponding settings (time and place).

General description

The program has the following functions adapted for the NutriNet-Santé study:

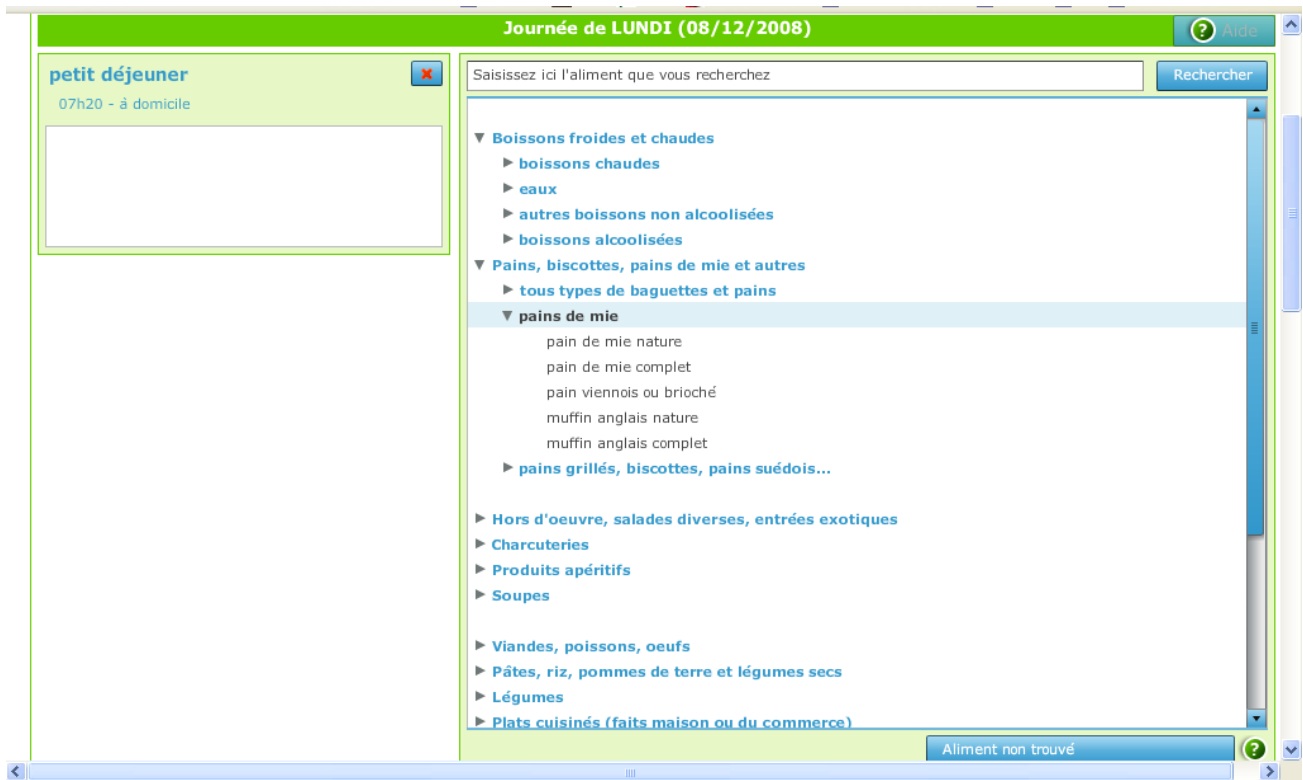
- **Completeness:** for all food and beverage items consumed over a period of 24h, the participants can provide detailed information on the type (including recipes), actual quantities, and corresponding settings (time and place).
- **Interactivity:** the program is designed in a user-friendly format adapted for online self-administration. A system of alerts is active during the completion process and is supplementary to the user’s guide available for download in PDF format.
- **Accessibility:** the program is accessible from any location and at any time, and provides instant server-based back-up.
- **Control:** the program has a built-in control system which helps participants minimize the chance of forgetting to report consumed items.

Each reported consumption is linked to a certain time of day: 3 main meals (breakfast, lunch, dinner) and intake in-between meals (called «Other»). Each intake setting is reported in terms of place and time.

For each intake, participants use the following reporting procedure:

1. Enter the time (hour) of intake (modifiable)
2. Enter the place of intake (“at home,” “at a friends’ or family’s house,” “cafeteria,” etc.)
3. Enter all consumed food items without their quantities using the following 2 modalities:
 - “Food browser”: the types of food are organized according to categories representing a “tree” where volunteers can select the appropriate branch before reaching the “leaves” (ie, generic food items);
 - “Search engine”: the food item is searched via a “matching” process which ignores spelling errors and/or incomplete typing. The results are provided in the form of a “tree” just like the “food browser” case.
 - In case the volunteer cannot find a particular food item, he/she has the possibility to enter it manually, including quantity consumed (via a scroll-down list with measurement units - ml, g, etc.).

For example:



4. For certain food items, the brand name can be specified (ie, from a drop-down list of popular brands), which would ensure more accurate data on the item's nutritional value. Brand names (if not on the list) can also be entered manually.
5. For each entered food item, the program displays a list of commonly consumed side dishes. It is also possible from there to go back to the main list, but in that case the association "food item-side dish" will not be saved. A message is also displayed indicating that such side dishes are meant to provide data on food seasoning/additives, such as salt and fats.
6. Following the entry of food intake information, the program will automatically match the entered items with those commonly consumed (ie, water, bread, etc.). In order to prevent/limit omissions, as a reminder tool the program displays a list of corresponding items that have not been entered.
7. Next, the consumed quantity of each item is entered, as follows: portion sizes for each item are displayed in pictures to assure an accurate estimation of the consumed quantity (this step relies on a validated photograph manual developed for the SU.VI.MAX study). Participants can also choose intermediate or extreme portion sizes, which are clearly displayed.

For example:



The photograph manual (SU.VI.MAX study) also includes various volumes/amounts of fluids. In NutriNet-Santé, however, partial volumes (1/4, 1/8, etc.) have been increased.

The possibility to enter weight (in g) is also available.

