# Prioritization of the biomarkers to be analyzed in the French biomonitoring program

Clemence Fillol, Robert Garnier, Jean-Ulrich Mullot, Céline Boudet, Isabelle Momas, Louis Rachid Salmi, Stéphanie Vandentorren

Research Article Open Access

Clémence Fillol, Robert Garnier, Jean-Ulrich Mullot, Céline Boudet, Isabelle Momas, Louis Rachid Salmi, Stéphanie Vandentorren*

# Prioritization of the biomarkers to be analyzed in the French biomonitoring program

**Abstract:** The aim of this work was to develop a comprehensive prioritization method to select the biomarkers to be monitored in the French national biomonitoring program. The first step consisted in building an exhaustive list of biomarkers. The next step involved prioritizing the initial list of biomarkers according to specific scientific questions about human exposure to chemicals in the environment, and meet logistical, feasibility and budgetary constraints. The Delphi consensus method was used to prioritize biomarkers and was developed in three phases: i) the definition of relevant criteria for selecting biomarkers; ii) the prioritization of the biomarker list based on these criteria and iii) the validation of the list by the stakeholders. Among the eight relevant criteria for selecting biomarkers, hazard identification and social perception were the highest-rated and lowest-rated criteria, respectively. After scoring each criterion for each group of biomarkers, and discussing the relative ranking of each group during a round table meeting, the final prioritized list obtained contained both historic (e.g. dioxins or lead) and emerging substances (e.g. phthalates, bisphenol A). Combining rigor and flexibility, our method has clearly helped to build a prioritized list shared and supported by many international actors.

**Keywords:** Biomonitoring – Biomarkers – Environmental epidemiology – Delphi process

DOI 10.2478/bimo-2014-0010
Received August 13, 2014; accepted October 16, 2014

## 1 Introduction

A biomarker is defined as "any substance, structure or process that can be measured in the body or its products and that may influence or predict the incidence or outcome of disease" [1]. Different categories of biomarkers measure exposure, effect and susceptibility. It is important to clearly define the framework for using biomarkers (e.g. pollution burden or health risk) in order to establish the most appropriate decisions as regards defining the baseline exposure information to collect, rather than creating conclusions about human health risk [2]. While research projects more often examine effect biomarkers and genetic factors, population studies generally focus on exposure biomarkers. Accordingly, we will discuss only exposure biomarkers in this paper.

Biomarkers provide a direct measure of total exposure to environmental pollutants and integrate the different sources and pathways of exposure. They help answer specific questions regarding the characteristics of exposure to a specific pollutant (nature or conditions of exposure), for example, urinary arsenic can be measured to study recent arsenic exposure, cotinine concentrations in urine help measure tobacco smoke, and lead concentrations in blood helps confirm reduced exposure to lead from environmental sources.

***Corresponding author: Stéphanie Vandentorren:** Institut de veille sanitaire, Département Santé environnement, 12 rue du Val d'Osne, 94415 Saint Maurice, France Tel : +33 1 44 02 08 29 - Fax : +33 1 44 02 06 76, E-mail: Stephanie.VANDENTORREN@ars.sante.fr
**Clémence Fillol:** Institut de veille sanitaire, Département Santé environnement, 12 rue du Val d'Osne, 94415 Saint Maurice, France
**Robert Garnier:** Paris poison center. Hôpitaux universitaires Saint-Louis-Lariboisière-Fernand Widal, AP-HP & Université Paris-Diderot, Paris, France
**Jean-Ulrich Mullot:** Laboratoire d'Analyses de Surveillance et d'Expertise de la Marine, 83000 Toulon, France
**Céline Boudet:** Institut National de l'Environnement Industriel et des Risques, (INERIS), 60 550 Verneuil en Halatte, France
**Isabelle Momas:** Université Paris Descartes, Laboratoire Santé Publique et Environnement – EA 4064, Paris, France
**Louis Rachid Salmi:** Univ. Bordeaux, ISPED, Centre INSERM U897-Epidemiologie-Biostatistique, F-33000 Bordeaux, France
**Louis Rachid Salmi:** INSERM, ISPED, Centre INSERM U897-Epidemiologie-Biostatistique, F-33000 Bordeaux, France
**Louis Rachid Salmi:** CHU de Bordeaux, Pole de sante publique, Service d'information medicale, F-33000 Bordeaux, France

The interpretation of the measurements of biomarker levels relies on the availability of specific data: toxicity, toxicokinetics, dose effect relationship, etc. There are two levels of interpretation: the first addresses the question of exposure while the second focuses more on the health risk. Concerning the former, if no clear toxicity threshold exists, the question (generally) asked is whether the exposure level is higher than a given percentile (often 95$^{th}$) of the level observed in the general population. To answer this, the concentrations of biomarkers measured are compared with reference values established in general population (measurements obtained in a reference population, typically with no known exposure or only minimal exposure to the toxicant of concern [3]). Concerning health risk, the question asked is whether the exposure level is higher than the International threshold level for safe health. Threshold values are based on specific scientific knowledge derived from the literature and from expert committees (e.g. HBM values in Germany or biomonitoring equivalents). Additional factors such as age, gender, body weight must be taken into account when evaluating exposure biomarker concentration results. Biomarkers are used for monitoring occupationally-exposed populations and in general population surveys. In the latter case, biomonitoring has different aims: i) to describe the exposure level of environmental pollutants in the general population; ii) to define national reference values; iii) to search for exposure determinants; iv) to analyze geographical and temporal variations through repeated surveys and v) to evaluate the impact of public policies on reducing exposure.

