To which mixtures are French pregnant women mainly exposed? A combination of the second French total diet study with the EDEN and ELFE cohort studies

Thiéma Traoré, Anne Forhan, Véronique Sirot, Manik Kadawathagedara, Barbara Heude, Marion Hulin, Blandine de Lauzon-Guillain, Jérémie Botton, Marie-Aline A Charles, Amélie Crépet

To which mixtures are French pregnant women mainly exposed? A combination of the second French Total Diet Study with the EDEN and ELFE cohort studies.

T. Traoré[1], A. Forhan[2, 3], V. Sirot[1], M. Kadawathagedara[2], B. Heude[2, 3], M. Hulin[1], B. de Lauzon-Guillain[2, 3], J. Botton[2, 4], M. A. Charles[2, 3], A. Crépet[1,a]

[1] ANSES, French Agency for Food, Environmental and Occupational Health & Safety, 14 rue Pierre et Marie Curie, 94701 Maisons-Alfort, France.

[2] INSERM, UMR1153 Epidemiology and Biostatistics Sorbonne Paris Cité Center (CRESS), Early Origin of Child Health and Development Team (ORCHAD), Paris, F-75014 France.

[3] Paris Descartes University, France.

[4] Faculty of Pharmacy, Paris-Sud University, Paris-Saclay University, 92296 Châtenay-Malabry, France.

Abstract

Pregnant women and their unborn child are exposed to a large number of substances during pregnancy. Some of these substances may cross the placenta, resulting in exposure of the foetus. There is growing evidence that certain substances could interact to produce a mixture effect. It is therefore essential to identify the main mixtures mothers are exposed to.

This study aimed to identify the major mixtures French pregnant women included in EDEN and ELFE cohorts were exposed to, on the basis of the 441 substances analysed in the second French total diet study. Exposure systems and the composition of substances were identified from co-

Abbreviations

BMI, body mass index; EDEN, study on the pre- and early postnatal determinants of child health and development; ELFE, French longitudinal study since childhood; FFQ, food frequency questionnaire; INCA, Individual and National Food Consumption Survey; LB, lower bound; SNMU, sparse non-negative matrix under-approximation; TDS, Total Diet Study

[a] Corresponding author: A. Crépet; email address: amelie.crepet@anses.fr

exposures using sparse non-negative matrix under-approximation to generate the main mixtures. Individuals were clustered to define clusters with similar co-exposure profiles.

Six clusters associated with eight mixtures were identified. For example in ELFE, cluster 2 comprising 10% of the population was characterised by mixtures “Pest-1” mainly contains pesticides and ”TE-F-PAH” contains trace elements, furans and polycyclic aromatic hydrocarbons. Five other clusters were also described with their associated mixtures. Similar results were observed for EDEN.

This study helps to prioritise mixtures for which it is crucial to investigate possible toxicological effects and to recommend epidemiological studies concerning health effects.

Keywords

Co-exposure assessment, sparse non-negative matrix under-approximation, chemical mixtures, pregnant women, individual clusters.

Highlights

- Exposures to 200 substances were assessed for French women before and during their pregnancy
- Co-exposure data was reduced into two matrices using SNMU method: exposure systems and individuals’ coefficients
- Eight main mixtures were defined from exposure systems and six of them were common
- Six clusters of mothers with similar pattern of contaminant intake were identified

1 Introduction

Populations worldwide are exposed to a wide range of chemicals through their diet, and pregnant women are no exception. The foetal and early postnatal periods are considered to involve increased susceptibility to long-term effects (Sly and Flack, 2008). Numerous studies have shown associations between prenatal exposure to environmental contaminants that cross the placental barrier, in particular endocrine disruptor chemicals, and postnatal growth (Mendez et al., 2011; Valvi

et al., 2012) or between acrylamide exposure and prenatal growth (Pedersen et al., 2012; Duarte-Salles et al., 2013). Heavy metals have also been found to have an impact, in particular on cognitive, motor and intellectual development (Canfield et al., 2003; Dorea and Donangelo, 2006).

In most cases, chemicals are studied individually and the potential interactions or additive effects of substances are not taken into account. However, some recent epidemiological studies have highlighted relationships between exposure to mixtures of substances and health effects. For example, Czarnota et al. (2015) showed an association between exposure to a mixture of five polychlorinated biphenyls (PCBs), seven polycyclic aromatic hydrocarbons (PAHs) and 15 pesticides and the risk of non-Hodgkin lymphoma. On the basis of a literature review, Claus Henn et al. (2014) noted several interactions between substances and health effects in children. For example, an interaction between cadmium and lead on reproductive hormones and neurodevelopment was highlighted. The authors also emphasised the need to improve statistical tools to study exposure to mixtures. It is still a challenge to determine which substances should be studied together. For example, EFSA (2013) proposed cumulative assessment groups (CAGs) based on grouping pesticides by their organ toxicity or specific effects. However, these CAGs do not take into account the probability of co-exposure of an individual to substances in a particular CAG. Over the last seven years, methods have been proposed to identify major mixtures from environmental exposures. Given the high number of substances to which populations are exposed, it is necessary to develop adapted statistical methods to reduce the dimension, to consider and to identify mixtures. Crépet and Tressou (2011) used a Bayesian non-parametric model to determine major pesticide mixtures. More recently, an approach based on non-negative matrix factorisation (NMF) (Lee and Seung, 2001) and clustering methods was used to identify the primary mixtures associated with specific diets (Béchaux et al., 2013; Traoré et al., 2016). The NMF and clustering methods were used to define dietary patterns and clusters of individuals with similar diets (Zetlaoui et al., 2011; Sy et al., 2013; Gazan et al., 2016).

The objective of this study was to describe the main mixtures pregnant women in France are exposed to, using a modified version of the NMF method, called sparse non-negative matrix under-approximation (SNMU) (Gillis and Plemmons, 2013). Exposures to up to 200 substances were estimated from two datasets: firstly, the dietary consumption of pregnant women in France before and during pregnancy on the basis of the EDEN and ELFE mother-child cohorts; secondly, substance concentration levels in foods provided by the second French Total Diet Study (TDS 2). The SNMU was completed by hierarchical clustering to classify groups with similar exposure profiles. Mixtures identified from the various exposure data are described in this paper for both cohorts. The exposure profiles are described only for ELFE cohort.

2 Materials and methods

2.1 Food Consumption

Pregnant women before 24 weeks of amenorrhea ($N = 2{,}002$) aged from 18 to 45 years were included in the EDEN mother-child cohort. Recruitment took place between February 2003 and January 2006 in two university hospitals in France in the cities of Nancy and Poitiers (Heude et al., 2016). Participants completed a validated (Deschamps et al., 2009) food frequency questionnaire (FFQ) at two time points. The first FFQ was recorded at inclusion and corresponds to maternal diet in the year before pregnancy. The second FFQ was completed at delivery and concerned maternal diet during the last trimester of pregnancy. The consumption frequency of 137 food and beverage items was collected with a 7-item scale from “never” to “more than once a day”. Portion sizes were determined using pictures for 12 food types (meat, chips, pasta, vegetables, cakes, cheese) on a three-level scale and for other food types, sizes were standard portions for the adult population in France (Hercberg et al., 2002). Individual daily intake of each food item was assessed by combining the frequency of intake and portion size. Subjects for whom more than three food items of the FFQ were missing were excluded. For the subjects with only one or two missing items, the median values

of daily intakes were attributed. In all, 1,806 FFQs for the diet before pregnancy and 1,666 for the diet during the last trimester of pregnancy were included in this analysis.

The ELFE cohort study is another mother-child investigation on a larger and highly representative sample of pregnant women in mainland France in 2011 (Vandentorren et al., 2009). Women attending 320 randomly selected maternity hospitals were asked to participate. Recruitment was carried out in four waves of 4 to 8 days. An FFQ modified from the EDEN FFQ was used to collect data on dietary habits for 18,042 women during the last trimester of pregnancy. The consumption frequency of 125 food and beverage items was collected using a 7-item scale from "never" to "more than once a day". The portion sizes of 75 food items were estimated using photos for different food types on a 5-level scale based on the SU.VI.MAX portion book (Hercberg, 2000). Individual daily intake of each food item was assessed by combining frequency of intake and portion size. Individuals with more than 10 missing food items in their FFQ were excluded. For the subjects with fewer than ten missing items, the median values of daily intakes were attributed. Finally, intake data were available for 15,226 women.

2.2 Concentration data

The French total diet study (TDS 2) (Sirot et al., 2009) provided concentration levels for 441 substances in 212 core foods (Arnich et al., 2012; Bemrah et al., 2012; Nougadère et al., 2012; Sirot et al., 2012a; Sirot et al., 2012b; Sirot et al., 2013; Veyrand et al., 2013; Bemrah et al., 2014; Rivière et al., 2014). The 212 core foods cover about 90% of the whole diet of the French population, as estimated from the second French national food consumption survey (INCA 2) (Dubuisson et al., 2010; Lioret et al., 2010). In order to be representative of consumer habits in France, each type of food considered was composed of 15 samples of the same food, so as to cover different varieties, purchase locations, preparation methods, cooking methods, etc. In all, 19,785 food products were purchased in eight regions in France over different seasons from 2007 to 2009 to make up the 1,319 composite samples of core foods to be analysed for additives, environmental contaminants, pesticide residues, trace elements and minerals, mycotoxins, phytoestrogens and acrylamide.

As recommended in GEMS/Food-EURO (2013), non-detected values were replaced by 0 and detected values unable to be quantified, by the limit of detection. With this scenario, called the lower bound (LB) scenario, the exposure of the whole population is equal to zero for 210 pesticides, four perfluoroalkyl acids and seven mycotoxins that were not considered in this study. In addition, 10 minerals (Ca, Cu, Fe, K, Mg, Mn, Mo, Na, Se and Zn) out of 13 analysed in TDS 2 were not considered in the present study, in the absence of potential adverse effects on development or health. The remaining 210 out of 441 substances were considered, including:

- 21 trace elements and minerals: aluminium (Al), antimony (Sb), barium (Ba), cadmium (Cd), cobalt (Co), gallium (Ga), germanium (Ge), lithium (Li), lead (Pb), nickel (Ni), silver (Ag), strontium (Sr), tellurium (Te), tin (Sn) and vanadium (V). Separate analyses were performed to take into account the proportion of inorganic and organic arsenic (As) and mercury (Hg), and the proportion of trivalent and hexavalent chromium (Cr). Inorganic and organic arsenic (Asi and Aso) and mercury (HgI and MeHg) and trivalent and hexavalent chromium (CrIII and CrVI) were therefore considered instead of total arsenic, chromium and mercury;
- 17 congeners of polychlorinated dibenzo-p-dioxins (or dioxins) and polychlorinated dibenzofurans (or furans) (PCCD/F) (HCDD/F-123478, HCDD/F-123678, HCDD/F-123789, HCDD/F-1234678, OCDD/F, PCDD/F-12378, TCDD/F-2378, HCDF-1234789, HCDF-234678, PCDF-23478);
- 12 congeners of 'dioxin-like' polychlorinated biphenyls (PCBs) (PCB-77, 81, 105, 114, 118, 123, 126, 156, 157, 167, 169, 189);
- six congeners of 'non dioxin-like' polychlorinated biphenyls (PCB-28, 52, 101, 138, 153, 180);
- 12 perfluoroalkyl acids (PFAAs): nine carboxylates (PFDA, PFDoA, PFHpA, PFHxA, PFNA, PFOA, PFTeDA, PFTrDA, PFUnA) and three sulfonates (PFBS, PFHxS, PFOS);
- 14 brominated flame retardants (BFRs): eigth polybrominated diphenyl ether congeners (PBDE-28, 47, 99, 100, 153, 154, 183, 209), three polybrominated biphenyl congeners (PBB-52, 11, 153) and three hexabromocyclododecane congeners (HBCD-alpha, beta, gamma);
- 18 mycotoxins: fumonisins B1 and B2 (FB1, FB2), ochratoxin A, B (OTA, OTB) and patulin (Pat), trichothecenes from group A including T2-toxin, HT2-toxin, diacetoxyscirpenol (DAS) and monoacetoxyscirpenol (MAS); group B including nivalenol (NIV), deoxynivalenol (DON), de-epoxy derivative of DON (DOM-1), 3-acetyldeoxynivalenol (DON3), 15-acetyldeoxynivalenol (DON15), and fusarenon X (FusX), zearalenone (Zea) and its metabolites: alpha-zearalanol and alpha-zearalenol;

- 11 phytoestrogens: biochanin A, coumestrol, daidzein, enterolactone, equol, , formononetin, genistein, glycitein, matairesinol, resveratrol, and secoisolariciresinol;
- 73 active pesticide residues;
- 4 additives: nitrites, sulphites, tartaric acid and annatto;
- 21 heat-induced contaminants including acrylamide and 20 congeners of polycyclic aromatic hydrocarbons (PAHs): anthrancene (AN), benzo[a]anthracene (BaA), benzo[a]pyrene (BaP), benzo[b]fluoranthene (BbF), benzo[c]fluorine (BcFL), benzo[g,h,i]perylene (BghiP), benzo[j]fluoranthene (BjF), benzo[k]fluoranthene (BkF), chrysene (CHR), cyclopenta(c,d)pyrene (CPP), dibenzo[a,h]anthracene (DBahA), dibenzo[a,e]pyrene (DbaeP), dibenzo[a,h]pyrene (DbahP), dibenzo[a,i]pyrene (DbaiP), dibenzo[a,l]pyrene (DbalP), fluoranthene (FA), indeno[1,2,3-cd]pyrene (IP), 5-methylchrysene (MCH), phenanthrene (PHE) and pyrene (PY);
- and bisphenol A (BPA).