8. When the dietary intake information is entered for the particular day, the “Nutrinaute” is asked to indicate whether that represents typical or unusual intake. If the latter, he/she is asked to indicate whether the intake was more or less than usual and also to indicate the reason(s).

For any meal, the following “tree” structure is presented:

- Water, other cold and hot beverages
- Bread, biscuits, and other similar items
- Various salads, exotic dishes, appetizers
- Processed meat (cold cuts)
- Aperitifs
- Soups

- Meat, fish, eggs
- Pasta, rice, potatoes, legumes
- Vegetables
- Cooked dishes (homemade or take-out)
- Fast food, pizza, sandwiches, pies and others

- Dairy products (milk, yogurt and cheese)
- Sweets (breakfast, dessert, pastries)
- Fruit

- Dressing, seasoning, cooking oil, added fat, salt
- Sugar, flour
- Diet products

Architecture of the site

The questionnaire modules regarding data entry, presentation and treatment are separated. The program relies on 2 different, relational databases: one for dietary data (INM-DB) and the other for participant data (PREC-DB). Moreover, the program operates on a server and relies on multiple support systems in order to ensure data quality control. The website interface is implemented in the FLEX programming language. There is interaction, controlled by the server, between the databases and the interface. All programs are developed in JAVA, whereas the operating system for the servers is Linux.

All entered data (along with their inter-relationships) are backed up. Nonetheless, any data deleted by the participant while completing the dietary questionnaire is likewise deleted from the database.

Interface

The interface represents RIA, implemented in FLEX and AJAX which enable the following:

- Drag-and-drop: the item is picked up and dropped into the “quick list”;
- Analogous interface for PC/MAC
- No page reload
- Data entry can be interrupted and later resumed without losing any information (draft saving)

- Browsers: Microsoft Internet Explorer 6/ Mozilla Firefox 2
- Installed Plugin Flash: Adobe Flash Player 9.0.124
- Software download links are also provided to the participants.

2013 Update: In May 2013, an accurate salt intake estimation module was implemented.

4.5.1.2 Sociodemographic and lifestyle questionnaire

The following information is collected :

- Marital status
- Number of children and grand-children
- Number and relationship of individuals in the household
- Employment situation
- Current or most recent occupation
- Educational level
- Spouse (occupation, education, etc., if applicable)
- Income level
- Smoking status (type, quantity, duration, passive smoking, etc.)
- Alcohol consumption (type, quantity, frequency, etc.)
- Seafood consumption

4.5.1.3 Health status questionnaire

The following information is collected :

- Reproductive history (women only): monthly periods, pregnancies, menopause, contraception, hormone replacement therapy, etc.
- Medical history: cancer, CVD, hypertension, diabetes, hyperlipidemia, neurological and psychiatric disorders, respiratory diseases, bone and joint disorders, retinopathy, nephropathy, endocrine system disorders (thyroid, adrenal gland, etc.)
- Hospitalization due to cancer or CVD (most recent hospital admission date, name of hospital, department, location)
- Contact information for specialists
- Medication use
- Dietary supplement use
- Family disease history (1st degree relatives): cancer, CVD, hypertension, diabetes, hyperlipidemia, neurological and psychiatric disorders, respiratory diseases, bone and joint disorders, retinopathy, nephropathy, endocrine system disorders (thyroid, adrenal gland, etc.)
- Causes of death of 1st degree relatives (when applicable)

4.5.1.4 Anthropometrics questionnaire

The “Nutinautes” are asked to measure their weight and height

They are also asked to provide the following information :

- Recent weight change
- Weight fluctuation history
- Dieting (type and reason, duration)
- Intentional weight loss
- Observing own weight
- Body image (Sorensen scale)

In order to ensure standardized measurement, instruction on how to perform the anthropometric measurements will be supplied to the participants. An instructional video is also available on the website: www.etude-nutrinet-sante.fr.

4.5.1.5 Physical activity and sedentary behavior questionnaire

Physical activity and sedentary behaviors are assessed via the online version of the “International Physical Activity Questionnaire” (IPAQ). Sedentary behaviors are assessed by the time spent in front of a TV, computer, etc. The IPAQ can be used to categorize physical activity into 3 levels of intensity (walking, moderate and vigorous), expressed in number of days per week and duration per day.

The physical activity data can be used to calculate weekly energy expenditure (relying on norms provided by the IPAQ development team) via metabolic equivalents (MET) in min/week.

MET (min/week) estimates according to the intensity of physical activity (IPAQ):

Physical activity	Frequency days/week (D=day)	Mean duration/day (in min) (T= time)	Metabolic equivalents (MET-minutes/week)
Vigorous	D1	T1	$MET1 = 8.0 \times D1 \times T1$
Moderate	D2	T2	$MET2 = 4.0 \times D2 \times T2$
Walking	D3	T3	$MET3 = 3.3 \times D3 \times T3$

Using METs, participants’ physical activity level can be classified into “low,” “moderate” and “high.” A decision was made to retain the “moderate” category of the IPAQ as the minimum acceptable level in accordance with the PNNS objectives regarding physical activity. The proportion of participants categorized as having a “high” level of physical activity is also noted.

Definition of physical activity categories as applied in IPAQ

Categories of physical activity	Definition
High	<ul style="list-style-type: none"> • $D1 \geq 3$ and $MET1 \geq 1\,500$ or • $D1 + D2 + D3 \geq 7$ and $MET1 + MET2 + MET3 \geq 3\,000$
Moderate	<ul style="list-style-type: none"> • $D1 \geq 3$ and $T1 \geq 20$ or • $D2 \geq 5$ and $T2 \geq 30$ or $D3 \geq 5$ and $T3 \geq 30$ or • $D1 + D2 + D3 \geq 5$ and $MET1 + MET2 + MET3 \geq 600$
Low	<ul style="list-style-type: none"> • No physical activity reported • None of the above criteria

4.5.2 Additional monthly questionnaires

Each month, the participants will receive automated e-mail messages regarding the study progress as well as any upcoming questionnaires to be completed via a built-in link. There will be no more than one questionnaire per month with the due date set as a function of the date of enrollment (ie, on a rolling basis).