Several countries in Europe and North America have been developing biomonitoring programs for many years. The United States uses the National health and nutrition examination survey (NHANES) [4,5], a program designed in the early 1960s by the Centers for Disease Control and Prevention (CDC), to assess the health and nutritional status of adults and children and their exposure to various pollutants. During each survey period (2 years), different chemicals and/or their metabolites (i.e. biomarkers) are measured in approximately 2,500 people representative of the general population. In Germany, the Federal Environment Agency (UBA) implemented a cross-sectional nationwide population study in the mid-1980s, entitled the GerES (German Environmental Survey) to assess population exposure to environmental pollutants. The GerES is conducted in close collaboration with the National Health Interview and Examination Surveys (NHIES) of the Robert Koch Institute [6,7]. Sweden also has a long tradition of monitoring environmental pollutants, which in recent years has included biomonitoring (Karolinska Institute) of persistent organic pollutants in breast milk [8], blood lead levels in children since the 1970s, and emerging pollutants such as phthalates and perfluorinated compounds which are currently measured in sensitive groups (e.g. pregnant women, children ...).

Other European countries [9,10] and Canada have developed their own biomonitoring programs over recent years. The Canadian survey of health measures, CHMS (Canadian Health Measures Survey) [11,12] is a nationwide study conducted by Statistics Canada that collects information on the general health of 5,000 Canadian individuals, aged between 6 and 79 years old, through interviews and clinical measurements. The agency Health Canada has included a biomonitoring component in this survey, to provide national exposure data for a series of environmental chemicals. As part of the European Environmental Health 2004-2010 program, Europe started to harmonize biomonitoring practices in order to provide results that are comparable among EU countries.

In France, biomarkers were initially used as tools in both local [13,14] and multicenter polluted soil studies [15]. They were also used in many studies relating to occupational health risk assessment (including studies on polluted sites) [16].

The National Nutrition and Health Survey (ENNS) [17,18] was conducted in 2006-2007 by the French Institute for Public Health Surveillance (InVS). It provided the first reference values for the French population's exposure to pesticides and various metals. More recently, the Grenelle law (n ° 2009-967 of August 3, 2009) led to the development of a French National Biomonitoring program. This program was designed to estimate the population's exposure to various substances present in the environment (including in food) and to improve the understanding of the determinants of exposure. It consists of two distinct studies:

- The analysis of biomarkers in children selected through the Elfe cohort (Longitudinal Study from Childhood) [19], which constitutes the perinatal component of the program,

- The analysis of biomarkers in the French metropolitan population aged between 6 and 74 years, in a cross-sectional survey (Environmental Health Biomonitoring Physical Activity and Nutrition Survey, called "Esteban").

Generally, the selection of pertinent exposure biomarkers for a specific pollutant is based on the analysis of several criteria. Some of these are specific to the intrinsic characteristics of the biomarker: sensitivity, specificity, expected levels in the general population, biological half-life (a biomarker with a short half-life is used to study recent exposure, for example, urinary

arsenic reflects exposure in the previous 3 to 4 days), correlation between chemical concentrations measured in environmental media and in biological matrix, intra-and inter-individual variability. Other criteria are related to feasibility considerations (the invasiveness of the samples as blood samples, the minimum required blood/urine sample volume, the transport of the samples, the cost, etc.) and to analytical procedures. For many substances in the environment there are currently no relevant or validated exposure biomarkers [1]. Consequently, selecting an appropriate exposure biomarker list is often an arduous task.

In this context, elaborating a method to choose and prioritize biomarkers to use in the Elfe and Esteban studies was essential. As far as possible, we try to find a rigorous method that would not only formalize the procedures for considering knowledge and opinions and making decisions about which biomarkers to use, but would also maintain the flexibility required for negotiation and cooperation among international actors.

We present the results of this work, whose aim was to develop a prioritization method based on consensual selection criteria that would be applied in a formalized approach in order to reach a final list of biomarkers to be used in the Elfe and Esteban studies. The protocol for what is entitled "Biomarker Choice and Prioritization" is the cornerstone of the French National Biomonitoring program, as it will lead to the definition of the list of pollutants and associated biomarkers that will be monitored in the French population for the next ten years.

# 2 Methods

The entire process that we applied, detailed in the following text, is described in Figure 1.

## 2.1 Step A. Creating the list of priority biomarkers

We proceeded in two steps:

- The aim of the first step was to build the most exhaustive list of useful exposure biomarkers. The three teams detailed below worked on this, and ultimately arrived at a list of groups of pollutants of interest with corresponding 50 groups of biomarkers.