2.3 Matching the TDS 2 with FFQ data from the EDEN and ELFE cohort studies

Since some differences were observed between food items from TDS 2 and the two cohorts, matching between the nomenclatures of these studies was conducted. In most cases (65% for EDEN and 55% for ELFE), a food item from the two cohorts was comparable to a food item from TDS 2. In other cases, different scenarios were used:

- Correlation scenario: the food items from the two cohorts were more detailed than in TDS 2. For example, in the two cohorts, carrots are separated into “cooked carrot” and “raw carrot”, while in TDS 2 the food item “carrots” covers cooked and raw carrots. In this case, the concentration of the TDS 2 food item “carrots” was used for both detailed food items in the cohorts.
- Grouping scenario: a cohort food item is covered by several food items from TDS 2. In this case, the weighted mean of the concentrations of related TDS 2 food items was attributed to these cohorts’ food items. The weights were estimated by the ratios of consumed quantity of each food item by the female population aged 18 to 45 years of the INCA 2 study. This scenario was applied for example to the category “dried vegetables” in the two cohorts and equivalent to “white bean” + “lentil” in TDS 2.
- Recipes scenario: this case concerns only the category considered as composite foods in the two cohorts. For these products, the concentrations were calculated by using recipes. For example, it is applied to the “gratin dauphinois” equivalent to “potatoes” + “semi-skimmed milk” + “eggs” + “butter” + “cream” in TDS 2.

After the matching process, the five food items in EDEN (avocado, pumpkin, prepared light meals, sweetener and whisky) and the six in ELFE (avocado, pumpkin, mushroom, other starches, prepared light meals and exotic fruits) with no TDS 2 correspondence were excluded. Finally, 207 substances were associated with 132 food items in EDEN, and 210 substances with 119 food items in ELFE.

2.4 Exposure assessment

Exposure e_{np} of woman n to contaminant p was assessed (see equation below) by combining the concentration c_{fp} in food f of contaminant p with the average daily intake q_{nf} of woman n to food f and summing over all the food items reported as consumed by each woman.

$$e_{np} = \sum_{f=1}^{F} q_{nf} \times c_{fp}$$

The exposure of the population of N individuals to P substances can be expressed as a matrix E of dimension $(P \times N)$.

Three different exposure matrices were obtained: two in the EDEN cohort (i) before pregnancy (EDEN before pregnancy) and (ii) at the end of pregnancy (EDEN during pregnancy), and (iii) one in the ELFE cohort at the end of pregnancy (ELFE during pregnancy). To deal with differences of variability between the exposures, each exposure matrix was standardised, i.e. E was multiplied by the $(P \times P)$ inverse diagonal matrix of the standard deviations of the exposure to each substance across individuals S_E: $E^* = S_E^{-1}E$.

2.5 Mixture identification

2.5.1 NMF with under-approximation and sparsity constraints: the SNMU

Non-negative matrix factorisation (NMF) (Lee and Seung, 2001) consists in factorising a non-negative matrix by the product of two low-rank nonnegative matrices. The NMF is performed using an optimisation method with a non-negativity constraint. Applied to the $(P \times N)$ exposure matrix E and for a given number of factorial dimensions K designating the number of exposure systems, the

optimal approximation W and H of dimensions $(P \times K)$ and $(K \times N)$, respectively was obtained by minimising the following equation:

$$\min_{W \in \mathbb{R}^{P \times K},\ H \in \mathbb{R}^{K \times N}} \|E - WH\|_F^2 \ such\ that\ W \geq 0\ and\ H \geq 0.$$

In the exposure systems matrix W, each column represents an exposure system and each element w_{pk} gives the contribution of the substance p in the exposure system k. The matrix H defines the individuals' coefficient matrix and each element h_{kn} gives the contribution of the exposure system k to an individual's overall exposure n. The NMF equation was solved by a multiplicative algorithm based on a gradient descent approach (Lee and Seung, 2001) and exposure systems were extracted at the same time.

Recently, Gillis and Glineur (2010) introduced a new approach called non-negative matrix under-approximation (NMU) and proposed a recursive algorithm based on Lagrangian relaxation to solve the NMF equation. The principle of the recursive algorithm was to find exposure systems one by one from the exposure matrix E. To ensure non-negativity, Gillis and Glineur (2010) added a new constraint to the optimisation process: $WH \leq E$ i.e. each exposure e_{pn} must be greater than its approximation $w_{pk}h_{kn}$.

As substances can, in some cases, have small coefficients relative to the other system contributors, Gillis and Plemmons (2013) proposed adding a sparsity constraint to enforce separation of the lowest contributors using a penalty $\mu \geq 0$ on the exposure matrix W. Thus, the new model, called sparse NMU (SNMU), is defined by the following optimisation problem (Gillis and Plemmons, 2013):

$$\min_{W \in \mathbb{R}^{P \times K},\ H \in \mathbb{R}^{K \times N}} \|E - WH\|_F^2 + \mu \|W\|_0 \ such\ that\ W \geq 0, H \geq 0\ and\ WH \leq E,$$

Finally, to define the mixture from an exposure system, substances were ordered by decreasing contribution to the exposure system expressed in percentage (%Subst in Tables 1, 2, 3 and 4). For practical purposes, it was chosen to present the first 25 substances in each mixture.

The SNMU was recoded on R (Team, 2016) software using the implementation algorithm on Matlab software by Gillis and Plemmons (2013). The original code is available on their website: https://sites.google.com/site/nicolasgillis/code.

2.5.2 Hierarchical clustering method

Following the SNMU, a clustering method was applied to identify clusters of women with similar co-exposure profiles. Hierarchical agglomerative clustering was applied, using the *hclust* function in R (Team, 2016) software, on the matrix H containing the contributions of each exposure system the individuals' exposures. The Euclidian distance between individuals and the Ward aggregation criterion were used as parameters in the *hclust* function. The principle of the *hclust* method is to start with N clusters: every individual n represents a cluster. In the second step, two clusters are grouped into a bigger one, on the basis of their distance. After step $N-1$, a single cluster composed of all the individuals is obtained. Given that at each level of aggregation, some inter-class inertia is lost, the optimal number of clusters is that preceding a significant loss of inertia.

The *catdes* function from the *FactoMineR* package (Lê et al., 2008) was used to determine the principal exposure systems characterising each cluster. It was also used to compare the mean of characteristics (age, BMI before pregnancy and weight gain) and exposure of the individuals from the different clusters with those of the overall population. The *p*-value from the F-test in one-way analysis of variance of the *catdes* function was compared with the standard *p*-value 0.05 to indicate a significant difference between individuals in a cluster and the whole population.

2.5.3 Selection of the number $\boldsymbol{K}$ of exposure systems and sparsity parameter

The determination of the optimal number of exposure systems K remains a challenge with NMF. In their studies, Zetlaoui et al. (2011), Sy et al. (2013) and Béchaux et al. (2013) used a criterion based on the residual sum of squares between the exposure and the estimated matrices. With the large number of substances, this criterion does not produce results allowing an optimal choice of K but gives indications on candidate values to the optimal choice. In their studies, Gazan et al. (2016) and Traoré et al. (2016) supplemented the residual sum of squares criterion using their indications by

the relevance and quality of interpretation of clustering results. In this study, the residual sum of squares criterion was used initially and consists in applying the SNMU on exposure matrix E with a set of values K and of plotting the residual sum of squares for each value K. The criterion based on the relevance and quality of interpretation consisted in using the candidate values obtained from the residual sum of squares criterion and in examining how the exposure systems characterise the clusters. The K values for which one or more exposure systems were not used to characterise a cluster were rejected.

As proposed by Gillis and Plemmons (2013), the SNMU requires human supervision to adjust the penalty parameter μ. This parameter was applied to enforce sparsity on each column of the exposure matrix W. Gillis and Plemmons (2013) introduced a parameter $\lambda \in [0, 1[^k$ and each λ_k is proportional to the penalty parameter. A lower and upper bound, respectively δ and Δ, were imposed on the sparsity λ_k of each column of the exposure systems matrix W: new parameters were then defined $0 \leq \delta < \Delta \leq 1$. The value of parameter λ_k is decreased (resp. increased) when the lower (resp. upper) bound is reached for a given iteration (Gillis and Plemmons, 2013). The SNMU was applied to the exposure data E and several values of the parameters λ, δ and Δ were tested regarding their effect on the mixture weights.

3 Results

3.1 The number K of exposure systems and the sparsity coefficient

The residual sum of squares criterion was first applied to exposure data E from EDEN before pregnancy and candidate values to the optimal choice of K were around $K = 8$. Secondly, the criterion based on the relevance and quality of interpretation was applied with values $K = 7, 8, 9$ and 10. With 7 and 8, each exposure characterises at least one cluster, whereas with 9 and 10 one or two exposure systems were unused to characterise a cluster. Thus, the optimal choice of the number K of exposure systems was set to 8.

For each k, the lower bound δ and the upper bound Δ were set to 0.2 and 0.99. The selected values for λ were $0.95, 0.8, 0.65, 0.5, 0.4, 0.4, 0.4$ and 0.4 for the different k values.

3.2 EDEN and ELFE mixtures

The mixtures obtained on the basis of data from EDEN for before (Table 1) and during pregnancy (Table 2) and those from ELFE data (Table 3) were very similar. Overall, six mixtures were common to the three datasets. The tables 1 to 3 present the contribution (%) of each substance to the mixture composed of 25 substances. The total contribution of the 25 substances per exposure system is also presented.

3.2.1 Common mixtures

The "TE-F-PAH" mixture, identified from EDEN exposure data before pregnancy and composed mostly of trace elements and furans, was also identified from EDEN and ELFE exposure data during pregnancy. The substances common to each mixture were composed mostly of 11 trace elements (inorganic arsenic, chromium VI, vanadium, lead, tellurium, chromium III, cobalt, nickel, cadmium, barium and aluminium), five furans (OCDF, HCDF-234678, 1234789, 123478 and 1234678) and the PAH pyrene. In EDEN before pregnancy, these substances were also associated with two other trace elements (germanium and strontium), five PAHs (benzo[g,h,i]perylene, indeno[1,2,3-cd]pyrene, cyclopenta(c,d)pyrene, benzo[a]pyrene and fluoranthene) and PFOA. In EDEN during pregnancy, they were also associated with trace elements (germanium and strontium), PAHs (benzo[g,h,i]perylene and fluoranthene), dioxins (HCCD-123678 and 1234678), furan (HCDF-123678) and PFOA. In ELFE, they were also associated with trace elements (antimony and lithium), dioxins (HCDD-123678, 1234678 and 234678), furans (HCDF-123678 and 123789) and zearalenone.

The "PCB-BFR-Aso-MeHg" mixture identified from EDEN before pregnancy was also identified from EDEN and ELFE during pregnancy. For the common substances, the mixtures were composed of three NDL-PCBs (PCB-101, 153 and 180), three DL-PCBs (PCB-156, 167 and 189), seven BFRs (PBDE-28, 47, 100, and 154; PBB-52, 101 and 153) and two trace elements (organic arsenic and methyl mercury). In EDEN before pregnancy, the mixture also contained NDL-PCB 138, five DL-PCBs (PCB-77, 105, 123,

157, and 169), three PFAAs (PFOS, PFUnA and PFTrDA) and a furan (TCDF 2378). In EDEN during pregnancy, the mixture also contained three PFAAs (PFUnA, PFTrDA and PFOS), NDL-PCB 28, two DL-PCBs (PCB-81 and 123), two BFRs (PBDE-99 and 153), a furan (PCDF-12378) and inorganic mercury. In ELFE during pregnancy, the mixture also contained six DL-PCBs (PCB-77, 105, 114, 118, 123 and 157), two NDL-PCBs (PCB-52 and 138) and two furans (PCDF-12378 and TCDF-2378).

The “Pest-1” mixture composed mainly of pesticides and identified from EDEN before pregnancy was also identified from EDEN and ELFE during pregnancy. It was composed of four carbamates (carbendazim, thiabendazole, pirimicard and methomyl), four organophosphates (phosmet, phosalone, azinphos-methyl and chlorpyrifos-ethyl), two dicarboximides (captan and folpet), three benzoylureas (diflubenzuron, triflumuron and teflubenzuron), seven other pesticides (propargite, ethoxyquin, diphenylamine, tebufenozide, fludioxonil, boscalid and bifenthrin) and a mycotoxin (patulin). This mixture was also associated with four pesticides (pyrimethanil, cyprodinyl, imazalil and iprodione) in EDEN before pregnancy and four pesticides (pyrimethanil, phenylphenol, lambda-cyhalothrin and cyprodinyl) in ELFE, whereas in EDEN during pregnancy, it was also associated with two pesticides (ethion and imazalil), a PFAA (PFNA) and a phytoestrogen (daidzein).

In addition, the “Pest-2” mixture identified from EDEN before pregnancy and composed mostly of pesticides was also identified from EDEN and ELFE during pregnancy. It was composed of two pyrethroids (acrinathrin and lambda-cyhalothrin), two strobilurns (kresoxim-methyl and azoxystrobin), three organophosphates (dichlorvos, dimethoate and chlorpyrifos-ethyl), three triazoles (penconazole, tebuconazole and fenbuconazole), two dicarboximides (procymidone and iprodione) and seven other pesticides (endosulfan, bupirimate, mepanipyrim, fenhexamid, cyprodinyl, metalaxyl and boscalid). This mixture was also associated with six pesticides (myclobutanil, pyrimethanil, tetradifon, pyriproxyfen, diethofencarb and chlorothalonil) in EDEN before pregnancy, three pesticides (fludioxonyl, phosmet and chlorothalonil), two trace elements (nickel and silver) and an additive (sulphites) in EDEN during pregnancy. In ELFE, ”Pest-2” was also

associated with four pesticides (phosmet, fludioxonil, chlorthal dimethyl and pyriproxyfen), a trace element (nickel) and an additive (sulphites).