In addition, *ad hoc* questionnaires can also be developed and administered to all participants.

4.5.3. Collection of follow-up data

Some of the information collected at baseline will be updated on a regular basis. Follow-up is scheduled for at least 10 years. Each year, on the same date as the initial enrollment, a modified version of the “Baseline kit” will be administered. Modifications will pertain to the following aspects:

- Updated SES and lifestyle questionnaires
- Dietary data will be collected again at 12-month post-baseline and then annually (on the same date as the initial enrollment) over 5 years. It will include three 24h dietary records on randomly selected days (2 weekdays and 1 weekend day) over a 15-day period. Participants will be informed in advance about the assigned dates and will be asked to maintain their usual eating habits
- Physical activity
- Anthropometrics (every 6 months)
- Health status update.

4.5.4. Follow-up and validation of health events

4.5.4.1. Self-reporting of health events

Within the main website, a specific page has been included where participants can self-report the occurrence of health events (diseases, conditions) (ie, “Supplementary personal file”).

Dossier personnel complémentaire
Etude NutriNet-Santé

Vous êtes sur le site réservé aux Nutrinautes vous permettant d'accéder à votre **dossier personnel complémentaire NutriNet-Santé** dédié à la déclaration des événements de santé, des traitements et des examens médicaux.
Pour avoir accès à votre espace personnel complémentaire et confidentiel il est **nécessaire de vous inscrire à nouveau** lors de la première utilisation.
Pour cela cliquez sur [première connexion](#).

Inscription première connexion

Vous avez déjà un mot de passe pour ce site complémentaire

Identifiant Nutrinet:

Mot de passe: [Mot de passe oublié](#)

Connexion

This site allows the “Nutrinaute” to report a new health event, to provide additional information regarding an event already reported, to report new treatment(s), and medical tests (labs, X-rays, etc.).

Dossier personnel complémentaire
Etude NutriNet-Santé

Identifiant: uren
Dernière connexion: 15/04/2010 23:03

[Accueil](#)
[Changer de mot de passe](#)
[Se déconnecter](#)

Déclaration d'un évènement de santé

Cardiovasculaire → Infarctus du myocarde, angine de poitrine ou angor, insuffisance cardiaque...

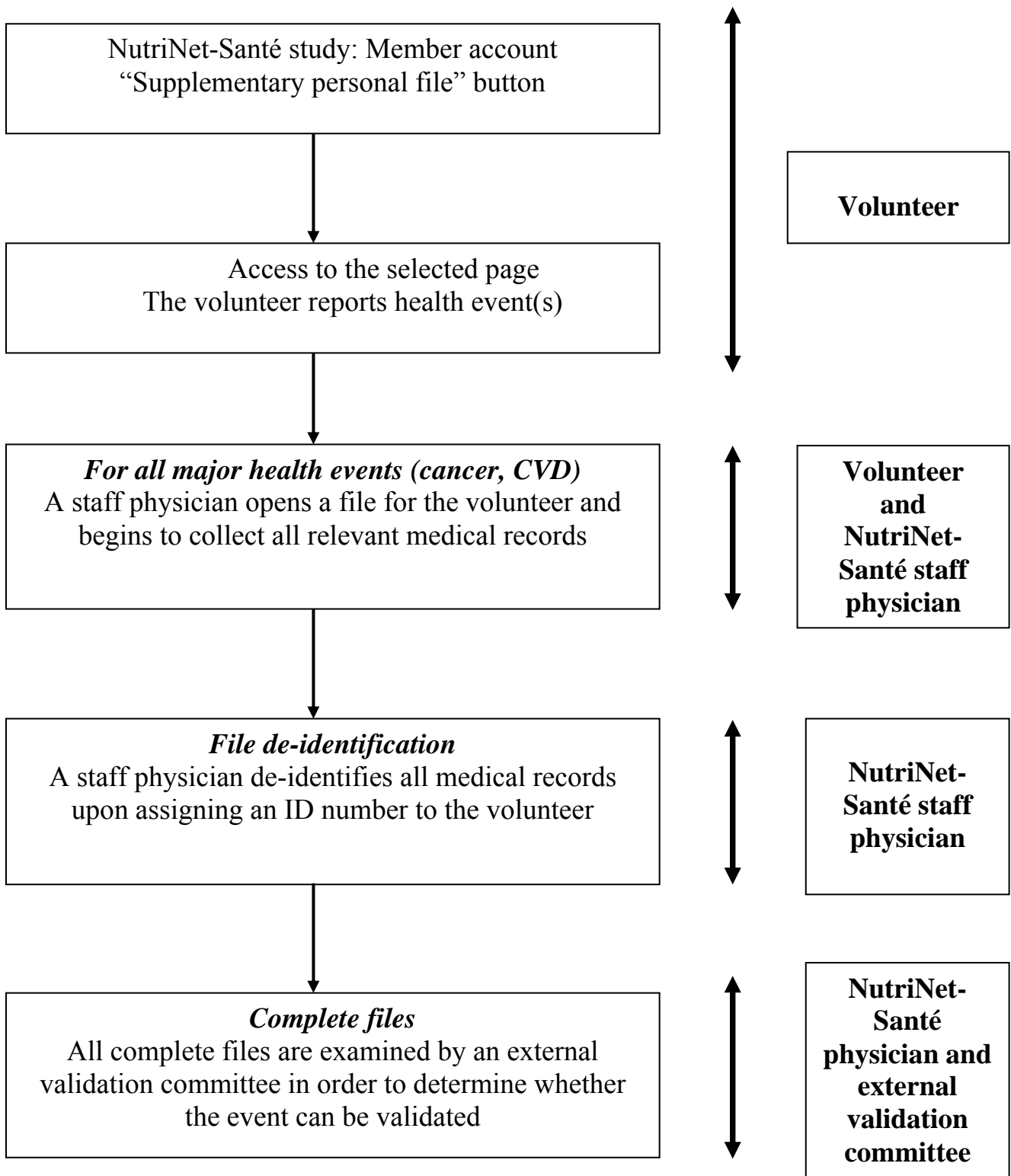
Neurovasculaire → Accident vasculaire cérébral (attaque, hémorragie ou congestion cérébrale, artérite des membres inférieurs...)