In 2009-2010, a working group (Team # 1) was created. It included members of the Department of Health and Environment in the French Institute for Public Health

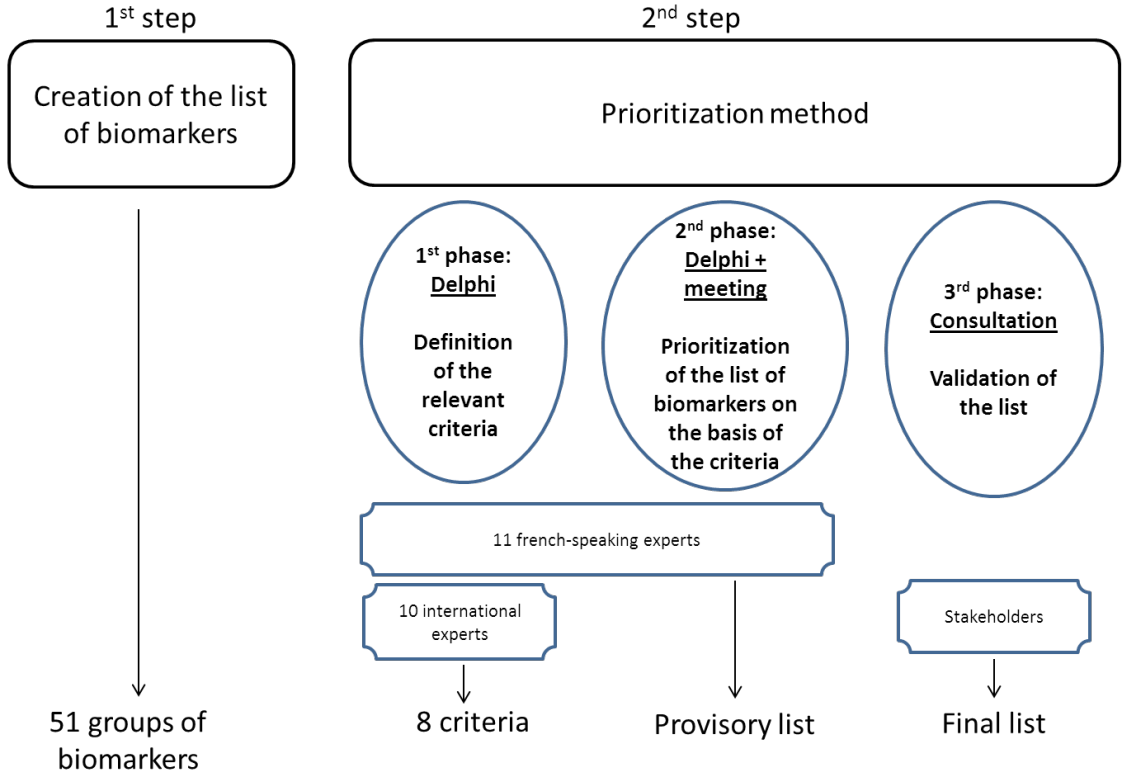


**Figure 1:** Process applied for prioritization of biomarkers in the French Biomonitoring program

Surveillance (InVS) who had recognized experience in biomonitoring studies, various ministries (health, ecology and work), and other public health agencies. The objective of this group was to validate a first set of pollutants (more than 100) on the basis of biomonitoring feasibility (from international and French experience), relevance (according to key information on the toxicity of substances, for instance, carcinogenicity and endocrine disruption effects[1]) and existing regulations for the compounds in air or in water. This list was then extended to pollutants that members of the working group considered of major interest. This selection was based on toxicity and priorities in terms of health effects or routes of exposure.

This larger list of pollutants was classified according to associated, known sets of biomarkers (based on chemical properties, exposure relevance, toxicity and analytical techniques), leading to a list of 50 biomarker groups established with the help of the Scientific Council of the French national program of biomonitoring. This council included 13 members with epidemiology, analytical and clinical toxicology, ethical, and public health knowledge.

## 2.2 Step B. Prioritization method

The second step of the process consisted in prioritizing the classified list of 50 groups of biomarkers for each of pollutant groups in order to answer some scientific questions: knowledge (about chemical properties, toxicity, etc..), respond to logistical constraints and feasibility (e.g. Are the collected matrix blood/urine sample volumes sufficiently large?) and meet budgetary constraints (e.g. minimum number of subjects, pooling analyses).

The Delphi consensus method was used for prioritization and was developed in three phases: i) the definition of relevant criteria for selecting biomarkers, as defined by the Delphi consensus [21-26]; ii) the prioritization of the biomarker list on the basis of these criteria and iii) the validation of the list by the stakeholders.

### 2.2.1 First phase: criteria definition

In this step, a group of 11 French-speaking experts (team # 2) from different fields was selected to define the criteria for biomarker selection. They were selected as a function of distinct parameters: i) their scientific knowledge of the field of biomarker development ; ii) their affiliation to different institutions related to the field; iii) their affiliation to NGOs engaged in the fields of public health and/or the environmental and especially concerned by the biomarker use. This list of 11 experts was proposed to the French Scientific Council for ratification. It included 3 toxicologists, 1 expert in occupational medicine, 3 epidemiologists, 1 expert in pollutant exposure, 1 expert from the chemical industry and 2 environmental NGOs were selected. Using the Delphi method [21-26], a list of selection criteria for the Elfe and Esteban studies, which the scientific council of the French National Program of Biomonitoring judged to be relevant, was established. Via an e-mail questionnaire, all the experts of team # 2 were then asked to give their opinion regarding the relevance of each of eight criteria which would be used to classify the biomarkers: from 0 if the criterion was, in their opinion, not relevant to 10 if it was very relevant. The eight criteria were defined as:

- The hazardous properties to health of the substances exposure to be measured. This criterion included the known or potential toxic effects of substances and their severity, especially carcinogenicity, reproductive toxicity, mutagenicity, neurotoxicity, immunotoxicity, and endocrine-disrupting effects;
- The exposure characteristics. This criterion included: i) the nature of contamination sources (anthropogenic and natural); ii) the characteristics of contamination (dispersive or confined); iii) the potential human exposure and the characteristics of the exposed population (general population, workers only or vulnerable populations: children, pregnant women, etc.), and iv) the possibility of multi-method/multiple sources of/multiple types of exposure (soil, air, water, etc.);
- The social perception. This criterion reflected the level of public concern (Were exposure to the particular pollutant and its potential effects a concern for the public authorities? Were the dangers of this substance given media coverage?);
- The biomarker characteristics. This criterion included the meaning of the marker (i.e. does it reflect current exposure and/or the internal dose accumulated, and/or the biologically active internal dose?) and also took into account the sensitivity, specificity, and the intra-individual and/or inter-individual variability of the biomarker.
- The results' interpretation. This criterion included the availability of information for interpreting the results of biomarker exposure measurement, such

1 key information on the toxicity of endocrine disruptor compounds were based on the Inserm Collective expertise [20], even though the potential toxic effects were very difficult to estimate because of the lack of regulatory or recognized methods to assess specific pollutant mechanisms.

as: the distribution of biomarker levels in a reference population; knowledge of the relationship between the biomarker level and external exposure and/or adverse effects; the toxicokinetics of the xenobiotic, and of the biomarker when not the xenobiotic itself (ideally integrated in a physiologically-based pharmacokinetic (PBPK) model); the individual and environmental factors that may influence the fate of the xenobiotic, in vivo analysis (co-exposures, food habits, genetic determinants, body mass index, etc.);
- The logistic and analytical feasibility. This criterion included the sampling method's human invasiveness, the blood or urine sample volume required to analyze biomarkers, the conditions for collection (transport, storage, etc.), the availability of a validated assay method with sufficient information to analyze biomarkers, such as the existence of a detection limit and a quantification limit adapted for the interpretation, and the cost of analysis;
- The feasibility of prevention. This criterion included the availability of European or national regulations, the availability of a toxicity reference value (TRV), as well as the current feasibility of exposure reduction, taking into account its techno-economic and social implications, the possibility of supporting a pre-defined public health policy, etc;
- The contribution in terms of new knowledge in France, considering the gaps of knowledge at the national and international levels and the national specificities in terms of exposure, behavior, susceptibility to exposure, etc., the need for national data for harmonization and international comparisons.

Each expert had to justify the reasons for his/her choices and was invited to suggest additional criteria and their rated relevance (again, from 0 to 10), if needed. Moreover, these questionnaires were also sent to 10 international experts (team # 3) - identified as leaders on biomonitoring in their home countries - in order to compare the results of the 11 French-speaking experts with international ones. This allowed us to produce a working list for the rating of each of the criteria based on each expert's choices. According to the Delphi method, this list was then sent back to all 21 experts to compare their own ratings for criteria with those of all their counterparts. In a second questionnaire, they were asked whether or not they wished to modify or not their selection after considering the answers of the other experts, and to elaborate the arguments supporting their conclusions (i.e. agreement or disagreement with the other experts' choices). All 3 rounds (the second questionnaire was sent twice) were conducted anonymously. The anonymity of the synthesis ensured the independence of each expert in the overall group (teams 2 and 3) throughout the process. No expert was aware of the identity of the other group members until the end of the first phase. This methodological decision allowed us to produce a consensus on the criteria that should be applied to each biomarker in order to finish/complete the prioritization process.

#### 2.2.2 Second phase: the biomarker prioritization process

The finalized list which described the overall rating (obtained after combining each expert's final scores for each selection criterion) was submitted to all 21 experts and their collective approval obtained. They were then asked to prioritize the different biomarkers using this list. To that purpose we asked each expert to rate the 51 groups of biomarkers (identified in Step A above) according to each of the 8 criteria (see above). The importance for each criterion was defined as follows: 0.8 if the whole group of biomarkers fitted the criterion; 0.6 if the answer was somewhat true; 0.4 if the answer was somewhat untrue; and 0.2 if none of the biomarkers of the group fitted the criterion. The sum of each criterion's scores yielded a total rating for each of the 50 biomarker groups. These global scores were then used to rank the various biomarker groups and thus produce a prioritized list. The next step was the organization of a meeting, gathering all the French-speaking experts. At this meeting, on May 3rd, 2011, the ranking of each group of biomarkers on the list was discussed. During this discussion, the experts were asked to state the adjustments they wished to make. *In fine*, the new prioritized list of biomarkers was submitted to them by e-mail in May 2011 for final approval.

### 2.3 Finalization of the list

In accordance with the commitments made at the Grenelle Environment Forum, the Government adopted a second National Environmental Health Plan (PNSE) on June 24th, 2009, for the subsequent period 2009-2013.

This second PNSE consisted of three working groups involving elected officials, local authorities, and the concerned State agencies, but also representatives of associations, trade unions and companies, experts, qualified individuals, and professionals in the health system and in health insurance. The actions of this second PNSE were developed along two main structural axes: reducing exposure to diseases with a high impact on

health (e.g. cancer, cardiovascular diseases, respiratory diseases and neurological diseases, etc) and reducing environmental inequalities. One of the three working groups, the "emerging risks" group, entailed ensuring continuous dialogue on emerging risks (including nanotechnologies, electromagnetic waves, and endocrine disruptors). We submitted our prioritized list to this "emerging risks" group in May 2011 for their comments, as the PSNE was responsible for funding the Elfe and Esteban studies. The list was slightly modified according to the feedback received. The final list was finally presented by email to the members of the National Biomonitoring Program's Scientific council, and then to its Steering Committee on July 1st, 2011.