The mixture “Pest-3” identified from EDEN before pregnancy and composed mainly of pesticides, was also identified from EDEN and ELFE during pregnancy. It was composed of two pyrethroids (cyfluthrin and bifenthrin), three triazoles (tetraconazole, triadimenol and myclobutanil), two carbamates (methomyl and metalaxyl) and nine other pesticides (trifloxystrobin, spiroxamine, quinoxyfen, etofenprox, tebufenpyrad, pyrimethanil, fenhexamid, teflubenzuron and cyprodinyl). In EDEN before pregnancy, this mixture was also associated with four PFAAs (PFHxA, PFBS, PFHxS and PFHpA) and five pesticides (cyproconazole, sulphur, pyriproxyfen, diethofencarb and chlorothalonil). Eight pesticides (mepanipyrim, penconazole, boscalid, chlorpyrifos-methyl, chlorpyrifos-ethyl, iprodione, azoxystrobin and lambda-cyhalothrin) and a trace element (silver) were also associated with this mixture in EDEN during pregnancy. In ELFE, nine pesticides (mepanipyrim, penconazole, boscalid, iprodione, chlorpyrifos-ethyl, lambda-cyhalothrin, procymidone, azoxystrobin and chlorpyrifos-methyl) were also associated with this mixture.

The main substances identified in the “PFAA-Ge-Li” mixture from EDEN before pregnancy were also identified from EDEN and ELFE during pregnancy. This mixture was composed of five PFAAs (PFHxS, PFBS, PFHxA, PFHpA and PFOA) and two trace elements (germanium and lithium). In EDEN before pregnancy, this mixture was also associated with seven pesticides (pyriproxyfen, tetradifon, sulphur, diethofencarb, chlorothalonil, cyproconazole and bifenthrin), six other trace elements (strontium, vanadium, inorganic arsenic, gallium, inorganic mercury and antimony), three other PFAAs (PFOS, PFTeDA and PFDoA), a PAH (dibenzo[a,h]pyrene) and a mycotoxin (patulin). In EDEN during pregnancy, this mixture also contained six other trace elements (strontium, inorganic arsenic, vanadium, chromium VI, inorganic mercury and tellurium), six phytoestrogens (formononetin, equol, enterolactone, matairesinol, glycitein and resveratrol), a PAH (5-methylchrysene), two mycotoxins (DON15 and FB2) and three pesticides (malathion, phenylphenol and carbaryl). In ELFE, this mixture was also associated with nine other pesticides (carbendazim, phenylphenol, pyriproxyfen, tetradifon,

lindane, sulphur, carbofuran, diethofencarb and chlorothalonil), three other trace elements (chromium VI, tin and gallium), four mycotoxins (FB1, OTA, FB2 and DON15), a PFAAs (PFOS) and a phytoestrogen (equol).

3.2.2 Other Mixtures

Five mixtures were found in only one or two datasets.

The “Myco-Pest-PAH” mixture identified from EDEN before pregnancy was also found in EDEN during pregnancy. It contained eight mycotoxins (alpha-zearalanol, alpha-zearalenol, DAS, DON3, FusX, OTB, OTA and HT2-toxin), three pesticides (chlorpyrifos-methyl, cyproconazole and pirimiphos-methyl) and four PAHs (benzo[g,h,i]perylene, benzo[e]pyrene, cyclopenta(c,d)pyrene and indeno[1,2,3-cd]pyrene). In EDEN before pregnancy, these substances were associated with nine other pesticides (pyriproxyfen, tetradifon, sulphur, chlorothalonil, diethofencarb, flutriafol, iprodione, ethion and bifenthrin) and an additive (sulphites). In EDEN during pregnancy, these substances were associated with three other mycotoxins (DON, DON15 and zearalenone), a PAH (pyrene), two phytoestrogens (daidzein and genistein), a trace element (gallium), a pesticide (sulphur) and two PFAAs (PFBS and PFHxS).

The “Mixt-1” mixture from EDEN before pregnancy was composed of acrylamide, seven mycotoxins (T-2 toxin, nivalenol, OTB, FB1, MAS, FB2, and DON15), six PAHs (dibenzo[a,i]pyrene, anthracene, benzo[c]fluorine, 5-methylchrysene, dibenzo[a,l]pyrene and benzo[k]fluoranthene), four pesticides (chlorpropham, hexachlorobenzene, lindane and phenylphenol), but also some additives (nitrites and tartaric acid), phytoestrogens (glycitein, enterolactone, equol and formononetin) and a trace element (tin).

The “Mixt-2” mixture from EDEN during pregnancy was composed of nine DL-PCBs (PCB-77, 105, 114, 118, 126, 156, 157, 167 and 169), three NDL-PCBs (PCB-52, 138 and 153), two BFRs (HBCD alpha and PBDE-47), seven PAHs (chrysene, benzo[a]anthracene, benzo[j]fluoranthenen, benzo[b]fluoranthene, benzo[k]Fluoranthene, benzo[c]fluorine and bibenzo[a,e]pyrene), inorganic mercury, a furan (TCDF-2378), bisphenol A and T2 toxin.

Two other mixtures were identified from ELFE during pregnancy only and were composed of heterogeneous substances: the “Mixt-3” and ”Mixt-4” mixtures. The “Mixt-3” mixture was composed of PFAAs (PFTeDA, PFDoA, PFDA, PFOS and PFNA), fourteen PAHs (benzo[k]fluoranthene, benzo[b]fluoranthene, benzo[j]fluoranthene, dibenzo[a,i]pyrene, chrysene, benzo[a]anthracene, dibenzo[a,e]pyrene, benzo[a]pyrene, indeno[1,2,3-cd]pyrene, dibenzo[a,h]anthracene, dibenzo[a,l]pyrene, benzo[c]fluorene, benzo[g,h,i]perylene and fluoranthene), trace elements (silver, organic arsenic and cadmium), BFRs (HBCD beta and gamma) and dioxin OCDD. The “Mixt-4” mixture contained mostly pesticides with among others chlorothalonil, three carbamates (diethofencarb, carbofuran and metalaxyl), two dicarboximides (procymidone and iprodione), two pyrethroids (bifenthrin and lambda-cyhalothrin) and three triazoles (cyproconazole, flutriafol and tebuconazole). “Mixt-4” also contained trace elements (tin and lithium) and a phytoestrogen (secoisolariciresinol).

Table 1: Mixtures identified from EDEN before pregnancy. The percentage of explained variance by each mixture is presented in parentheses. The percentage of each substance in the mixture is given in the column %Subst. The contribution of the 25 substances to the exposure system composed of the total number of substances is presented in "total contribution".

EDEN mixtures before pregnancy															
TE-F-PAH (30.5%)	**% Subst**	**PCB-BFR-Aso-MeHg (11.3%)**	**% Subst**	**Pest-1 (8.6%)**	**% Subst**	**Mixt-1 (15.8%)**	**% Subst**	**Pest-2 (8.4%)**	**% Subst**	**Pest-3 (7.6%)**	**% Subst**	**PFAA-Ge-Li (11.6%)**	**% Subst**	**Myco-Pest-PAH (6.2%)**	**% Subst**
Inorganic arsenic	5.51	PCB 153	4.54	Carbendazim	5.26	Acrylamide	7.10	Acrinathrin	6.83	Cyfluthrin	9.20	PFHxS	17.1	Alpha-zearalanol	15.1
Germanium	5.37	PCB 157	4.51	Thiabendazole	4.91	T-2 toxin	6.52	Endosulfan	6.67	Tetraconazole	9.20	PFBS	17.1	Alpha-zearalenol	15.1
Chromium VI	5.26	PCB 138	4.50	Propargite	4.89	Nivalenol	6.16	Bupirimate	6.65	Trifloxystrobin	9.20	PFHxA	16.7	DAS	15.1
Strontium	4.83	PCB 167	4.49	Phosmet	4.84	Dibenzo[a,i]pyrene	5.52	Kresoxim-methyl	6.64	Spiroxamine	9.20	PFHpA	16.4	DON3	15.1
Vanadium	4.55	PCB 156	4.43	Pirimicarb	4.79	Anthracene	5.46	Dichlorvos	6.62	Quinoxyfen	9.20	PFOA	7.85	FusX	15.1
Lead	4.42	PCB 123	4.43	Ethoxyquin	4.74	OTB	5.26	Penconazole	6.43	Etofenprox	9.20	Strontium	5.91	OTB	2.77
Tellurium	4.35	PBDE 47	4.41	Phosalone	4.72	FB1	5.24	Mepanipyrim	6.19	Tebufenpyrad	9.20	Germanium	4.68	OTA	2.33
Chromium III	4.19	PCB 180	4.35	Diphenylamine	4.72	Chlorpropham	4.93	Azoxystrobin	5.61	Triadimenol	9.14	Lithium	2.64	Chlorpyrifos-methyl	1.95
Cobalt	4.17	PCB 189	4.31	Azinphos-methyl	4.71	Nitrites	4.90	Fenhexamid	4.77	Myclobutanil	4.11	Pyriproxyfen	1.09	Benzo[g,h,i]perylene	1.94
Nickel	3.87	PCB 101	4.24	Captan	4.71	Glycitein	4.68	Tebuconazole	4.61	Pyrimethanil	3.54	Tetradifon	1.06	Benzo[e]pyrene	1.76
Cadmium	3.78	PBDE 100	4.14	Diflubenzuron	4.71	MAS	4.18	Dimethoate	4.60	Fenhexamid	3.36	Sulphur	1.04	Pyriproxyfen	1.39
Benzo[g,h,i]perylene	3.71	PBDE 154	4.12	Tebufenozide	4.71	Benzo[c]fluorene	4.13	Fenbuconazole	4.57	Bifenthrin	2.65	Diethofencarb	0.97	Tetradifon	1.33
Barium	3.69	PBDE 28	4.00	Triflumuron	4.71	FB2	4.00	Procymidone	4.18	Teflubenzuron	1.65	Chlorothalonil	0.97	Sulphur	1.29
Pyrene	3.67	PBB 153	3.97	Folpet	4.71	Tartaric acid	3.63	Myclobutanil	4.14	Methomyl	1.28	Vanadium	0.95	Cyclopenta(c,d)pyrene	1.14
Indeno[1,2,3-cd]pyrene	3.67	Organic arsenic	3.97	Methomyl	4.54	DON15	3.33	Cyprodinyl	3.81	Metalaxyl	1.21	Cyproconazole	0.80	Cyproconazole	1.11
OCDF	3.60	PFOS	3.94	Teflubenzuron	4.39	Hexachlorobenzene	3.07	Iprodione	3.26	Cyprodinyl	1.02	Inorganic arsenic	0.74	Chlorothalonil	1.08
HCDF 234678	3.55	PCB 169	3.92	Chlorpyrifos-ethyl	4.05	Lindane	2.81	Lambda-cyhalothrin	3.06	PFHxA	0.90	PFOS	0.67	Diethofencarb	1.08
HCDF 1234789	3.55	TCDF 2378	3.80	Fludioxonil	3.92	Phenylphenol	2.77	Metalaxyl	2.88	PFBS	0.90	Dibenzo[a,h]pyrene	0.62	Indeno[1,2,3-cd]pyrene	0.89
Aluminium	3.51	PBB 52	3.57	Boscalid	3.88	5-methylchrysene	2.70	Boscalid	2.46	PFHxS	0.90	Bifenthrin	0.42	Pirimiphos-methyl	0.87
HCDF 123478	3.51	PCB 77	3.52	Bifenthrin	2.91	Enterolactone	2.50	Chlorpyrifos-ethyl	1.15	PFHpA	0.89	Gallium	0.41	Flutriafol	0.71
PFOA	3.51	PBB 101	3.48	Patulin	2.52	Dibenzo[a,l]pyrene	2.45	Pyrimethanil	1.06	Cyproconazole	0.84	PFTeDA	0.40	Iprodione	0.65
HCDF 1234678	3.48	PCB 105	3.47	Pyrimethanil	1.95	Equol	2.42	Tetradifon	0.99	Sulphur	0.83	PFDoA	0.40	Ethion	0.63
Cyclopenta(c,d)pyrene	3.46	Methyl mercury	3.38	Cyprodinyl	1.74	Benzo[k]fluoranthene	2.25	Pyriproxyfen	0.99	Pyriproxyfen	0.82	Inorganic mercury	0.40	Sulphites	0.61
Benzo[a]pyrene	3.41	PFUnA	3.25	Imazalil	1.72	Tin	2.18	Diethofencarb	0.91	Diethofencarb	0.79	Patulin	0.39	Bifenthrin	0.51
Fluoranthene	3.40	PFTrDA	3.23	Iprodione	1.24	Formononetin	1.82	Chlorothalonil	0.91	Chlorothalonil	0.79	Antimony	0.37	HT2	0.47
Total contribution = 44%		**Total contribution = 87%**		**Total contribution = 95%**		**Total contribution = 72.7%**		**Total contribution = 95%**		**Total contribution = 87.8%**		**Total contribution = 93.6%**		**Total contribution = 94.4%**	
79 substances with %Subst > 0		**42 substances with %Subst > 0**		**42 substances with %Subst > 0**		**71 substances with %Subst > 0**		**47 substances with %Subst > 0**		**78 substances with %Subst > 0**		**60 substances with %Subst > 0**		**54 substances with %Subst > 0**	

Table 2: Mixtures identified from EDEN during pregnancy. The percentage of explained variance by each mixture is presented in parentheses. The percentage of each substance in the mixture is given in the column %Subst. The contribution of the 25 substances to the exposure system composed of the total number of substances is presented in "total contribution".