Cancer →

Autre → Hypertension, diabète, maladies de la thyroïde, maladies respiratoires, problèmes digestifs, autres pathologies...

[Retour à la gestion de vos évènements](#)

The flowchart below outlines the entire process, from the self-report of a health event by the participant to the validation of such events



For each disease/condition the volunteer must report:

- Disease name
- Date of diagnosis
- Hospital admission information (if applicable)
- Hospital contact information
- Contact information for physician(s) who had provided treatment for the disease/condition

Regarding major health events, the “Nutrinaute” can use any of the following options in order to obtain/provide the relevant medical records:

- If already available, the documents can be sent by postal mail, fax or e-mail
- The participant can designate a NutriNet staff physician to act on his/her behalf for obtaining the records
- Letter templates and useful Internet links are provided to the participant in order to aid him/her in retrieving the relevant medical records

This site allows the “Nutrinaute” to provide additional information regarding an event already reported. In choosing this option, a new window opens, displaying a recapitulation of all reported health events.

4.5.4.2. Health event management

A health status database is created using both data entered on the main website and data entered in the “Supplementary personal file.” This separate database includes a “Physician access” interface, recapitulating all reported health events to be followed up and/or validated as well as progress reports for each event. Treatment and lab test data are also stored and linked to the respective volunteer. The health status database provides access to: all reported information, including all health status questionnaires completed by the volunteer, all reported contact information for physicians and/or hospitals, and the contact information for the volunteer, to be used if needed. For each event, all associated medical records are de-identified, and an ID number is assigned.

Following the self-report of a health event, a staff physician opens a file for the volunteer and begins to collect all relevant medical records, to be presented to the validation committee. The physician highlights all major elements in the file which bear on the potential validation of the event. The following health events are considered “major” and require verification:

CVD:

- Myocardial infarction
- Acute coronary syndrome

Neurovascular diseases:

- Stroke (ischemic, hemorrhagic)

Cancer

- All sites except basal cell carcinoma

Once all relevant records have been collected and examined, the physician marks the files as “Investigation completed.” The next step is presenting the file for discussion/validation at the next scheduled committee meeting.

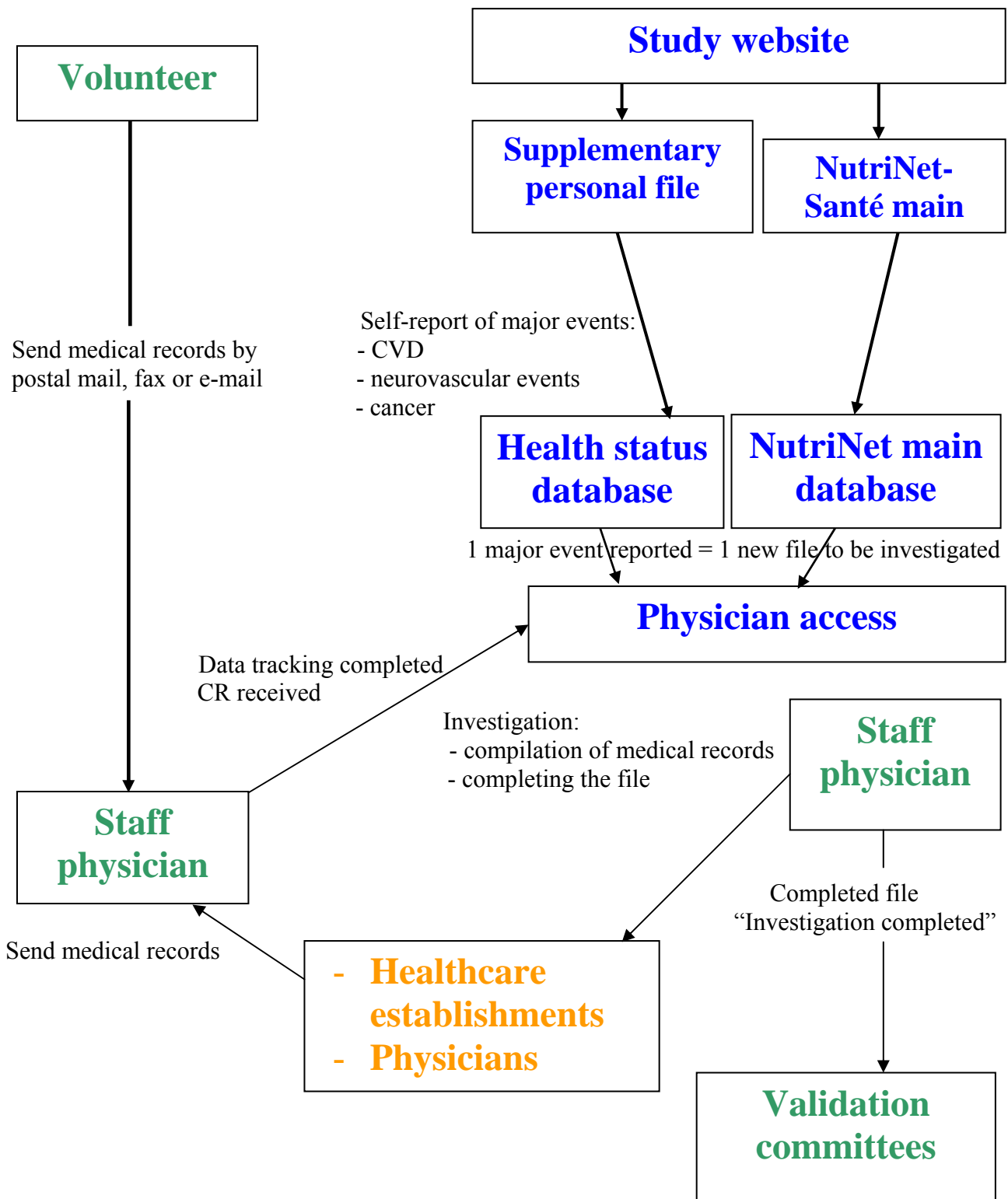
4.5.4.3. Health event validation committees

The following three External Validation Committees have been set up:

- CVD event committee (cardiologists)
- Neurovascular event committee (neurologists)
- Cancer incidence committee (oncologists)

Upon request by the coordinating team and as a function of the number of health event files to be validated, these committees are convened several times per year. At least 2 members of each committee must be present in order to proceed with the validation, employing an international disease coding scheme. These members sign the validation sheet once agreement has been reached as to the type of event, its disease code, and the incidence date.