# 3 Results

## 3.1 Criteria definition

The final scores for each criterion are presented in Figure 2 for all French-speaking experts. As the results were similar between French-speaking experts and foreign experts, the scores of the latter are not presented. Hazard identification was the highest-rated criterion and social perception the lowest-rated. Although the final scores for each criterion were not very different from one another, classification was still possible. The second least-rated criterion was "the feasibility of prevention." The following extra criteria were proposed by the experts but were considered as already

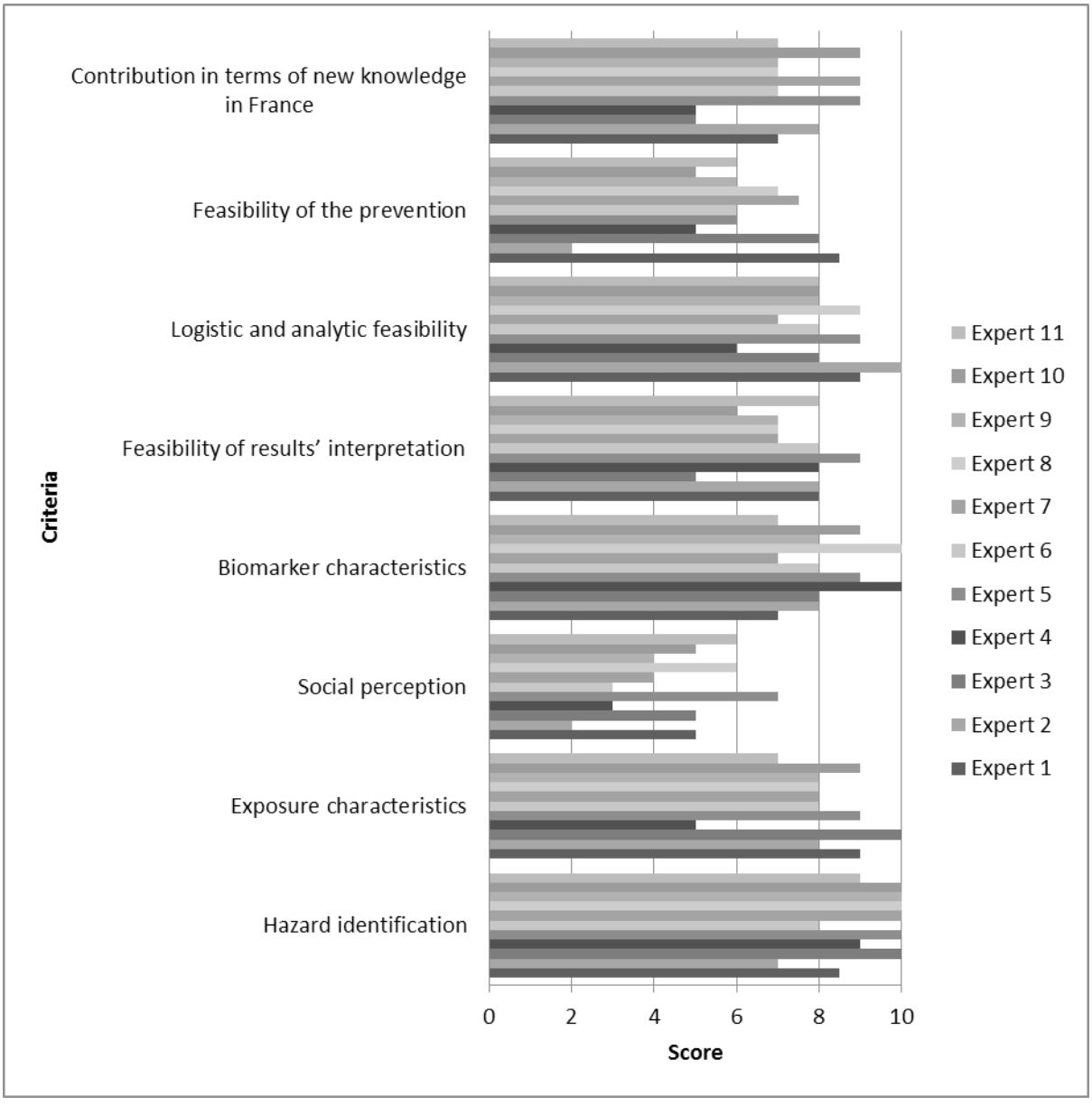


**Figure 2:** Results of the questionnaire for French speaking experts

included into the previously eight identified criteria: worldwide concern in social perception; bioaccumulation or half-life, bioavailability, bio-persistence of the pollutant, knowledge of the circadian rhythm, and lack of alternative biomarker to estimate exposures in biomarker characteristics; existence of human toxicity data and health hazard in hazard identification; heterogeneity of pollutant exposures and target population based on a critical age window in exposure characteristics; new contaminants contributing to new knowledge in France.