EDEN mixtures during pregnancy														
TE-F-PAH (39.1%)	**% Subst**	**Mixt-2 (19.5%)**	**% Subst**	**Pest-1 (7.8%)**	**% Subst**	**Pest-3 (4.2%)**	**% Subst**	**PCB-BFR-Aso-MeHg (5.1%)**	**% Subst**	**Pest-2 (5.5%)**	**% Subst**	**Myco-Pest-PAH (7.2%)**	**% Subst**	**PFAA-Ge-Li (11.5%)**	**% Subst**
Inorganic arsenic	5.70	PCB 126	4.74	Phosalone	6.23	Myclobutanil	6.26	PFTrDA	6.24	Acrinathrin	7.04	Alpha-zearalanol	12.2	PFHxS	15.5
Chromium VI	5.56	PCB 118	4.65	Diphenylamine	6.21	Triadimenol	6.09	PFUnA	6.15	Endosulfan	6.85	Alpha-zearalanol	12.2	PFBS	15.5
Vanadium	5.37	PCB 52	4.55	Azinphos-methyl	6.21	Cyfluthrin	6.07	Methyl mercury	5.97	Bupirimate	6.84	DAS	12.2	PFHxA	14.7
Germanium	5.29	PCB 114	4.43	Tebufenozide	6.21	Tetraconazole	6.07	PBB 101	5.83	Dichlorvos	6.84	DON3	12.2	PFHpA	14.3
Strontium	4.94	PCB 77	4.43	Diflubenzuron	6.21	Spiroxamine	6.07	PBB 52	5.36	Kresoxim-methyl	6.65	FusX	12.2	PFOA	7.50
Chromium III	4.34	PCB 105	4.38	Captan	6.21	Quinoxyfen	6.07	PBB 153	4.54	Azoxystrobin	6.47	OTB	4.61	Strontium	6.00
Tellurium	4.29	HBCD alpha	4.32	Triflumuron	6.21	Etofenprox	6.07	PBDE 100	4.42	Penconazole	5.15	OTA	4.23	Germanium	4.52
Lead	4.23	Chrysene	4.30	Folpet	6.21	Trifloxystrobin	6.07	PFOS	4.41	Tebuconazole	4.87	Chlorpyrifos-methyl	3.93	Lithium	3.01
Cobalt	3.92	Benzo[a]anthracene	4.28	Pirimicarb	6.21	Tebufenpyrad	6.07	PBDE 154	4.23	Dimethoate	4.86	Benzo[g,h,i]perylene	3.58	Formononetin	2.56
HCDF 1234789	3.84	Benzo[j]fluoranthene	4.26	Ethoxyquin	6.19	Fenhexamid	6.03	PBDE 28	4.17	Fenbuconazole	4.83	HT-2 toxin	2.83	Equol	2.38
HCDF 234678	3.78	Benzo[b]fluoranthene	4.07	Propargite	5.89	Pyrimethanil	5.15	PCB 28	4.14	Procymidone	4.50	Benzo[a]pyrene	2.82	Enterolactone	2.29
HCDF 1234678	3.75	PCB 169	4.01	Thiabendazole	5.49	Bifenthrin	4.22	PCB 101	3.79	Mepanipyrim	4.38	Pirimiphos-methyl	2.81	Matairesinol	2.27
OCDF	3.75	Inorganic mercury	3.99	Carbendazim	4.80	Teflubenzuron	4.21	PBDE 153	3.61	Boscalid	4.02	Cyclopenta(c,d)pyrene	2.43	Glycitein	1.74
HCDF 123478	3.73	TCDF 2378	3.97	Phosmet	4.29	Methomyl	3.95	PCB 123	3.58	Lambda-cyhalothrin	3.96	Indeno[1,2,3-cd]pyrene	2.22	Inorganic arsenic	1.22
PFOA	3.67	Benzo[k]fluoranthene	3.95	Methomyl	3.73	Mepanipyrim	3.92	PBDE 99	3.52	Cyprodinyl	3.31	DON	1.68	Vanadium	1.05
Pyrene	3.65	Benzo[c]fluorene	3.83	Teflubenzuron	3.43	Cyprodinyl	3.40	Organic arsenic	3.46	Iprodione	3.30	Zearalenone	1.07	5-methylchrysene	0.98
HCDF 123678	3.51	PCB 157	3.76	Chlorpyrifos-ethyl	2.63	Penconazole	3.08	PCB 189	3.42	Metalaxyl	2.68	Pyrene	1.05	DON15	0.74
Barium	3.48	Bisphenol A	3.74	Patulin	1.79	Metalaxyl	1.99	PCB 81	3.21	Chlorpyrifos-ethyl	2.62	Daidzein	0.87	Chromium VI	0.57
Aluminium	3.40	PCB 138	3.70	Boscalid	1.50	Boscalid	1.80	PCB 180	3.19	Fludioxonyl	1.98	Genistein	0.85	FB2	0.55
Nickel	3.36	PCB 156	3.64	Fludioxonil	1.46	Chlorpyrifos-methyl	1.73	PCDF 12378	3.13	Nickel	1.84	Gallium	0.69	Inorganic mercury	0.55
Cadmium	3.34	PCB 153	3.54	Ethion	0.77	Chlorpyrifos-ethyl	1.62	PCB 167	2.98	Silver	1.84	Cyproconazole	0.68	Tellurium	0.52
Benzo[g,h,i]perylene	3.34	Dibenzo[a,e]pyrene	3.41	Imazalil	0.64	Iprodione	1.40	PBDE 47	2.94	Phosmet	1.80	PFBS	0.68	Malathion	0.43
HCDD 123678	3.26	PCB 167	3.39	Bifenthrin	0.59	Azoxystrobin	1.09	PCB 153	2.70	Fenhexamid	1.41	PFHxS	0.68	Phenylphenol	0.40
HCDD 1234678	3.26	PBDE 47	3.33	PFNA	0.54	Lambda-cyhalothrin	0.91	PCB 156	2.53	Sulphites	1.00	DON15	0.65	Carbaryl	0.40
Fluoranthene	3.24	T-2 toxin	3.31	Daidzein	0.32	Silver	0.63	Inorganic mercury	2.47	Chlorothalonil	0.97	Sulphur	0.64	Resveratrol	0.37
Total contribution = 51.8%		**Total contribution = 48.4%**		**Total contribution = 96.5%**		**Total contribution = 94.2%**		**Total contribution = 75.9%**		**Total contribution = 88.9%**		**Total contribution = 91.6%**		**Total contribution = 93%**	
64 substances with %Subst > 0		**94 substances with %Subst > 0**		**53 substances with %Subst > 0**		**49 substances with %Subst > 0**		**71 substances with %Subst > 0**		**61 substances with %Subst > 0**		**56 substances with %Subst > 0**		**58 substances with %Subst > 0**	

Table 3: Mixtures identified from ELFE during pregnancy. The percentage of explained variance by each mixture is presented in parentheses. The percentage of each substance in the mixture is given in the column %Subst. The contribution of the 25 substances to the exposure system composed of the total number of substances is presented in "total contribution".

ELFE mixtures during pregnancy														
TE-F-PAH (31.4%)	**% Subst**	**PCB-BFR-Aso-MeHg (10.7%)**	**% Subst**	**Pest-1 (12.9%)**	**% Subst**	**Pest-3 (6.9%)**	**% Subst**	**PFAA-Li-Ge (16.0%)**	**% Subst**	**Pest-2 (7.5%)**	**% Subst**	**Mixt-3 (4.4%)**	**% Subst**	**Mixt-4 (10.2%)**	**% Subst**
Chromium VI	4.92	PCB 123	4.61	Propargite	6.27	Myclobutanil	5.71	PFHxA	13.9	Bupirimate	8.32	PFTeDA	7.57	Chlorothalonil	11.2
Chromium III	4.66	PBDE 47	4.46	Thiabendazole	6.17	Fenhexamid	5.64	PFHpA	13.9	Dichlorvos	8.28	Benzo[k]fluoranthene	7.55	Diethofencarb	11.2
Lead	4.62	PBDE 100	4.42	Pirimicarb	6.12	Triadimenol	5.48	PFHxS	13.8	Endosulfan	8.25	Benzo[b]fluoranthene	7.50	Carbofuran	11.2
Cobalt	4.62	PCB 52	4.42	Phosalone	6.10	Etofenprox	5.42	PFBS	13.8	Acrinathrin	8.24	Benzo[j]fluoranthene	7.31	Tetradifon	11.2
Barium	4.43	PCB 101	4.40	Diphenylamine	6.09	Trifloxystrobin	5.42	PFOA	10.4	Kresoxim-methyl	8.05	Dibenzo[a,i]pyrene	7.12	Pyriproxyfen	11.1
Tellurium	4.43	PBDE 154	4.40	Diflubenzuron	6.09	Quinoxyfen	5.42	Carbendazim	5.42	Azoxystrobin	7.42	PFDoA	6.34	Chlorthal-dimethyl	9.11
Vanadium	4.38	PCB 77	4.36	Folpet	6.09	Spiroxamine	5.42	Chromium VI	2.87	Penconazole	5.47	Chrysene	6.01	Procymidone	6.93
Nickel	4.36	PCB 153	4.30	Azinphos-methyl	6.09	Cyfluthrin	5.42	FB1	2.42	Boscalid	4.91	Benzo[a]anthracene	4.94	Pyrimethanil	4.71
Inorganic arsenic	4.15	PBB 153	4.23	Captan	6.09	Tetraconazole	5.42	OTA	2.11	Tebuconazole	4.75	Dibenzo[a,e]pyrene	4.79	Metalaxyl	4.29
Antimony	3.98	PCB 105	4.23	Ethoxyquin	6.09	Tebufenpyrad	5.42	FB2	2.04	Dimethoate	4.67	Benzo[a]pyrene	4.58	Iprodione	3.35
HCDF 123678	3.96	PBDE 28	4.20	Triflumuron	6.09	Methomyl	4.88	Lithium	1.81	Fenbuconazole	4.65	Silver	4.16	Bifenthrin	2.54
Pyrene	3.94	PBB 52	4.15	Tebufenozide	6.09	Teflubenzuron	4.63	Phenylphenol	1.65	Mepanipyrim	4.33	Indeno[1,2,3-cd]pyrene	3.78	Sulphur	2.44
Aluminium	3.89	PCB 157	4.14	Carbendazim	5.86	Pyrimethanil	4.48	Tin	1.49	Procymidone	3.99	Dibenzo[a,h]anthracene	3.32	Cyproconazole	2.01
HCDF 123478	3.83	TCDF 2378	4.07	Phosmet	3.54	Cyprodinyl	4.22	Pyriproxyfen	1.45	Phosmet	3.00	HBCD beta	2.74	Tin	1.33
HCDF 1234789	3.83	PCB 167	4.02	Patulin	2.96	Bifenthrin	4.11	PFOS	1.40	Chlorpyrifos-ethyl	2.74	PFDA	2.74	Cyprodinyl	1.32
OCDF	3.77	PBB 101	3.98	Fludioxonil	2.45	Mepanipyrim	3.99	Germanium	1.30	Metalaxyl	2.66	Dibenzo[a,l]pyrene	2.61	Lambda-cyhalothrin	1.15
Lithium	3.70	PCB 114	3.85	Chlorpyrifos-ethyl	2.21	Metalaxyl	3.80	Tetradifon	1.18	Lambda-cyhalothrin	2.48	PFOS	2.45	Fludioxonil	1.10
HCDF 123789	3.68	PCDF 12378	3.74	Teflubenzuron	1.75	Penconazole	3.23	DON15	1.17	Cyprodinyl	2.11	Benzo[c]fluorene	2.30	Secoisolariciresino	0.71
HCDD 123678	3.64	PCB 138	3.73	Boscalid	1.66	Boscalid	2.55	Lindane	1.15	Fludioxonyl	1.61	Organic arsenic	2.11	Flutriafol	0.68
HCDD 123789	3.62	PCB 189	3.65	Methomyl	1.26	Iprodione	2.03	Sulphur	1.15	Iprodione	1.23	PFNA	2.00	Imidacloprid	0.68
HCDF 1234678	3.60	Organic arsenic	3.51	Bifenthrin	1.14	Chlorpyrifos-ethyl	1.96	Gallium	1.14	Fenhexamid	0.67	Benzo[g,h,i]perylene	1.98	Boscalid	0.49
HCDF 234678	3.58	Methyl mercury	3.43	Pyrimethanil	1.06	Lambda-cyhalothrin	1.43	Equol	1.13	Sulphites	0.62	Cadmium	1.86	Tebuconazole	0.40
HCDD 123478	3.51	PCB 118	3.39	Phenylphenol	0.96	Procymidone	1.37	Carbofuran	1.12	Nickel	0.62	Fluoranthene	1.76	Lithium	0.37
Cadmium	3.51	PCB 180	3.21	Lambda-cyhalothrin	0.88	Azoxystrobin	1.28	Diethofencarb	1.12	Chlorthal-dimethyl	0.49	HBCD gamma	1.26	Chlorfenvinphos	0.33
Zearalenone	3.40	PCB 156	3.10	Cyprodinyl	0.86	Chlorpyrifos-methyl	1.27	Chlorothalonil	1.12	Pyriproxyfen	0.43	OCDD	1.22	Lindane	0.28
Total contribution = 47%		**Total contribution = 68%**		**Total contribution = 93.5%**		**Total contribution = 92.1%**		**Total contribution = 82.8%**		**Total contribution = 95%**		**Total contribution = 85.2%**		**Total contribution = 97.2%**	
85 substances with %Subst > 0		**80 substances with %Subst > 0**		**52 substances with %Subst > 0**		**58 substances with %Subst > 0**		**72 substances with %Subst > 0**		**58 substances with %Subst > 0**		**62 substances with %Subst > 0**		**50 substances with %Subst > 0**	

3.3 ELFE exposure profiles

The clustering method applied to matrix H from the factorisation of the ELFE exposure matrix led to division of the $15{,}226$ subjects in the overall study population into six clusters. A cluster is based on women with similar co-exposure profiles. For each cluster, co-exposure profiles were linked to the identified mixtures of the ELFE cohort presented in Table 3. Within each cluster, mixtures covering at least 65% of the total exposure were described. The exposure (mean, P95 and P99) of the cluster and of the whole population was assessed for each mixture compound (Table 4). The average age, BMI before pregnancy and weight gain during pregnancy are also presented (Table 4). The whole study population of $15{,}226$ pregnant women in France had a mean age of 30.7 years $[P_{2.5}; P_{97.5}] = [21.1; 40.7]$, a mean BMI before pregnancy of $23.5\ kg/m^2$ $[P_{2.5}; P_{97.5}] = [17.6; 36.3]$, and a mean weight gain during pregnancy of $13.1\ kg$ $[P_{2.5}; P_{97.5}] = [2.0; 25.0]$.