The flowchart below summarizes the health event validation steps:



4.5.5. Health status follow-up using national medical registries

4.5.5.1. Follow up for mortality and cause of death

Information on vital status and causes of death will be obtained via order # 98-37, granting access to the national RNIPP registry and the CépiDc-INSERM database. Access to such data is relatively easy and straightforward, except for individuals born abroad.

Later on, electronic certification of medical causes of death will be used (INSERM/DGS). Such novel procedures will allow access in real time to information on cause of death, validated by a physician. As regards the physician, the electronic procedures have been designed to closely resemble paper certification while improving data quality. Online support is also available for the completion of the medical certificate, and especially as regards the clear outline of the primary and secondary causes of death. The electronic system has been pilot-tested for 6 months, demonstrating a high degree of user-friendliness. The goal of reducing time lags was clearly achieved.

4.5.5.2 Follow up for health events

Health events follow-up will likewise rely on access to national medical registries. A decree from the State Council and authorization by CNIL currently allow the collection of the social security number (NIR) from participants who explicitly give consent. The NIR, in turn, will allow access to the national CNAMTS database.

- Regarding health events, the SNIIR-AM (*Système national d'information inter régimes de l'assurance maladie*) database could provide access to standardized and coded individual medical records. Information is available on health insurance expenditures, stratified by geographical area, type of expenditure, and type of medical establishment/professional.
- The national health insurance database provides prescription information, however, no data are available on actual diseases/conditions, self-medication, or conditions for which reimbursement was not sought.
- The long-term illness database (ALD, *Affections de Longue Durée*) contains information on individuals whose medical expenses are covered at 100% by the social security system, following approval by the respective service which is charged with coding the disease/condition in question using the International Classification of Diseases (ICD-10).
- The hospitalization database (PMSI) contains information on each hospitalization, including the main diagnosis, age, sex, and the most costly diagnostic tests performed. The diagnosis code is obtained from the ICD-10, whereas diagnostic test codes are obtained from the CCAM.

5. Ancillary protocols

In response to the evolution of research interests and potential collaborations, it is entirely feasible to add a variety of *ad hoc* questionnaires and ancillary protocols pertaining to selected subsamples or to the entire cohort. The “Nutrinautes” will receive automated e-mail messages regarding any upcoming questionnaires to be completed via a built-in link.

Interface features on the website will permit the administrator(s) to screen and select «Nutrinautes» corresponding to specific criteria (ie, all non-smoking men aged 45+ years) and then to provide access to different working groups to the selected subsamples. These subsamples could then be targeted for further study via instruments administered on a rolling or fixed-term basis.

6. Volunteer retention protocol

Each month, the “Nutrinautes” will receive automatic e-mails about the study progress and any upcoming questionnaires. These e-mail messages will contain random facts about the cohort (ie, total weight of all “Nutrinautes” in thousands of tons; total number of consumed bananas, etc.).

- The study website: www.etude-nutrinet-sante.fr will provide information regarding the cohort on a regular basis, and will also provide general information about health and nutrition (“News” blog)
- Special features of the interface permit the “Nutrinautes” to contact the study team via e-mail (without displaying the recipient’s e-mail address) with one of the following subject lines:
 - Contact information change
 - Technical problems
 - Request for information on data confidentiality
 - Questions about the study assessment instruments:
 - Dietary questionnaires
 - Sociodemographic and lifestyle questionnaire
 - Health status questionnaire
 - Anthropometric measures
 - Physical activity
 - Other questionnaire(s)
 - Notice of leaving the cohort

Frequently asked questions (FAQ)

- FAQ by “Nutrinautes” (registered and/or enrolled) concerning the site’s terms of use
- FAQ by the general public
- Different means will be in place in order to create a sense of belonging (and a sense of community) among the participants. Each year, the «Nutrinautes» will receive a «Diploma» following the wedding anniversary tradition of bronze, silver, gold, etc.
- A NutriNet-Santé membership card will be available to download and print.
- A random drawing will take place each year, with the prize of airline tickets for different destinations (depending on tenure in the cohort).
- Each «Nutrinaute» can refer friends/family to the study via pre-formatted e-mail messages.
- A special website called «The Nutrinaute Clan» will allow participants to access a variety of games, quizzes, cartoons and videos on the study, and on the topics of health and nutrition. In addition, participants will have access to forums where they can connect with other members of the cohort and share interests, ideas, etc.

7. Data quality control and instrument quality

A variety of personal information is requested at baseline (ie, address, telephone, family physician, etc.) in order to limit false information in the questionnaires. A number of quality control and validation procedures are in place.

8. IT aspects and data security

Various IT aspects (server structure, data security, etc.) have been presented to and approved by the CNIL. All information concerning the “Nutrinautes” is treated in a confidential manner and only the coordinating team has preferential access to data that have not been de-identified. At all times, the “Nutrinautes” have the right to access and rectify their nominative data (by contacting the coordinating team), as specified in law n°78-17 from January 6, 1978, concerning computer science and privacy, amended by law n°94-548 from July 1, 1994, concerning personal nominative data used for health research purposes.

The raw data collected via the questionnaires (stored in 1 file per questionnaire) will be exported to a server in a CSV format via a SSH (Secure Shell) link. The export will be automatic and performed on a daily basis.