## 3.2 Biomarker prioritization

After calculating the scores for each of the 50 groups of biomarkers and discussing the list position (ranking) of each biomarker group at the meeting, the French-speaking experts produced a list of prioritized biomarkers (see Methods and Table1). The group of biomarkers classed in first position (i.e. highest ranking importance) was serum "dioxin-like polychlorinated biphenyls, dioxins and furans". As shown in Figure 3, this was mostly, but not only, due to the high rating obtained in the social perception criterion (for example, urinary "bisphenol A" also had a strong social perception and was in 11$^{th}$ position). This result shows that even though social perception had the lowest overall ranking, it still modified the ranking of biomarkers exhibiting similar scores. Biomarker characteristics and the feasibility of interpreting the results were two high-rated criteria. Many biomarker groups obtained the maximum score for the criterion "Hazard identification": for example, "lead" in blood was classed in second position in the final list. Urinary "cadmium", urinary "benzene", urinary "mercury" and the blood VOCs "tetrachloroethylene and trichloroethylene" were classed in the third, fourth, sixth and seventh positions, respectively. The group of blood trihalogenomethanes "chloroform, bromoform, bromodichloromethane and dibromochloromethane" obtained the highest ranking for the criterion "exposure characteristics" (14$^{th}$ position).

Concerning the criterion "logistic and analytic feasibility", the highest biomarker rating was obtained for urinary "cadmium". Urinary "cotinine," which is globally listed in fifth position, had the highest rating for the criterion "feasibility of prevention". Serum "PFOA" was in 15$^{th}$ position in the final list but had the highest rate for the criterion "new knowledge in France".

However, the positioning of some groups of biomarkers in the list was challenged in discussions during the expert meeting. The meeting ended in the

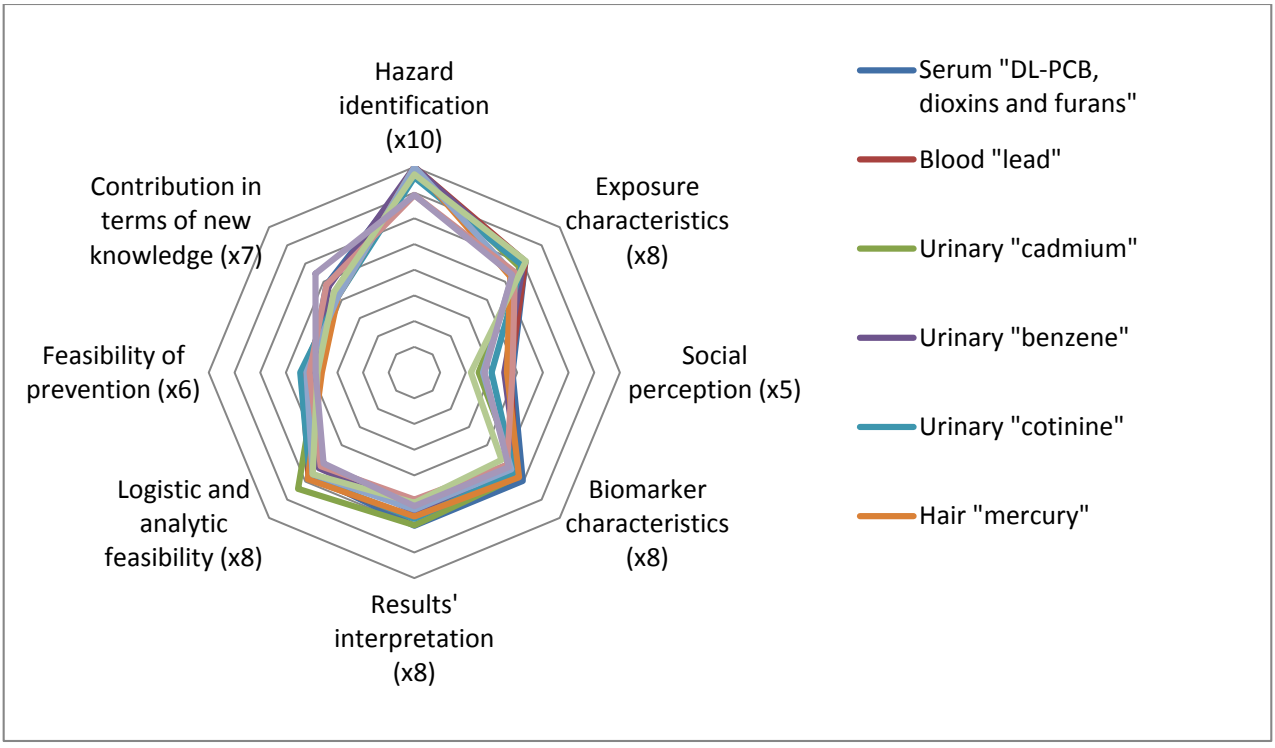


**Figure 3:** Illustration of the results of the biomarker prioritization

production of a consensual, prioritized list of biomarkers to be included in the French national biomonitoring program (Table 1).

- The following biomarkers were downgraded in the priority list because they were undetectable in previous French studies, because there was no argument for significant exposure in the French general population and/or because of previous difficulty to measure them.
  - "tin in urine" ;
  - Cesium ;
  - 3-Hydroxybenzo[a]pyrene ;
  - Tetrachloroethylene and trichloroethylene.
- The biomarkers cited below were finally ranked higher in the prioritized list because a previous study conducted of the French general population (ENNS[17,18]) had shown high contamination and had enabled measures to be taken to reduce domestic exposure or because they had been grouped with other biomarkers for analytical reasons:
  - Deca BDE 209 ;
  - PFOA ,
  - 1,4-Dichlorobenzene ,
  - Chlorophenols, and organophosphate metabolites.

Unfortunately, the international experts did not participate in the round table meeting and so they could not discuss their results with those of the French-speaking experts.