Cluster 1 contains $5{,}090$ women (33.4%) with a mean age of 30.4 years and a mean weight gain during pregnancy of $13.0\ kg$. This cluster is characterised at 36.6% by "PFAA-Ge-Li" mixture and TE-F-PAH" mixture "at 31%. Significantly higher exposures were observed for PFAAs and carbendazim from mixture "PFAA-Ge-Li" for individuals in this cluster compared to the whole ELFE population.

Members of cluster 2, which represents nearly 10% $(1{,}520$ women) of the overall study population, were significantly older (mean age of 32.0 years) and had a mean BMI ($22.9\ kg/m^2$) significantly lower than the whole ELFE population. This cluster is characterised by the "Pest-1" mixture (40.5%) and "TE-F-PAH" mixture (24.8%). Individuals in this cluster were significantly more exposed to substances from "Pest-1" mixture compared to the whole ELFE population.

Cluster 3 is composed of $5{,}183$ women (34%) with a mean age of 30.1 and a mean BMI of $23.6\ kg/m^2$. This cluster is characterised by "TE-F-PAH" mixture (49.5%), PCB-BFR-Aso-MeHg" mixture "(11.4%) and "Pest-1" mixture (9.9%). This cluster has the specificity of having no significantly higher exposure compared to the whole population, irrespective of the substance.

Cluster 4, composed of $1{,}039$ women (6.8%) with a mean age of 31.0 years was characterised by "TE-F-PAH" mixture (37.8%), "PCB-BFR-Aso-MeHg mixture" (24%) and "Mixt-3" mixture (9.7%).

Individuals in this cluster were significantly more exposed to all selected substances in these mixtures than the overall study population.

Cluster 5 includes 7% of the whole population. Members were significantly older than the overall study population (31.4 years old on average). Individuals in this cluster were mostly characterised by "Pest-3" mixture (21.4%), "Te-F-PAH" mixture (19.9%) and "Pest-2" mixture (18.6%). Women in this cluster were also significantly more exposed to all substances in these mixtures than the whole ELFE population.

Cluster 6 is composed of women significantly older (mean age of 31.8 years) and with a lower mean BMI ($23.1\ kg/m^2$) than the whole ELFE population. This cluster was associated to "Mixt-4" mixture (38.2%), "TE-F-PAH" mixture (24.4%) and "PFAA-Ge-Li" mixture (13.3%). Members of this cluster were significantly more exposed to substances from "Mixt-4" mixture compared to the whole ELFE population.

Table 4: Result of the clustering of 15,226 individuals: number of individuals (*N*), individual characteristics of the cluster (variables with * have a significantly different mean compared to the overall study population), mixtures associated with the cluster, the major substances combined in the mixture, and the exposure level comparison between cluster and whole population (value in bold corresponds to a significant difference in mean *p* value < *0.05*). The exposure unit is on unit /day.

Individuals Group		Mixtures			Cluster exposure			Population exposure		
Cluster	**Description**	**Major Mixtures**	**Major substances (unit)**	**%Subst**	**Mean**	**P95**	**P99**	**Mean**	**P95**	**P99**
1		PFAA-GE-LI (36.6 %)	**PFHxA (ng)**	11.5	**6.59**	15.5	19.5	3.26	9.93	18.6
			PFHpA (ng)	11.5	**9.57**	22.4	28.2	4.79	14.5	27.4
	N = 5,090		**PFHxS (ng)**	11.4	**4.73**	11.3	14.1	2.29	7.08	13.52
			PFBS (ng)	11.4	**1.80**	4.30	5.37	0.87	2.70	5.15
			PFOA (ng)	8.59	**2.08**	4.35	5.83	1.40	3.33	5.34
			Carbendazim (µg)	4.49	**1.18**	2.36	3.76	1.02	3.10	4.75
	Age = 30.4*		Chromium VI (µg)	2.38	54.5	96.6	143	55.5	98.9	150
	[$P_{2.5}$; $P_{97.5}$] = [21.2; 40.2]		FB1 (ng)	2.00	1120	3269	4959	1148	3335	5943
			OTA (ng)	1.75	12.9	26.8	34.3	13.1	27.6	36.1
			FB2 (ng)	1.69	609	1791	3077	626	1831	3267
			Lithium (µg)	1.50	46.5	90.2	139	53.0	103	145
	BMI = 23.6		Phenylphenol (µg)	1.37	3.53	10.0	19.5	3.84	10.5	20.2
	[$P_{2.5}$; $P_{97.5}$] = [17.6; 36.7]		Tin (µg)	1.23	120	314	783	136	389	919
			Pyriproxyfen (µg)	1.20	0.02	0.05	0.10	0.03	0.14	0.26
			PFOS (ng)	1.16	4.29	9.30	15.1	4.27	10.4	22.9
			Germanium (µg)	1.08	5.30	11.9	17.4	6.08	12.9	19.7
	WeightGain = 13.0*		Tetradifon (µg)	0.98	0.01	0.04	0.07	0.02	0.11	0.21
	[$P_{2.5}$; $P_{97.5}$] = [2.0; 24.0]		DON15	0.97	40.3	153	182	41.5	154	300
			Lindane (µg)	0.95	0.04	0.08	0.26	0.04	0.09	0.27
			Sulphur (µg)	0.95	4.34	17.9	29.2	5.48	19.5	38.1
			Gallium (µg)	0.95	0.13	0.31	0.36	0.14	0.32	0.39
			Equol	0.94	7232	18316	29931	7606	20476	31496
			Carbofuran (µg)	0.93	0.02	0.05	0.09	0.03	0.14	0.29
			Diethofencarb (µg)	0.93	0.25	0.72	1.36	0.46	2.06	4.12
			Chlorothalonil (µg)	0.93	0.07	0.21	0.41	0.14	0.61	1.23
		TE-F-PAH (31.0 %)	See cluster 4		No higher significant difference					
2		Pest-1 (40.5 %)	**Propargite (µg)**	5.86	**35.7**	57.2	69.6	9.25	34.8	62.9
			Thiabendazole (µg)	5.77	**43.9**	76.8	86.1	11.9	45.6	78.3
	N = 1,520		**Pirimicarb (µg)**	5.72	**0.12**	0.20	0.20	0.03	0.10	0.20
			Phosalone (µg)	5.71	**0.79**	1.29	1.29	0.18	0.65	1.29
			Diphenylamine (µg)	5.70	**16.9**	27.5	27.5	3.83	13.8	27.5
			Diflubenzuron (µg)	5.70	**0.005**	0.008	0.008	0.001	0.004	0.008
	Age = 32.0*		**Folpet (µg)**	5.70	**0.18**	0.30	0.30	0.04	0.15	0.30
	[$P_{2.5}$; $P_{97.5}$] = [22.9; 41.6]		**Azinphos-methyl (µg)**	5.70	**0.61**	0.99	0.99	0.14	0.50	0.99
			Captan (µg)	5.70	**0.28**	0.45	0.45	0.06	0.23	0.45
			Ethoxyquin (µg)	5.70	**6.96**	11.3	11.3	1.58	5.67	11.3
			Triflumuron (µg)	5.70	**0.16**	0.25	0.25	0.04	0.13	0.25
	BMI = 22.9*		**Tebufenozide (µg)**	5.70	**0.02**	0.04	0.04	0.005	0.02	0.04

Individuals Group		Mixtures			Cluster exposure			Population exposure		
Cluster	**Description**	**Major Mixtures**	**Major substances (unit)**	**%Subst**	**Mean**	**P95**	**P99**	**Mean**	**P95**	**P99**
	$[P_{2.5}; P_{97.5}] = [17.6; 35.2]$		**Carbendazim (µg)**	5.48	**2.51**	4.25	4.74	1.02	3.10	4.75
			Phosmet (µg)	3.31	**0.67**	1.20	2.39	0.28	1.20	1.96
			Patulin (ng)	2.77	**72.9**	148	205	40.9	111	215
			Fludioxonil (µg)	2.29	**6.34**	16.0	23.0	3.58	12.2	22.4
	WeightGain = 13.1		**Chlorpyrifos-ethyl (µg)**	2.07	**1.82**	4.39	7.95	1.13	4.92	8.94
	$[P_{2.5}; P_{97.5}] = [17.6; 35.2]$		**Teflubenzuron (µg)**	1.64	**0.16**	0.28	0.36	0.10	0.36	1.26
			Boscalid (µg)	1.55	**1.64**	3.32	4.80	1.38	4.73	10.9
			Methomyl (µg)	1.18	**0.007**	0.01	0.02	0.005	0.02	0.07
			Bifenthrin (µg)	1.07	**0.20**	0.57	0.89	0.18	0.65	1.23
			Pyrimethanil (µg)	0.99	1.91	5.34	7.22	1.80	6.66	13.7
			Phenylphenol (µg)	0.90	**4.16**	10.4	16.8	3.84	10.5	20.2
			Lambda-cyhalothrin (µg)	0.82	**0.46**	1.41	2.27	0.37	1.32	2.15
			Cyprodinyl (µg)	0.80	2.03	5.63	8.41	2.16	7.68	15.2
		TE-F-PAH (24.8 %)	See cluster 4		No higher significant difference					
3			Chromium VI (µg)	2.31	47.5	81.3	102	55.5	98.9	150
			Chromium III (µg)	2.19	291	506	685	338	626	1041
	N = 5,183		Lead (µg)	2.17	13.0	21.5	29.1	15.7	27.9	44.7
			Cobalt (µg)	2.17	13.6	23.2	29.7	15.8	28.8	46.2
			Barium (µg)	2.08	562	953	1241	653	1203	1975
			Tellurium (µg)	2.08	2.28	3.88	4.98	2.36	4.25	6.19
			Vanadium (µg)	2.06	73.0	131	176	77.9	147	224
	Age = 30.1*		Nickel (µg)	2.05	154	265	348	194	361	601
	$[P_{2.5}; P_{97.5}] = [20.7; 40.3]$		Inorganic arsenic (µg)	1.95	30.6	54.9	74.1	31.9	60.8	95.8
			Antimony (µg)	1.87	1.92	3.61	5.23	2.13	4.20	6.94
			HCDF 123678 (pg)	1.86	4.29	7.76	9.95	5.16	9.74	15.9
			Pyrene (ng)	1.85	414	764	1003	482	919	1469
		TE-F-PAH (34 %)	Aluminium (µg)	1.83	3012	5655	8073	3591	7017	12051
	BMI = 23.6*		HCDF 123478 (pg)	1.80	7.32	13.0	16.9	8.79	16.4	28.2
	$[P_{2.5}; P_{97.5}] = [17.6; 36.3]$		HCDF 1234789 (pg)	1.80	5.23	9.33	11.7	6.22	11.6	19.6
			OCDF (pg)	1.77	25.6	45.7	57.3	30.5	57.3	94.6
			Lithium (µg)	1.74	48.8	91.3	116	53.0	103	145
			HCDF 123789 (pg)	1.73	1.06	1.96	2.48	1.30	2.50	4.11
			HCDD 123678 (pg)	1.71	5.94	10.9	14.2	7.22	14.0	24.0
	WeightGain = 13.2		HCDD 123789 (pg)	1.70	2.07	3.87	4.97	2.50	4.87	7.81
	$[P_{2.5}; P_{97.5}] = [1.0; 25.0]$		HCDF 1234678 (pg)	1.69	9.31	16.8	21.9	11.2	21.1	37.3
			HCDF 234678 (pg)	1.68	2.02	3.56	4.42	2.46	4.57	7.91
			HCDD 123478 (pg)	1.65	1.58	2.94	3.86	1.90	3.74	6.35
			Cadmium (µg)	1.65	10.9	18.7	24.6	13.4	25.0	44.8
			Zearalenone (ng)	1.60	438	838	1197	489	983	1619
			PCB 123 (pg)	3.14	604	1423	1990	960	2580	4575
			PBDE 47 (ng)	3.04	8.31	20.3	29.1	13.6	37.4	68.9
		PCB-BFR-Aso-MeHg (11.4 %)	PBDE 100 (ng)	3.01	1.68	4.34	6.30	2.82	8.07	14.7
			PCB 52 (pg)	3.01	9784	22288	31568	15425	39948	79671
			PCB 101 (pg)	3.00	14632	38057	56236	24856	71555	136572