9. NutriNet-Santé Biobank: Collection of biological specimens

The purpose of the NutriNet-Santé Biobank is to collect biological data (as well as certain clinical information) from a subsample of “Nutrinautes,” which will allow a wide array of analyses in the years to come. Such analyses will be aimed at shedding light on numerous mechanistic hypotheses within the nutrition-health relationship.

The development of the protocol regarding the collection, treatment and storage of biological samples takes into consideration the fact that there is no “ideal” way of addressing all potential scientific questions in all areas of research. In addition, it is not possible at present to determine the most useful biomarkers for disease detection. Finally, the protocol reflects technical, analytic, and financial limitations. For example, instead of 24h urine samples, spot urine samples are collected. Whereas no collection of saliva or hair samples is planned, this could be part of future ancillary protocols. Due to limitations, certain sample treatment methods (ie, stabilization at the Local Sample Collection Centers as, for example, treatment of vitamin C with metaphosphoric acid), or certain storage methods (eg, with liquid nitrogen) are also not planned.

For the most part, the Biobank protocol is based on the UK Biobank program (an ongoing national Biobank which includes 500 000 participants and 35 local centres). The selected criteria were defined following extensive consultations within the scientific community and a peer review. Then, a detailed instruction manual was put together in order to ensure that the selected methodology corresponds to the objectives of the Biobank. The Standard Operating Procedures manual of the UK center outlines in detail the methods of sample collection, on-the-spot treatment of the samples, their transportation under strict temperature requirements, the treatment at the central (coordinating) laboratory, the number of aliquots, and sample storage. This manual has been adapted to the needs of the NutriNet-Santé Biobank.

9.1. Population subsample

The purpose of the NutriNet-Santé Biobank is to collect, treat, and store biological data from a subsample of at least 20 000 NutriNet-Santé participants throughout the country. Standardized

procedures as regards collection, treatment, transport and storage of the samples have been implemented.

9.2 Logistics

We employ a Laboratory Information Management System (LIMS) for the management of all aspects of the Biobank, from specimen collection to storage.

9.2.1. Centralized participant scheduling system

Study volunteers are invited to visit a participating Local Sample Collection Center (LSCC) equipped for blood sample draws (approximately 90 LSCC have been opened, with a total of 4 LSCC being open at any one time).

An online, centralized participant invitation/scheduling system coordinated by U557 INSERM (University of Paris 13) is in place. It relies on a dedicated website accessible only by «Nutrinautes». Via that website, the participants receive all pertinent information regarding the objectives of the Biobank, conditions for the blood draws, sample volume, LSCC access, copies of the consent forms, etc.

An e-mail invitation with a built-in link is sent to the «Nutrinautes», allowing them to modify the date/time of the appointment. An automatic confirmation e-mail is sent once the appointment has been confirmed; an additional reminder e-mail is sent 48h before the appointment and will include information on permitted duration of fasting, permitted foods, etc. A final reminder e-mail is sent the day before the appointment.

In order to ensure acceptable participation rates, participants have the opportunity to take advantage of cholesterol (HDL and LDL), triglyceride, and diabetes screening tests (sending of the results for all participants with a special notice in case of abnormal results).

Each day, a total of 15 participants can be treated at the 4 open LSCC (ie, 60 participants per day).

9.2.2 LSCC procedures and centralized specimen collection

The LSCC are in charge of receiving the participants and the field staff in charge of collecting the biological samples (blood and urine), collecting additional clinical data, on-the-spot treatment of the specimens, and transportation to the central laboratory for aliquot production and storage.

9.2.2.1. Coding of test tubes and urine collection vessel

The complete set of samples collected from each participant will be labeled with a barcode. The local clinics are equipped with printers for the label sheets with the barcodes (10 labels per person), which contain a sufficient number of digits allowing the following:

- Identification of the sample collection LSCC;
- De-identified and identical labeling of all samples for each volunteer;
- Identification of the sample collection containers (plastic tubes, EDTA K2 tubes, lithium-heparin tubes, urine collection vessels, etc.)

9.2.2.2. Types of test tubes

Blood samples are collected using the Vacutainer® system. A total of 43 ml of blood (ie, 5 vacutainers) is collected from each volunteer. A spot urine sample is also collected in a vacutainer. A variety of tubes containing different kinds of anti-coagulants or separators are used. The anti-coagulants used (EDTA K2 and lithium-heparin) as well as the obtained sera will make feasible a variety of analyses. Blood from each volunteer are collected into: two 9 ml tubes containing EDTA K2, one 9 ml tube containing lithium-heparin, and two 8 ml plastic clot activator serum separation tubes (as shown in the Table below).

The various vacutainers are used to obtain plasma, serum, buffy coat (for DNA extraction) and red blood cells. The contents in each tube are in turn fractioned into a sufficient number of aliquots in order to permit various analyses to be carried out in the future. In addition, the lithium-heparin tubes and the plastic test tubes contain an inert gel which facilitates separation of the cellular components, thus preventing potential changes in plasma or serum during the interval between centrifugation at the LSCC and aliquot production at the central laboratory.

Type of sample (number of test tubes)	Collection order	Volume collected/tube (ml)	Temperature until centrifugation and aliquot production (°C)
EDTA K2 (1)	1	9	4
Lithium-heparin tube with plasma separation gel plug (1)	2	9	4
Plastic tubes with clot activator and serum separation gel plug (2)	3	8	4
EDTA K2 (1)	4	9	4
Total sample volume		43	
Urine (1)	-	9	4

Table 1 - Collection order, volume collected and temperature until centrifugation and aliquot production.

9.2.2.3. Test tube pre-analysis treatment in LSCC

All LSCC are supplied with refrigerated centrifuges to ensure the separation of plasma or serum within a set and standardized interval, which will prevent any degradation of the samples (eg, contamination from cell lysis if set interval is not observed). For each participant, the following information will be entered in the Biobank database: date, time of sample collection, adherence to recommended fasting period, potential problems during blood draw, etc.