### 3.3 Stakeholders

The prioritized biomarker list was presented to the PNSE2 group for comments. This group, (described in 1.b iii above) made the following recommendations:

- All 51 biomarker groups in the list deserve to be analyzed and if this is not possible, as many pesticide exposure biomarkers as possible should be used.
- Biomarkers should be prioritized inside the 50 different biomarker groups.

Based on the biomarker group scores and the stakeholders' advice, each of the 50 groups of biomarkers was finally ranked in order of priority. The new list was constructed taking into account the ranking of the first group of biomarkers by family. Within families, "priority A" (first half of the list) and "priority B" (second half of the list) were identified.

This was the list finally presented to the members of the scientific council and to the Steering Committee of the National Biomonitoring Program, as described above.

This final list is presented in supplementary data.

## 4 Discussion

This prioritization process was long and sometimes arduous: in the beginning, there were occasionally discrepancies in the understanding of some criteria. The common definition of the various criteria was improved upon at the expert meeting, thanks to common understanding and agreement used for the final biomarker ranking. The selection of the experts was a fundamental step of our criteria-definition process. The limits of the Delphi method have already been addressed in various papers [27]. Obviously, the experts were representative of their individual scientific/political/societal domain and recognized as such. Since the experts came from different disciplines and backgrounds, they did not have expertise on all proposed biomarker groups and some found this situation difficult. However, this difficulty was overcome by asking the experts not to rate groups of which they had no information or knowledge. As a consequence, the score obtained for each group of biomarkers was the average of the ratings given by the experts who had ranked that particular group. After considering many responses, we finally chose not to retain the criterion "cost" to prioritize the list of biomarkers. We preferred to prioritize based only on scientific criteria, as financial conditions are always taken into account in general when conducting a study. The experts also raised other questions/raised other points concerning the initial establishment of the list of the 50 biomarker groups (this was especially true for pesticides) and the fact that the same list of biomarkers was ranked for two studies with two different populations: adults (Esteban) and pregnant women and newborns (Elfe). The medians of all the criteria were similar except for the extreme criterion "hazard identification" [highest median] and for the criterion "social perception" [lowest median]. According to some experts, the criteria did not seem discriminatory. However, all criteria were chosen for their relevance. Finally, the correlation between the 11 experts on the rating criteria was fairly good with a Kendall coefficient of concordance of 0.59 ($p<0.0001$). In the end, each group of biomarkers was rated for each criterion: for example, social perception could be a low-rated criterion, but it could be more relevant to one group of biomarkers than to another.

Finally, the final prioritized list obtained contained both historic pollutants (i.e. already measured in previous studies, such as dioxins or lead) and emerging pollutants (phthalates, bisphenol A). This list certainly reflects the French population's preoccupations in terms of exposure to ubiquitous environmental pollutants. When reviewing the prioritized list obtained, the experts did not contest

either the biomarkers listed at the top of the prioritized list (biomarkers for which analysis was considered a priority) or at the bottom of the list (biomarkers that are not considered a priority to measure).

Our selection and prioritization procedure is different from the methods used in other studies:

In GerES [6,7] (see above), a scientific committee conducts the selection of biomarkers according to criteria such as level of toxicity, health risk, relevance to environmental policy, exposure of the population, existence of reliable sampling procedures, existence of analytical capacities, relevance of public policy, and costs. Unlike our process, the criteria are not defined by a method of consensus.

In Nhanes [4,5], the biomarker selection process is also long and involves several steps, but the method used is participatory:

- Production of a Federal Register that can provide each substance being followed; the removal of pollutants from the registry involves a complex procedure and should meet multiple conditions, e.g. the stability of the biomarker in the population or its presence in less than 5% of the population.

Classification of pollutants into categories by a scientific committee according to criteria such as risk of exposure, health effects, analytical capacities.

In Canada, ECMS [11] selects 137 biomarkers, integrating even so-called "emerging" biomarkers (Phthalates and Bisphenol A in particular), but the criteria for biomarker selection were not established using a consensus method. Informed by an advisory committee of experts, doctors, and laboratories, Statistics Canada considers both the expected follow-up and the cost of measurement when optimizing the cost of laboratory testing. Finally, the size of the investigated population varies from one group of biomarkers to another, and some biomarkers are measured every year, while others are evaluated only every 4 years.

A similar approach is used in the so-called "participatory" procedure in the Nhanes studies, but the present study is unique in that it utilizes a formalized consensus method to select and rate criteria. This is very useful to understand or describe the complex biomarker prioritization decision making process. Although methods vary from one country to another, the final lists of priority biomarkers are very similar. Beyond the nature of the pollutants and of the criteria, the study presented here has the advantage of bringing greater transparency about the implementation, practicality and efficacy of using a selection process that is not very well known in this domain.

# 5 Conclusion

The results we report here were generated by a panel of French-speaking experts. The formalized approach used for the prioritization of biomarkers is useful in terms of traceability of the final selection of biomarkers included in Human Biomonitoring programs. Given the political stakes, this method combines rigor and flexibility and clearly, in our case, helped to build a prioritized list that will be shared and supported by many if not all actors.

# References

[1] Toniolo P, Boffeta P, Shuker D, Rothman, Hulka, Pearce *et al*.. Application of Biomarkers in Cancer Epidemiology - Workshop Report. Lyon (France): IARC Scientific Publications, N°. 142; 1997: p.1.

[2] Paustenbach D, Galbraith D: Biomonitoring and biomarkers: exposure assessment will never be the same. *Environ Health Perspect* 2006, 114: 1143-1149.