Individuals Group		Mixtures			Cluster exposure			Population exposure		
Cluster	**Description**	**Major Mixtures**	**Major substances (unit)**	**%Subst**	**Mean**	**P95**	**P99**	**Mean**	**P95**	**P99**
			PBDE 154 (ng)	3.00	0.88	2.28	3.31	1.48	4.21	7.76
			PCB 77 (pg)	2.97	247	543	736	385	962	1974
			PCB 153 (pg)	2.93	56917	122694	158868	85576	208026	380119
			PBB 153 (ng)	2.88	0.04	0.10	0.15	0.06	0.19	0.38
			PCB 105 (pg)	2.88	6488	13891	18032	9718	23474	43399
			PBDE 28 (ng)	2.86	0.48	1.35	2.06	0.85	2.55	4.79
			PBB 52 (ng)	2.83	0.05	0.16	0.25	0.10	0.31	0.59
			PCB 157 (pg)	2.82	664	1421	1810	992	2378	4334
			TCDF 2378 (pg)	2.77	9.56	23.3	33.3	16.0	43.2	91.5
			PCB 167 (pg)	2.74	1626	3444	4438	2409	5747	10425
			PBB 101 (ng)	2.71	0.03	0.09	0.14	0.05	0.16	0.32
			PCB 114 (pg)	2.62	502	1061	1350	741	1765	3282
			PCDF 12378 (pg)	2.55	1.96	3.95	5.18	2.91	6.77	13.9
			PCB 138 (pg)	2.54	40495	84856	106790	59247	139002	251234
			PCB 189 (pg)	2.49	282	597	755	417	983	1828
			Organic arsenic (µg)	2.39	33.5	83.1	111	54.6	142	316
			Methyl mercury (µg)	2.34	1.05	3.41	4.70	1.82	5.45	9.99
			PCB 118 (pg)	2.31	22533	45994	57252	32575	74781	137130
			PCB 180 (pg)	2.19	20426	42775	53276	29689	69185	124568
			PCB 156 (pg)	2.11	2926	6038	7461	4225	9708	17959
		Pest-1 (9.9 %)	Propargite (µg)	5.86	3.96	9.75	14.6	9.25	34.8	62.9
			Thiabendazole (µg)	5.77	5.80	15.4	21.3	11.9	45.6	78.3
			Pirimicarb (µg)	5.72	0.01	0.03	0.03	0.03	0.10	0.20
			Phosalone (µg)	5.71	0.07	0.18	0.18	0.18	0.65	1.29
			Diphenylamine (µg)	5.70	1.54	3.86	3.86	3.83	13.8	27.5
			Diflubenzuron (µg)	5.70	0.001	0.001	0.001	0.001	0.004	0.008
			Folpet (µg)	5.70	0.02	0.04	0.04	0.04	0.15	0.30
			Azinphos-methyl (µg)	5.70	0.06	0.14	0.14	0.14	0.50	0.99
			Captan (µg)	5.70	0.03	0.06	0.06	0.06	0.23	0.45
			Ethoxyquin (µg)	5.70	0.63	1.59	1.59	1.58	5.67	11.3
			Triflumuron (µg)	5.70	0.01	0.04	0.04	0.04	0.13	0.25
			Tebufenozide (µg)	5.70	0.002	0.01	0.01	0.005	0.02	0.04
			Carbendazim (µg)	5.48	0.28	0.69	0.85	1.02	3.10	4.75
			Phosmet (µg)	3.31	0.14	0.34	1.52	0.28	1.20	1.96
			Patulin (ng)	2.77	33.0	105	210	40.9	111	215
			Fludioxonil (µg)	2.29	1.93	5.11	10.3	3.58	12.2	22.4
			Chlorpyrifos-ethyl (µg)	2.07	0.59	1.55	6.27	1.13	4.92	8.94
			Teflubenzuron (µg)	1.64	0.04	0.18	0.19	0.10	0.36	1.26
			Boscalid (µg)	1.55	0.71	1.96	2.84	1.38	4.73	10.9
			Methomyl (µg)	1.18	0.002	0.01	0.01	0.005	0.02	0.07
			Bifenthrin (µg)	1.07	0.09	0.24	0.39	0.18	0.65	1.23
			Pyrimethanil (µg)	0.99	0.78	2.17	2.81	1.80	6.66	13.7
			Phenylphenol (µg)	0.90	3.46	9.96	19.4	3.84	10.5	20.2
			Lambda-cyhalothrin (µg)	0.82	0.22	0.70	1.28	0.37	1.32	2.15
			Cyprodinyl (µg)	0.80	1.16	3.29	4.97	2.16	7.68	15.2
			Chromium VI (µg)	2.31	**69.4**	129	166	55.5	98.9	150
			Chromium III (µg)	2.19	**474**	966	1356	338	626	1041

Individuals Group		Mixtures			Cluster exposure			Population exposure		
Cluster	**Description**	**Major Mixtures**	**Major substances (unit)**	**%Subst**	**Mean**	**P95**	**P99**	**Mean**	**P95**	**P99**
4	N = 1,039	TE-F-PAH (37.8 %)	**Lead (µg)**	2.17	**20.7**	37.7	48.8	15.7	27.9	44.7
			Cobalt (µg)	2.17	**21.4**	41.6	58.6	15.8	28.8	46.2
			Barium (µg)	2.08	**858**	1686	2310	653	1203	1975
			Tellurium (µg)	2.08	**3.12**	5.59	7.30	2.36	4.25	6.19
			Vanadium (µg)	2.06	**103**	195	276	77.9	147	224
	Age = 31.0*		**Nickel (µg)**	2.05	**249**	496	699	194	361	601
	[$P_{2.5}$; $P_{97.5}$] = [21.1; 41.1]		**Inorganic arsenic (µg)**	1.95	**43.9**	82.9	116	31.9	60.8	95.8
			Antimony (µg)	1.87	**2.90**	6.08	8.37	2.13	4.20	6.94
			HCDF 123678 (pg)	1.86	**8.17**	15.6	23.5	5.16	9.74	15.9
			Pyrene (ng)	1.85	**737**	1453	2159	482	919	1469
			Aluminium (µg)	1.83	**4993**	10226	13818	3591	7017	12051
	BMI = 23.7		**HCDF 123478 (pg)**	1.80	**14.2**	28.2	41.0	8.79	16.4	28.2
	[$P_{2.5}$; $P_{97.5}$] = [21.1; 41.1]		**HCDF 1234789 (pg)**	1.80	**9.99**	18.9	28.0	6.22	11.6	19.6
			OCDF (pg)	1.77	**49.6**	91.9	137	30.5	57.3	94.6
			Lithium (µg)	1.74	**67.2**	126	153	53.0	103	145
			HCDF 123789 (pg)	1.73	**2.16**	3.96	5.80	1.30	2.50	4.11
			HCDD 123678 (pg)	1.71	**11.8**	23.2	33.7	7.22	14.0	24.0
	WeightGain = 13.3		**HCDD 123789 (pg)**	1.70	**3.96**	7.61	11.5	2.50	4.87	7.81
	[$P_{2.5}$; $P_{97.5}$] = [2.0; 26.4]		**HCDF 1234678 (pg)**	1.69	**18.3**	36.8	56.7	11.2	21.1	37.3
			HCDF 234678 (pg)	1.68	**4.10**	7.56	10.9	2.46	4.57	7.91
			HCDD 123478 (pg)	1.65	**3.03**	6.10	9.07	1.90	3.74	6.35
			Cadmium (µg)	1.65	**19.8**	41.0	59.7	13.4	25.0	44.8
			Zearalenone (ng)	1.60	**676**	1495	2183	489	983	1619
		PCB-BFR-Aso-MeHg (24 %)	**PCB 123 (pg)**	3.14	**2494**	5559	8654	960	2580	4575
			PBDE 47 (ng)	3.04	**36.2**	85.0	128	13.6	37.4	68.9
			PBDE 100 (ng)	3.01	**7.77**	18.4	28.6	2.82	8.07	14.7
			PCB 52 (pg)	3.01	**40056**	89801	131602	15425	39948	79671
			PCB 101 (pg)	3.00	**69910**	168748	243983	24856	71555	136572
			PBDE 154 (ng)	3.00	**4.06**	9.57	14.93	1.48	4.21	7.76
			PCB 77 (pg)	2.97	**989**	2129	3178	385	962	1974
			PCB 153 (pg)	2.93	**206682**	453129	642501	85576	208026	380119
			PBB 153 (ng)	2.88	**0.18**	0.43	0.63	0.06	0.19	0.38
			PCB 105 (pg)	2.88	**23300**	50623	72770	9718	23474	43399
			PBDE 28 (ng)	2.86	**2.45**	6.15	9.13	0.85	2.55	4.79
			PBB 52 (ng)	2.83	**0.30**	0.75	1.15	0.10	0.31	0.59
			PCB 157 (pg)	2.82	**2369**	5141	7359	992	2378	4334
			TCDF 2378 (pg)	2.77	**44.5**	99.7	153	16.0	43.2	91.5
			PCB 167 (pg)	2.74	**5703**	12236	17889	2409	5747	10425
			PBB 101 (ng)	2.71	**0.16**	0.40	0.62	0.05	0.16	0.32
			PCB 114 (pg)	2.62	**1746**	3763	5356	741	1765	3282
			PCDF 12378 (pg)	2.55	**6.99**	14.4	22.3	2.91	6.77	13.9
			PCB 138 (pg)	2.54	**137967**	293810	425007	59247	139002	251234
			PCB 189 (pg)	2.49	**975**	2094	3025	417	983	1828
			Organic arsenic (µg)	2.39	**149**	332	529	54.6	142	316
			Methyl mercury (µg)	2.34	**4.88**	10.9	18.2	1.82	5.45	9.99
			PCB 118 (pg)	2.31	**74621**	157596	220450	32575	74781	137130
			PCB 180 (pg)	2.19	**67873**	143174	207719	29689	69185	124568
			PCB 156 (pg)	2.11	**9587**	20271	28908	4225	9708	17959

Individuals Group		Mixtures			Cluster exposure			Population exposure		
Cluster	**Description**	**Major Mixtures**	**Major substances (unit)**	**%Subst**	**Mean**	**P95**	**P99**	**Mean**	**P95**	**P99**
		Mixt-3 (9.7 %)	**PFTeDA (ng)**	6.45	**0.03**	0.17	0.24	0.008	0.03	0.12
			Benzo[k]fluoranthene (ng)	6.43	**16.0**	44.74	64.89	7.67	16.2	38.4
			Benzo[b]fluoranthene (ng)	6.39	**41.7**	114	171	20.1	42.1	98.6
			Benzo[j]fluoranthene (ng)	6.23	**20.4**	53.38	80.12	9.92	21.4	49.0
			Dibenzo[a,i]pyrene (ng)	6.06	**0.19**	0.61	0.88	0.10	0.26	0.53
			PFDoA (ng)	5.40	**0.25**	1.09	1.54	0.07	0.21	0.83
			Chrysene (ng)	5.12	**84.0**	206	285	42.3	89.8	192
			Benzo[a]anthracene (ng)	4.21	**34.3**	81.58	114	18.2	40.4	77.0
			Dibenzo[a,e]pyrene (ng)	4.08	**1.47**	3.64	5.40	0.80	1.79	3.51
			Benzo[a]pyrene (ng)	3.90	**17.2**	37.87	50.81	9.58	20.1	37.0
			Silver (µg)	3.54	**167**	370	496	108	209	370
			Indeno[1,2,3-cd]pyrene (ng)	3.22	**19.0**	41.03	54.89	11.4	23.2	39.5
			Dibenzo[a,h]anthracene (ng)	2.83	**5.06**	11.18	15.20	3.14	6.56	11.4
			HBCD beta (ng)	2.33	**0.73**	1.55	2.33	0.37	0.80	1.53
			PFDA (ng)	2.33	**1.52**	5.22	17.52	0.66	2.75	8.70
			Dibenzo[a,i]pyrene (ng)	2.22	**0.32**	1.26	2.49	0.20	0.72	1.39
			PFOS (ng)	2.09	**9.90**	23.8	39.3	4.27	10.4	22.9
			Benzo[c]fluorene (ng)	1.96	**5.46**	13.5	19.5	2.60	6.39	12.8
			Organic arsenic (µg)	1.80	**149**	332	529	55	142	316
			PFNA (ng)	1.70	**0.41**	1.45	3.44	0.22	0.74	2.10
			Benzo[g,h,i]perylene (ng)	1.69	**38.6**	77.2	99.8	25.5	50.8	81.9
			Cadmium (µg)	1.59	**19.8**	41.0	59.7	13.4	25.0	44.8
			Fluoranthene (ng)	1.50	**326**	688	1011	190	377	683
			HBCD gamma (ng)	1.08	**1.31**	2.57	4.25	0.61	1.34	2.62
			OCDD (pg)	1.03	**182**	401	659	105	218	418
5		Pest-3 (21.4 %)	**Myclobutanil (µg)**	5.26	**1.44**	3.06	3.17	0.20	0.70	2.51
			Fenhexamid (µg)	5.20	**23.5**	51.0	52.1	3.17	14.6	39.0
			Triadimenol (µg)	5.05	**1.44**	2.99	3.17	0.20	0.56	2.94
	N = 1,080		**Etofenprox (µg)**	4.99	**1.00**	2.10	2.10	0.13	0.29	2.10
			Trifloxystrobin (µg)	4.99	**0.07**	0.14	0.14	0.009	0.02	0.14
			Quinoxyfen (µg)	4.99	**0.62**	1.31	1.31	0.08	0.18	1.31
			Spiroxamine (µg)	4.99	**0.09**	0.19	0.19	0.01	0.03	0.19
			Cyfluthrin (µg)	4.99	**0.06**	0.12	0.12	0.007	0.02	0.12
	Age = 31.4*		**Tetraconazole (µg)**	4.99	**0.06**	0.12	0.12	0.007	0.02	0.12
	$[P_{2.5}; P_{97.5}] = [20.9; 41.6]$		**Tebufenpyrad (µg)**	4.99	**0.02**	0.05	0.05	0.003	0.007	0.05
			Methomyl (µg)	4.50	**0.04**	0.08	0.08	0.005	0.018	0.07
			Teflubenzuron (µg)	4.27	**0.62**	1.36	1.36	0.10	0.36	1.26
			Pyrimethanil (µg)	4.13	**7.90**	15.7	19.5	1.80	6.66	13.7
			Cyprodinyl (µg)	3.89	**9.71**	19.3	25.1	2.16	7.68	15.2
	BMI = 23.6		**Bifenthrin (µg)**	3.79	**0.73**	1.48	1.98	0.18	0.65	1.23
	$[P_{2.5}; P_{97.5}] = [20.9; 41.6]$		**Mepanipyrim (µg)**	3.68	**1.11**	2.51	2.62	0.16	0.71	1.80
			Metalaxyl (µg)	3.50	**0.90**	1.79	2.22	0.21	0.77	1.43
			Penconazole (µg)	2.98	**0.54**	1.22	1.29	0.08	0.40	0.86
			Boscalid (µg)	2.35	**6.71**	14.5	17.1	1.38	4.73	10.9
			Iprodione (µg)	1.87	**31.1**	69.1	115	10.5	36.3	66.1
	WeightGain = 13.0		**Chlorpyrifos-ethyl (µg)**	1.81	**5.14**	10.6	11.1	1.13	4.92	8.94
	$[P_{2.5}; P_{97.5}] = [2.0; 25.0]$		**Lambda-cyhalothrin (µg)**	1.32	**1.15**	2.56	4.46	0.37	1.32	2.15