9.2.3. Transportation of specimens to central laboratory

Towards the end of every working day (Monday through Friday), each LSCC dispatches the collected specimens in special parcels, to be delivered at the central laboratory the following day (Tuesday through Saturday). Transportation takes place overnight, to ensure the arrival of test tubes at the central laboratory by the early morning (before 9 am). The daily collection and shipment of the test tubes are carried out by a qualified courier, using special containers (meeting current standards) which ensure a temperature of + 4°C during transportation to the central laboratory.

9.2.4 Procedures implemented at the central laboratory

The central laboratory is based at the Medical School of the University of Paris 13, adjacent to the NutriNet-Santé Biobank. Upon receipt, each barcode is scanned and matched with the data file from the LSCC in order to ensure that all test tubes collected the day before have arrived (a specimen check is also performed). An electronic connection between each LSCC and the central laboratory guarantees that all relevant information is available at the central laboratory even before the arrival and processing of test tubes. All transportation data are recorded at delivery.

A total of 30 aliquots are produced for each participant and stored at - 80°C For security purposes, aliquots are stored in two separate freezers. The aliquot production is carried out by an automated system in order to ensure rapid and high-quality processing, to reduce the possibility of error, to ensure accuracy and traceability of all samples. The reason for which samples are rapidly fractioned into a large number of aliquots is to protect them from freeze-thaw degradation and to store them under optimal conditions for future analyses.

Upon receipt in the laboratory, the EDTA tubes are centrifuged at 2 500 g during 10 minutes at 4°C. The tubes are treated with the aim that all samples are cryopreserved not more than 24h after collection (Table 2).

The aliquots produced from each vacutainer are stored in cryogenic tubes especially engraved with 2-dimensional (2D) barcodes. The process of aliquot production is fully automated to ensure accurate identification of the 2D-labeled tubes and to guarantee traceability from the sampling to the archiving (in the deep freezers) in the Biobank.

Vacutainer tube	Fractions	Number of aliquots frozen at -80°C
EDTA K2 x 2	Plasma	8
	White blood cell layer (buffy coat)	2
	Red blood cells	2
Lithium-heparin with plasma separation gel plug	Plasma	4
Plastic tubes with clot activator and serum separation gel plug x 2	Serum	6
Urine	Urine	8
Total number of aliquots		30

Table 2 - Fractions and aliquote collected from the vacutainers tubes and the spot urine sample.

9.2.5. Specimen storage

Aliquots are stored in the Biobank (located at the University of Paris 13, Bobigny). The various aspects of the Biobank (ie, quantity of tubes to be frozen, rigorous surveillance) necessitate the use of special software (LIMS) operated from local clinics, the central laboratory and the storage unit at the Biobank. Storage trays contain 96 cryogenic tubes with 2D barcodes, which are scanned in the automated system in order to track their exact location in the - 80°C freezers (compartment, rack and position).

The Biobank is physically located on the underground level of SMBH at UREN UMR INSERM U557/INRA U1125/CNAM/University of Paris 13, in an area made available for that purpose by the University of Paris 13. The premises consist of a big hall of 250 m² to house the freezers (a total of 40 freezers can be maintained) and the central laboratory. The Biobank is air-conditioned (to overcome the effect of heat released by the freezers).

9.3 Organization and management of the Biobank

The Biobank is under the direction of Dr. Pilar Galan (DR1 INRA, UMR INSERM/INRA/CNAM/University of Paris 13). It conforms to all legislation regulations in France, such as: reporting the collection to ARH at the Ministry of Research (article L 1243-3; order n°2007-1220), approval by AFSSAPS, CCP, CCTIRS and CNIL. The INSERM Public Health Institute has given its accord for INSERM's support for the study (file n°C09-42, 22/12/2009).

All procedures regarding the function and quality control of the NutriNet-Santé Biobank rely on the following:

- Recommendations published in NF S96-900 from July 2008 on the “Quality of Biological Resource Centers (CRB) - Management System of a CRB and Quality of Biological Resources from Human and/or Microorganism Origin.”
- OCDE guidelines regarding best practices among biological resource centers (OECD, April 2007).

10. RÉFÉRENCES

- (1) <http://www.cepidc.vesinet.inserm.fr/>
- (2) http://www.invs.sante.fr/presse/2008/communiqués/cancer_evolution/DPCancer21FEV08bd.pdf
- (3) Fédération Nationale des Observatoires Régionaux de la Santé. Le cancer dans les régions de France. Mortalité. Incidence. Affections de longue durée. Hospitalisations. 2006 (May)
<http://www.fnors.org/fnors/ors/travaux/synthesekcer.pdf>
- (4) Fédération Nationale des Observatoires Régionaux de Santé – Les maladies cardio-vasculaires dans les régions de France. Mortalité, affections de longue durée, hospitalisations, registres. 2006 (May).
- (5) Unité de Surveillance et d'Epidémiologie Nutritionnelle (USEN). Étude nationale nutrition santé (ENNS, 2006) – Situation nutritionnelle en France en 2006 selon les indicateurs d'objectif et les repères du Programme National Nutrition Santé (PNNS). Institut de Veille Sanitaire, Université de Paris 13, Conservatoire National des Arts et Métiers, 2007. 74 p. Available at: www.invs.sante.fr;
http://www.invs.sante.fr/publications/2007/nutrition_enns/RAPP_INST_ENNS_Web.pdf.
- (6) <http://www-good.ibl.fr/accueil/presse/html/campagne-obe-0108/docs/ObEpi%202006.pdf>
- (7) <http://www.insee.fr/fr/ffc/ipweb/ip1123/ip1123.html>
- (8) Caisse Nationale de l'Assurance Maladie. Dossier Epidémiologie du Diabète Traité. 2007 (June).
http://www.ameli.fr/fileadmin/user_upload/documents/Etude_diabete_2000-2005.pdf
- (9) Bonaldi C, Romon I, Fagot-Campana A – Impacts du vieillissement de la population et de l'obésité sur l'évolution de la prévalence du diabète traité : situation de la France métropolitaine à l'horizon 2016. Bull Epidémiol Hebd 2006;7:69-73.
- (10) <http://www.invs.sante.fr/entred/>
- (11) <http://www.who.int/mediacentre/factsheets/fs138/fr/index.html>
- (12) Ramaroson H, Helmer C, Barberger-Gateau P, Letenneur L, Dartigues JF; PAQUID. Prévalence de la démence et de la maladie d'Alzheimer chez les personnes de 75 ans et plus : données réactualisées de la cohorte PAQUID. Rev Neurol (Paris). 2003;159(4):405-11.
- (13) Ankri J. Epidémiologie des démences et de la maladie d'Alzheimer. Bull Epidémiol Hebd 2006 N°5-6 ; 4244-39.
- (14) Qiu C, De Ronchi D, Fratiglioni L. - The epidemiology of the dementias: an update. Curr Opin Psychiatry. 2007;20(4):380-5.
- (15) Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. Curr Opin Lipidol 2002; 13(1):3-9.
- (16) Jacques PF, Tucker KL. Are dietary patterns useful for understanding the role of diet in chronic disease? Am J Clin Nutr 2001; 73(1):1-2.