[3] Pirkle JL, Needham LL, Sexton K: Improving exposure assessment by monitoring human tissues for toxic chemicals. *J Expo Anal Environ Epidemiol* 1995, 5: 405-424.

[4] Calafat AM: The U.S. National Health and Nutrition Examination Survey and human exposure to environmental chemicals. *Int J Hyg Environ Health* 2012, 215: 99-101.

[5] Needham LL, Calafat AM, Barr DB: Uses and issues of biomonitoring. *Int J Hyg Environ Health* 2007, 210: 229-238.

[6] Becker K, Conrad A, Kirsch N, Kolossa-Gehring M, Schulz C, Seiwert M *et al*.: German Environmental Survey (GerES): human biomonitoring as a tool to identify exposure pathways. *Int J Hyg Environ Health* 2007, 210: 267-269.

[7] Seifert B, Becker K, Hoffmann K, Krause C, Schulz C: The German Environmental Survey 1990/1992 (GerES II): a representative population study. *J Expo Anal Environ Epidemiol* 2000, 10: 103-114.

[8] Verner MA, Sonneborn D, Lancz K, Muckle G, Ayotte P, Dewailly E *et al*.: Toxicokinetic modeling of persistent organic pollutant levels in blood from birth to 45 months of age in longitudinal birth cohort studies. *Environ Health Perspect* 2013, 121: 131-137.

[9] Cerna M, Krskova A, Cejchanova M, Spevackova V: Human biomonitoring in the Czech Republic: an overview. *Int J Hyg Environ Health* 2012, 215: 109-119.

[10] Schoeters G, Den HE, Colles A, Loots I, Morrens B, Keune H *et al*.: Concept of the Flemish human biomonitoring programme. *Int J Hyg Environ Health* 2012, 215: 102-108.

[11] Haines DA, Murray J: Human biomonitoring of environmental chemicals--early results of the 2007-2009 Canadian Health Measures Survey for males and females. *Int J Hyg Environ Health* 2012, 215: 133-137.

[12] Haines DA, Arbuckle TE, Lye E, Legrand M, Fisher M, Langlois R *et al*.: Reporting results of human biomonitoring of environmental chemicals to study participants: a comparison of approaches followed in two Canadian studies. *J Epidemiol Community Health* 2011, 65: 191-198.

[13] Fillol C, Dor F, Clozel B, Goria S, Seta N: Does arsenic in soil contribute to arsenic urinary concentrations in a French population living in a naturally arsenic contaminated area? *Sci Total Environ* 2010, 408: 6011-6016.
[14] Frery N, Maury-Brachet R, Maillot E, Deheeger M, de MB, Boudou A: Gold-mining activities and mercury contamination of native amerindian communities in French Guiana: key role of fish in dietary uptake. *Environ Health Perspect* 2001, 109: 449-456.
[15] Frery N, Vandentorren S, Etchevers A, Fillol C: Highlights of recent studies and future plans for the French human biomonitoring (HBM) programme. *Int J Hyg Environ Health* 2012, 215: 127-132.
[16] Dor F, Haguenoer JM, Zmirou D, Empereur-Bissonnet P, Jongeneelen FJ, Nedellec V *et al.*: Urinary 1-hydroxypyrene as a biomarker of polycyclic aromatic hydrocarbons exposure of workers on a contaminated site: influence of exposure conditions. *J Occup Environ Med* 2000, 42: 391-397.
[17] Falq G, Zeghnoun A, Pascal M, Vernay M, Le SY, Garnier R *et al.*: Blood lead levels in the adult population living in France the French Nutrition and Health Survey (ENNS 2006-2007). *Environ Int* 2011, 37: 565-571.
[18] Saoudi A, Zeghnoun A, Bidondo ML, Garnier R, Cirimele V, Persoons R *et al.*: Urinary arsenic levels in the French adult population: the French National Nutrition and Health Study, 2006-2007. *Sci Total Environ* 2012, 433: 206-215.
[19] Vandentorren S, Bois C, Pirus C, Sarter H, Salines G, Leridon H: Rationales, design and recruitment for the Elfe longitudinal study. *BMC Pediatr* 2009, 9: 58.
[20] Inserm. Reproduction et environnement, Editions Inserm, juin 2011, 735 p, Collection Expertise collective.
[21] Brockhoff K. The performance of forecasting groups in computer dialogue and face-to-face discussion. In: Linstone HA, Turoff M, editors. The Delphi Method: Techniques and Applications. Reading, MA: Addison-Wesley Publishing Company; 2002. pp. 285-311.
[22] Campbell SM, Braspenning J, Hutchinson A, Marshall M: Research methods used in developing and applying quality indicators in primary care. *Qual Saf Health Care* 2002, 11: 358-364.
[23] Fink A, Kosecoff J, Chassin M, Brook RH: Consensus methods: characteristics and guidelines for use. *Am J Public Health* 1984, 74: 979-983.
[24] Hsu C-C, Sandford B. The Delphi technique: making sense of consensus. Pract Assess Res Eval. 2007;12(10):1-8.
[25] Jones J, Hunter D: Consensus methods for medical and health services research. *BMJ* 1995, 311: 376-380.
[26] Streiner DL, orman G. Health measurement scales. A practical guide to their development and use. Oxford: Oxford University Press, 2003.
[27] Baker J, Lovell K, Harris N. How experts are the experts? An exploration of the concept of 'exper' within Delphi panel techniques. Nurse Researcher; 2006; 14, 1 ;59.