Individuals Group		Mixtures			Cluster exposure			Population exposure		
Cluster	Description	Major Mixtures	Major substances (unit)	%Subst	Mean	P95	P99	Mean	P95	P99
			Procymidone (µg)	1.26	**5.31**	11.6	16.9	1.71	5.82	9.87
			Azoxystrobin (µg)	1.18	**1.32**	2.94	4.18	0.27	0.81	2.71
			Chlorpyrifos-methyl (µg)	1.17	**0.72**	1.46	2.71	0.38	0.80	1.19
		TE-F-PAH (19.9 %)	Chromium VI (µg)	2.31	**83.6**	184	322	55.5	98.9	150
			Chromium III (µg)	2.19	**578**	1528	2767	338	626	1041
			Lead (µg)	2.17	**26.0**	60.8	109	15.7	27.9	44.7
			Cobalt (µg)	2.17	**27.4**	70.7	126	15.8	28.8	46.2
			Barium (µg)	2.08	**1104**	2673	4704	653	1203	1975
			Tellurium (µg)	2.08	**3.36**	7.94	11.9	2.36	4.25	6.19
			Vanadium (µg)	2.06	**121**	282	466	77.9	147	224
			Nickel (µg)	2.05	**364**	865	1467	194	361	601
			Inorganic arsenic (µg)	1.95	**50.7**	120	210	31.9	60.8	95.8
			Antimony (µg)	1.87	**3.45**	9.04	15.7	2.13	4.20	6.94
			HCDF 123678 (pg)	1.86	**8.38**	24.5	54.4	5.16	9.74	15.9
			Pyrene (ng)	1.85	**765**	2354	5244	482	919	1469
			Aluminium (µg)	1.83	**6255**	16044	31045	3591	7017	12051
			HCDF 123478 (pg)	1.80	**14.7**	43.8	107	8.79	16.4	28.2
			HCDF 1234789 (pg)	1.80	**10.3**	31.9	75.0	6.22	11.6	19.6
			OCDF (pg)	1.77	**51.1**	159	405	30.5	57.3	94.6
			Lithium (µg)	1.74	**80.2**	169	257	53.0	103	145
			HCDF 123789 (pg)	1.73	**2.18**	6.63	16.8	1.30	2.50	4.11
			HCDD 123678 (pg)	1.71	**12.1**	37.1	88.2	7.22	14.0	24.0
			HCDD 123789 (pg)	1.70	**4.09**	12.3	29.8	2.50	4.87	7.81
			HCDF 1234678 (pg)	1.69	**19.2**	57.9	147	11.2	21.1	37.3
			HCDF 234678 (pg)	1.68	**4.21**	12.7	30.9	2.46	4.57	7.91
			HCDD 123478 (pg)	1.65	**3.15**	9.12	23.0	1.90	3.74	6.35
			Cadmium (µg)	1.65	**24.7**	68.2	148	13.4	25.0	44.8
			Zearalenone (ng)	1.60	**727**	2029	4184	489	983	1619
		Pest-2 (18.6 %)	Bupirimate (µg)	7.90	**2.75**	6.63	7.64	0.44	1.20	6.52
			Dichlorvos (µg)	7.87	**0.11**	0.25	0.29	0.02	0.05	0.25
			Endosulfan (µg)	7.84	**1.67**	4.05	4.49	0.27	0.73	3.90
			Acrinathrin (µg)	7.83	**0.83**	1.89	2.31	0.14	0.57	1.86
			Kresoxim-methyl (µg)	7.65	**0.22**	0.53	0.59	0.03	0.10	0.51
			Azoxystrobin (µg)	7.05	**1.32**	2.94	4.18	0.27	0.81	2.71
			Penconazole (µg)	5.20	**0.54**	1.22	1.29	0.08	0.40	0.86
			Boscalid (µg)	4.66	**6.71**	14.5	17.1	1.38	4.73	10.9
			Tebuconazole (µg)	4.51	**0.29**	0.57	0.58	0.06	0.29	0.56
			Dimethoate (µg)	4.44	**3.71**	7.21	7.29	0.77	3.61	7.19
			Fenbuconazole (µg)	4.42	**0.70**	1.36	1.36	0.14	0.68	1.36
			Mepanipyrim (µg)	4.11	**1.11**	2.51	2.62	0.16	0.71	1.80
			Procymidone (µg)	3.79	**5.31**	11.6	16.9	1.71	5.82	9.87
			Phosmet (µg)	2.85	**1.06**	2.39	2.39	0.28	1.20	1.96
			Chlorpyrifos-ethyl (µg)	2.60	**5.14**	10.6	11.1	1.13	4.92	8.94
			Metalaxyl (µg)	2.53	**0.90**	1.79	2.22	0.21	0.77	1.43
			Lambda-cyhalothrin (µg)	2.36	**1.15**	2.56	4.46	0.37	1.32	2.15
			Cyprodinyl (µg)	2.00	**9.71**	19.3	25.1	2.16	7.68	15.2
			Fludioxonil (µg)	1.53	**10.0**	24.4	37.2	3.58	12.2	22.4

Individuals Group		Mixtures			Cluster exposure			Population exposure		
Cluster	**Description**	**Major Mixtures**	**Major substances (unit)**	**%Subst**	**Mean**	**P95**	**P99**	**Mean**	**P95**	**P99**
			Iprodione (µg)	1.17	**31.1**	69.1	115	10.5	36.3	66.1
			Fenhexamid (µg)	0.64	**23.5**	51.0	52.1	3.17	14.6	39.0
			Sulphites (µg)	0.59	**6025**	29117	59968	2542	14526	29022
			Nickel (µg)	0.59	**364**	865	1467	194	361	601
			Chlorthal-dimethyl (µg)	0.46	**0.02**	0.07	0.15	0.01	0.04	0.07
			Pyriproxyfen (µg)	0.40	**0.06**	0.18	0.27	0.03	0.14	0.26
6		Mixt-4 (38.2 %)	**Chlorothalonil (µg)**	10.9	**0.57**	1.23	1.41	0.14	0.61	1.23
			Diethofencarb (µg)	10.9	**1.92**	4.12	4.74	0.46	2.06	4.12
			Carbofuran (µg)	10.9	**0.13**	0.29	0.33	0.03	0.14	0.29
	N = 1,314		**Tetradifon (µg)**	10.9	**0.10**	0.21	0.24	0.02	0.11	0.21
			Pyriproxyfen (µg)	10.8	**0.13**	0.27	0.31	0.03	0.14	0.26
			Chlorthal-dimethyl (µg)	8.86	**0.04**	0.07	0.09	0.01	0.04	0.07
			Procymidone (µg)	6.74	**4.38**	8.61	11.9	1.71	5.82	9.87
			Pyrimethanil (µg)	4.58	**4.06**	7.82	9.19	1.80	6.66	13.7
	Age = 31.8*		**Metalaxyl (µg)**	4.17	**0.46**	0.88	1.08	0.21	0.77	1.43
	$[P_{2.5}; P_{97.5}]$ = [25.5; 40.8]		**Iprodione (µg)**	3.26	**22.6**	58.0	86.3	10.5	36.3	66.1
			Bifenthrin (µg)	2.47	**0.32**	0.76	1.15	0.18	0.65	1.23
			Sulphur (µg)	2.37	**12.4**	37.7	56.6	5.48	19.5	38.1
			Cyproconazole (µg)	1.95	**0.04**	0.12	0.18	0.01	0.06	0.12
	BMI = 23.1*		**Tin (µg)**	1.29	**190**	462	930	136	389	919
	$[P_{2.5}; P_{97.5}]$ = [17.6; 35.2]		**Cyprodinyl (µg)**	1.28	**3.26**	7.64	10.2	2.16	7.68	15.2
			Lambda-cyhalothrin (µg)	1.12	**0.61**	1.70	2.51	0.37	1.32	2.15
			Fludioxonil (µg)	1.07	**6.18**	19.1	28.9	3.58	12.2	22.4
			Secoisolariciresino (µg)	0.69	**22171**	51375	180293	16677	41757	177636
	WeightGain = 13.2		**Flutriafol (µg)**	0.66	**0.49**	1.62	3.23	0.27	0.71	2.54
	$[P_{2.5}; P_{97.5}]$ = [3.0; 24.0]		**Imidacloprid (µg)**	0.66	**0.07**	0.24	0.48	0.04	0.11	0.38
			Boscalid (µg)	0.48	**1.84**	3.55	4.78	1.38	4.73	10.91
			Tebuconazole (µg)	0.39	**0.09**	0.30	0.57	0.06	0.29	0.56
			Lithium (µg)	0.36	**55.7**	96.9	126	53.0	103	145
			Chlorfenvinphos (µg)	0.32	**0.12**	0.50	0.98	0.08	0.34	0.68
			Lindane (µg)	0.27	**0.05**	0.09	0.27	0.04	0.09	0.27
		TE-F-PAH 1 (24.4 %)	See cluster 4		No higher significant difference					
		PFAA-Ge-Li 5 (13.3 %)	See cluster 1		No higher significant difference					

a Population characteristics: age = 30.7, $[P_{2.5}; P_{97.5}]$ = [21.1; 40.7], BMI = 23.5, $[P_{2.5}; P_{97.5}]$ = [17.6; 36.3], WeightGain = 13.1, $[P_{2.5}; P_{97.5}]$ = [2.0; 25.0]

4 Discussion

The application of the SNMU and the hierarchical clustering method to the exposure datasets from the EDEN and ELFE cohorts and TDS 2 enabled us to identify the eight main chemical mixtures to which pregnant women in France are exposed before and during pregnancy. The method also made it possible to define groups of women with specific co-exposure profiles. Thus, six groups of pregnant women associated with the eight mixtures were identified. Except for specific compounds, similar mixtures were obtained concerning the periods before and during pregnancy using the EDEN survey. Moreover, these similar mixtures were also found from the ELFE cohort exposure data during pregnancy.

The SNMU and the choice of number of mixtures K

The choice of an optimal number of exposure systems K remains a major challenge. There is no theoretical result for determining this number in general, but different approaches have been proposed. Thus, in Bayesian NMF, Cemgil (2009) and Schmidt et al. (2009) evaluated a marginal likelihood for each candidate value of K and the value with highest marginal likelihood is considered as the optimal choice. Mørup and Hansen (2009), Yang et al. (2010) and Tan and Févotte (2013) also proposed methods in which the number of exposure systems is set to a large value and non-pertinent exposure systems are driven to zero during simulation. With the SNMU method, this problem is solved in part. This is because the recursive algorithm used allows us to identify exposure systems one by one. From the original exposure matrix $E = R_1$, the first approximation $W_{:1}H_{1:}$ is extracted and therefore subtracted from this matrix, i.e. $R_2 = R_1 - W_{:1}H_{1:}$. The same procedure is thus applied on the new obtained matrix R_2. In this way, another approximation is extracted corresponding to the first approximation for the matrix R_2 and to the second approximation for the original exposure matrix E. At each step k, an approximation $W_{:k}H_{k:}$ is extracted from a new residual matrix R_k. It corresponds to the k^{st} approximation of the original exposure matrix E and is identical, regardless of the number of exposure systems. Hence, the first approximations extracted, regardless

of the number of exposure systems, are identical. When applying the clustering method to the matrix H obtained with this procedure, the K values for which one or more exposure systems were not used to characterize a cluster were rejected. Therefore, the optimal number of K was defined as the largest of the non-rejected values. In this study, eight exposure systems were extracted on each exposure dataset and led to division of the overall study population into six clusters characterised by these exposure systems.

The sparsity coefficient and its bounds

As described by Gillis and Plemmons (2013), the SNMU requires human supervision to determine the sparsity coefficient and its bounds. No mathematical criterion to automatically determine these values is available. The values depend only on the degree of sparsity needed by the user to reduce the dataset. In our case study, we decided to set the bounds to 0.2 and 0.99 to make λ vary on a large range of values and we observed that results remained stable. We also observed that when k increased, the additional exposure systems were more sparse and the first exposure system was always less sparse. We set the coefficient lambda to high sparsity for the first system to force it to be sparse, whereas the values for the other k were set lower.

Spatial and temporal mixture comparison

The EDEN and ELFE cohorts were selected to identify main mixtures to which pregnant women in France were exposed during pregnancy. These two cohorts are the most recent and complete surveys available for this population. EDEN was the first survey conducted in a local area, whereas ELFE is a national survey including a large number of women. The FFQ used in the ELFE study was a modified version from the EDEN study. Some differences were observed, in particular on the number of food items: 137 in EDEN and 125 in ELFE. Some food items were considered individually in EDEN and grouped together in ELFE. For example, in the EDEN study, food items like “red and rosé wine”, “white wine”, “beer”, “cider”, “lemonade”, “light soda” and “light cola” were grouped in the ELFE study respectively in food items “white, red or rosé wine”, “cider and beer” and “light lemonade and soda”. Likewise, “fresh cream”, “butter” and “margarine”, light and non-light, were considered

individually in the EDEN study, whereas in the ELFE study, consumption of “fresh cream”, “butter” and “margarine” were recorded. Despite the differences in survey periods (2003-2006 for EDEN vs from 2011 for ELFE), regions (2 regions vs mainland France) and in the FFQs, comparable mixtures were obtained.

Similar results were also obtained on clustering from EDEN before and during pregnancy exposure data. The study population was divided into six clusters for each exposure data. Individuals in five clusters among the six were significantly more exposed to substances from mixtures characterising the clusters.