- (17) Jacobs DR, Jr., Steffen LM. Nutrients, foods, and dietary patterns as exposures in research: a framework for food synergy. *Am J Clin Nutr* 2003; 78(3 Suppl):508S-513S.
- (18) Kant AK. Dietary patterns and health outcomes. *J Am Diet Assoc* 2004; 104(4):615-635.
- (19) Waijers PM, Ocké MC, van Rossum CT, Peeters PH, Bamia C, Chloptsios Y, van der Schouw YT, Slimani N, Bueno-de-Mesquita HB. Dietary patterns and survival in older Dutch women. *Am J Clin Nutr*. 2006;83(5):1170-6.
- (20) Mizoue T, Yamaji T, Tabata S, Yamaguchi K, Shimizu E, Mineshita M et al. Dietary patterns and colorectal adenomas in Japanese men: the Self-Defense Forces Health Study. *Am J Epidemiol* 2005; 161(4):338-345.
- (21) Fung TT, Hu FB, Holmes MD, Rosner BA, Hunter DJ, Colditz GA et al. Dietary patterns and the risk of postmenopausal breast cancer. *Int J Cancer* 2005; 116(1):116-121.
- (22) Fung T, Hu FB, Fuchs C, Giovannucci E, Hunter DJ, Stampfer MJ et al. Major dietary patterns and the risk of colorectal cancer in women. *Arch Intern Med* 2003; 163(3):309-314.
- (23) Kerver JM, Yang EJ, Bianchi L, Song WO. Dietary patterns associated with risk factors for cardiovascular disease in healthy US adults. *Am J Clin Nutr* 2003; 78(6):1103-1110.
- (24) Osler M, Heitmann BL, Gerdes LU, Jorgensen LM, Schroll M. Dietary patterns and mortality in Danish men and women: a prospective observational study. *Br J Nutr* 2001; 85(2):219-225.
- (25) Schulze MB, Hoffmann K, Kroke A, Boeing H. Dietary patterns and their association with food and nutrient intake in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam study. *Br J Nutr* 2001; 85(3):363-373.
- (26) Terry P, Suzuki R, Hu FB, Wolk A. A prospective study of major dietary patterns and the risk of breast cancer. *Cancer Epidemiol Biomarkers Prev* 2001; 10(12):1281-1285.
- (27) Terry P, Hu FB, Hansen H, Wolk A. Prospective study of major dietary patterns and colorectal cancer risk in women. *Am J Epidemiol* 2001; 154(12):1143-1149.
- (28) Maskarinec G, Novotny R, Tasaki K. Dietary patterns are associated with body mass index in multiethnic women. *J Nutr* 2000; 130(12):3068-3072.
- (29) Zhuo XG, Watanabe S. Factor analysis of digestive cancer mortality and food consumption in 65 Chinese counties. *J Epidemiol* 1999; 9(4):275-284.
- (30) Slattery ML, Boucher KM, Caan BJ, Potter JD, Ma KN. Eating patterns and risk of colon cancer. *Am J Epidemiol* 1998; 148(1):4-16.
- (31) Osler M, Schroll M. Diet and mortality in a cohort of elderly people in a north European community. *Int J Epidemiol* 1997; 26(1):155-159.
- (32) Cottet V, Bonithon-Kopp C, Kronborg O, Santos L, Andreatta R, Boutron-Ruault MC et al. Dietary patterns and the risk of colorectal adenoma recurrence in a European intervention trial. *Eur J Cancer Prev* 2005; 14(1):21-29.
- (33) Kim MK, Sasaki S, Otani T, Tsugane S. Dietary patterns and subsequent colorectal cancer risk by subsite: A prospective cohort study. *Int J Cancer* 2005; 115(5):790-798.

- (34) Montonen J, Knekt P, Harkanen T, Jarvinen R, Heliovaara M, Aromaa A et al. Dietary patterns and the incidence of type 2 diabetes. *Am J Epidemiol* 2005; 161(3):219-227.
- (35) Sonnenberg L, Pencina M, Kimokoti R, Quatromoni P, Nam BH, D'Agostino R et al. Dietary patterns and the metabolic syndrome in obese and non-obese Framingham women. *Obes Res* 2005; 13(1):153-162.
- (36) Fung TT, Stampfer MJ, Manson JE, Rexrode KM, Willett WC, Hu FB. Prospective study of major dietary patterns and stroke risk in women. *Stroke* 2004; 35(9):2014-2019.
- (37) Whelton SP, He J, Whelton PK, Muntner P. Meta-analysis of observational studies on fish intake and coronary heart disease. *Am J Cardiol.* 2004;93(9):1119-23.
- (38) Khaw KT, Wareham N, Bingham S, Welch A, Luben R, Day N. Combined impact of health behaviours and mortality in men and women: the EPIC-Norfolk prospective population study. *PLoS Med* 5(1): e12. doi:10.1371/journal.pmed.0050012.