Comparison of mixture before and during pregnancy

Seven of the eight mixtures identified before and during pregnancy from EDEN exposures presented similarities and six of them were also observed in ELFE during pregnancy. In EDEN before pregnancy, “Mixt-1” mixture consisting of certain heat-induced contaminants (acrylamide and six PAHS), some mycotoxins, pesticides, phytoestrogens and additives was not identified from EDEN during pregnancy. This can be explained by the fact that the common foods contributing to exposure to these substances were coffee, cereal products (bread products, pasta, aperitif biscuits, semolina, etc.), potato products, meat and delicatessen meats, crustaceans and molluscs, fruit juices, milk and beans (Bemrah et al., 2012; Nougadère et al., 2012; Sirot et al., 2012a; Chan-Hon-Tong et al., 2013; Sirot et al., 2013; Veyrand et al., 2013). Consumption of coffee, meat and delicatessen meats, crustaceans and molluscs were reduced during pregnancy by the EDEN pregnant women population (Chan-Hon-Tong et al., 2013). This consumption reduction was also observed in pregnant women in Portugal (Pinto et al., 2009) for red meat, coffee and tea. In EDEN during pregnancy, “Mixt-2” mixture consisting mainly of DL-PCB, NDL-PCB, BFR and PAH compounds was not identified from EDEN before pregnancy. This result can be explained by the fact that the main food contributors to the exposure to these substances were fish products, butter, cheese, crustaceans and molluscs, cereal products, milk, meat and delicatessen meats (Sirot et al., 2012b; Chan-Hon-Tong et al., 2013; Veyrand et al., 2013; Rivière et al., 2014), even though consumption of these last food groups was

reduced during pregnancy. During pregnancy, women increased their consumption of fish, butter, milk, cheese and cereal products (Chan-Hon-Tong et al., 2013). This trend concerning the consumption of fish, milk and cereal products (especially bread) was also observed in pregnant women in Portugal (Pinto et al., 2009).

Comparison with the general adult population

The NMF method was also used by Traoré et al. (2016) to identify the main mixtures to which the general adult population in France is exposed through diet. In this study, dietary patterns were first defined using the LS-NMF method (Wang et al., 2006), a modified version of NMF (Lee and Seung, 2001) and solved by a multiplicative algorithm based on gradient descent, on individual consumption data, connected with mixtures and associated with individual clusters. In the present study, the SNMU method, solved by a recursive algorithm based on Lagrangian relaxation, was applied to exposure data concerning pregnant women in France to define mixtures. Despite the above differences, two mixtures identified from the three exposure datasets and one from ELFE during pregnancy in the present study presented some similarities with three mixtures identified in the general adult population in France by Traoré et al. (2016). These mixtures consisting of trace elements, PAHs and pesticides were associated with the “Mediterranean” diet which is mostly eaten by women. This comparison consolidates the method’s robustness since the consumptions were recorded using FFQ in the EDEN and ELFE studies, whereas in INCA 2, a seven-day record was used. The mixtures associated with the “Traditional”, the “Snacking” and the “Simplicity” diets in Traoré et al. (2016) were not found in the pregnant women cohorts, which could be explained by a reduction of the consumption of the associated foods during pregnancy.

Uncertainty Sources

Several sources of uncertainty may affect these results.

An FFQ record was considered for each woman. In EDEN, the first FFQ was collected at inclusion and concerned the diet in the year before pregnancy, and the second FFQ at delivery and concerned the diet during the third trimester of pregnancy. In ELFE, the FFQ was recorded at delivery and

described the dietary habits during the third trimester of pregnancy. Although the FFQ approach made it possible to collect information on a large number of individuals at lower cost, the consumed quantity is generally overestimated compared with other methods of recording (Rothenberg, 2009), especially when the number of foods recorded is high (Kim and Holowaty, 2003). This may have affected the exposures, which are probably over-estimated, and ultimately, the mixture compositions. Another issue concerning the populations in these cohorts is under-reporting by individuals. As recommended by the European Food Safety Authority (EFSA, 2014), under-reporting individuals, i.e. individuals who under-reported their consumption frequencies, were included in this study. This means that no exclusion on daily energy intake was applied (i.e. individuals with daily energy intake lower than the 3rd percentile or higher than the 97th percentile) and that may have affected the identification of the mixtures and the clusters with similar exposure profiles.

Another potential source of uncertainty is related to the concentration data collected using the TDS approach. As in any survey, TDS provided a snapshot of the general contamination of the population diet at a specific point in time. However, concentration data change over time, especially for pesticides and additives which constant changing regulations, but also for trace elements as observed for arsenic and lead when comparing TDS 1 (Leblanc et al., 2005) and TDS 2 (Arnich et al., 2012) in France. For example, due to the prohibition on lead in automotive gasolines, but also the reduction of lead presence in water pipes and interior paints, the exposure level is decreasing in industrialised countries (Etchevers et al., 2014) such as France. Climate change or different weather conditions could also have an effect on concentrations of mycotoxins as observed by Cano-Sancho et al. (2012) when compared European products with ethnical food such as Mexican corn-based foods. Therefore, regarding concentration data, the present mixtures come from observations over the period 2007-2009 (the TDS 2 sampling period) and should be updated with more recent data. Moreover, even though TDS 2 covered a large range of substances, some substances that may contribute to real-life mixtures were not analysed and the mixture could be incomplete. For example, substances such as phthalates, alkyl phenols and furane were not analysed in TDS 2 whereas they

were analysed during the French infant TDS (Hulin et al., 2014). However, the TDS protocol is based on previous experience of food contamination and should cover the major substances present in food. Also, some foods like exotic fruits, which were recorded in the EDEN and ELFE cohorts were not analysed in TDS 2 as they are not largely consumed by the general population, and this could lead to a slight underestimation of exposure to the substances present in these products. Dilution due to the pooling of samples is one of the issues in TDS. Actual values in one sample may be diluted below the LOD in composite samples and lead to censored data. The use of the lower bound (LB) scenario, i.e. undetected values were set to 0 and detected but unquantified values to the limit of detection, could lead to the exclusion of substances that could be present in a food item and to underestimation of exposure. Then lower analytical limits would be useful to precise the exposure to chemicals with a high censorship rate.

The mixtures were identified using chronic exposure assessment, as TDS data could not be used for acute exposure assessment. However, for some substances, the time frame related to acute exposure may be more appropriate for certain toxicological effects. As a result, it would be interesting to apply this approach to acute exposure data estimated with national monitoring programmes as was done for the general population in France (Crépet et al., 2013).

Moreover, it would be interesting to investigate other sources of exposure such as air or dust. For instance, it is possible that the occupational environment of pregnant women contributed to exposure. Finally, as this approach was based on external exposure, it may be useful to apply this method to internal exposure data for a large range of substances.

Conclusion

This study allowed us to point out the diversity of mixtures to which pregnant women in France are exposed before and during pregnancy through diet. It also showed new avenues for research and where efforts are to be made to prioritise combinations of substances for which it is essential to investigate possible combined toxicological effects. These mixtures can be used in epidemiology to study the association with health effects in children.

Funding sources

This research was supported by the French *Agence Nationale de la Recherche* (ANR) (Grant ANR-14-CE21-0002-02).

Acknowledgments

- The EDEN cohort is supported by grants from FRM, Inserm, IReSP, Nestlé, the French Ministry of Health, ANR, Univ. Paris-Sud, InVS, ANSES and MGEN. The funding source had no involvement in study design, in the collection, analysis, and interpretation of data, in the writing of the report, or in the decision to submit the paper for publication. The study received approval from the Ethics Committee of Kremlin-Bicêtre University Hospital and from the French National Commission for Data Protection (CNIL). The EDEN mother-child cohort study group includes: I. Annesi-Maesano, J.Y. Bernard, J. Botton, M-A. Charles, P. Dargent-Molina, B. de Lauzon-Guillain, P. Ducimetière, M. de Agostini, B. Foliguet, A. Forhan, X. Fritel, A. Germa, V. Goua, R. Hankard, B. Heude, M. Kaminski, B. Larroque, N. Lelong, J. Lepeule, G. Magnin, L. Marchand, C. Nabet, F. Pierre, R. Slama, M.J. Saurel-Cubizolles, M. Schweitzer, O. Thiebaugeorges.
- ELFE is a study conducted jointly by the National Institute of Demographic Studies (Ined), the French National Institute for Health and Medical Research (Inserm), French blood establishment (EFS), French Institute for Public Health Surveillance (InVS), French National Institute for Statistics and Economic Studies (Insee), General Directorate for Health (DGS, Ministry of Health), General Directorate for Risk Prevention (DGPR, Ministry of Environment), Directorate for Research, Studies, Evaluation and Statistics (Drees), and the French National Family Allowance Fund (Cnaf). It benefits from additional funding from the Ministry of Research, Committee on SHS data (CCDSHS) and Ministry of Culture and

Communication (Deps). As part of the RECONAI platform, ELFE benefits from ANR funding (ANR-11-EQPX-0038).

References:

Arnich, N., Sirot, V., Rivière, G., Jean, J., Noël, L., Guérin, T., and Leblanc, J. C. (2012). Dietary exposure to trace elements and health risk assessment in the 2nd French Total Diet Study. [Article]. *Food and Chemical Toxicology, 50*(7), 2432-2449. doi: 10.1016/j.fct.2012.04.016

Béchaux, C., Zetlaoui, M., Tressou, J., Leblanc, J. C., Héraud, F., and Crépet, A. (2013). Identification of pesticide mixtures and connection between combined exposure and diet. *Food and Chemical Toxicology, 59*, 191-198. doi: 10.1016/j.fct.2013.06.006

Bemrah, N., Jean, J., Rivière, G., Sanaa, M., Leconte, S., Bachelot, M., Deceuninck, Y., Bizec, B. L., Dauchy, X., Roudot, A. C., Camel, V., Grob, K., Feidt, C., Picard-Hagen, N., Badot, P. M., Foures, F., and Leblanc, J. C. (2014). Assessment of dietary exposure to bisphenol A in the French population with a special focus on risk characterisation for pregnant French women. *Food and Chemical Toxicology, 72*, 90-97. doi: 10.1016/j.fct.2014.07.005

Bemrah, N., Vin, K., Sirot, V., Aguilar, F., Ladrat, A. C., Ducasse, C., Gey, J. L., Rétho, C., Nougadere, A., and Leblanc, J. C. (2012). Assessment of dietary exposure to annatto (E160b), nitrites (E249-250), sulphites (E220-228) and tartaric acid (E334) in the French population: The second French total diet study. *Food Additives and Contaminants - Part A Chemistry, Analysis, Control, Exposure and Risk Assessment, 29*(6), 875-885. doi: 10.1080/19440049.2012.658525

Canfield, R. L., Henderson, C. R., Cory-Slechta, D. A., Cox, C., Jusko, T. A., and Lanphear, B. P. (2003). Intellectual Impairment in Children with Blood Lead Concentrations below 10 µg per Deciliter. *The New England journal of medicine, 348*(16), 1517-1526. doi: 10.1056/NEJMoa022848

Cano-Sancho, G., Ramos, A. J., Marín, S., and Sanchis, V. (2012). Occurrence of fumonisins in Catalonia (Spain) and an exposure assessment of specific population groups. [Article]. *Food Additives and Contaminants - Part A Chemistry, Analysis, Control, Exposure and Risk Assessment, 29*(5), 799-808. doi: 10.1080/19440049.2011.644813

Cemgil, A. T. (2009). Bayesian inference for nonnegative matrix factorisation models. [Article]. *Computational Intelligence and Neuroscience, 2009*. doi: 10.1155/2009/785152

Chan-Hon-Tong, A., Charles, M. A., Forhan, A., Heude, B., and Sirot, V. (2013). Exposure to food contaminants during pregnancy. [Article]. *Science of the Total Environment, 458-460*, 27-35. doi: 10.1016/j.scitotenv.2013.03.100

Claus Henn, B., Coull, B. A., and Wright, R. O. (2014). Chemical mixtures and children's health. *Current Opinion in Pediatrics, 26*(2), 223-229. doi: 10.1097/mop.0000000000000067

Crépet, A., and Tressou, J. (2011). Bayesian nonparametric model for clustering individual co-exposure to pesticides found in the French diet. *Bayesian Analysis, 6*(1), 127-144. doi: 10.1214/11-BA604

Crépet, A., Tressou, J., Graillot, V., Béchaux, C., Pierlot, S., Héraud, F., and Leblanc, J. C. (2013). Identification of the main pesticide residue mixtures to which the French population is exposed. *Environmental Research, 126*, 125-133. doi: 10.1016/j.envres.2013.03.008

Czarnota, J., Gennings, C., Colt, J. S., De Roos, A. J., Cerhan, J. R., Severson, R. K., Hartge, P., Ward, M. H., and Wheeler, D. C. (2015). Analysis of Environmental Chemical Mixtures and Non-Hodgkin Lymphoma Risk in the NCI-SEER NHL Study. *Environmental Health Perspectives*. doi: :10.1289/ehp.1408630

Deschamps, V., de Lauzon-Guillain, B., Lafay, L., Borys, J. M., Charles, M. A., and Romon, M. (2009). Reproducibility and relative validity of a food-frequency questionnaire among French adults and adolescents. *European Journal of Clinical Nutrition, 63*(2), 282-291. doi: 10.1038/sj.ejcn.1602914

Dorea, J. G., and Donangelo, C. M. (2006). Early (in uterus and infant) exposure to mercury and lead. *Clinical Nutrition, 25*(3), 369-376. doi: 10.1016/j.clnu.2005.10.007

Duarte-Salles, T., von Stedingk, H., Granum, B., Gützkow, K. B., Rydberg, P., Törnqvist, M., Mendez, M. A., Brunborg, G., Brantsæter, A. L., Meltzer, H. M., Alexander, J., and Haugen, M. (2013).

Dietary acrylamide intake during pregnancy and fetal growth-results from the Norwegian Mother and child cohort study (MoBa). [Article]. *Environmental Health Perspectives, 121*(3), 374-379. doi: 10.1289/ehp.1205396

Dubuisson, C., Lioret, S., Touvier, M., Dufour, A., Calamassi-Tran, G., Volatier, J. L., and Lafay, L. (2010). Trends in food and nutritional intakes of French adults from 1999 to 2007: Results from the INCA surveys. *British Journal of Nutrition, 103*(7), 1035-1048. doi: 10.1017/S0007114509992625

EFSA (2013). Scientific Opinion of the Panel on Plant Protection Products and their Residues (PPR) on the identification of pesticides to be included in cumulative assessment groups on the basis of their toxicological profile. EFSA Journal; 11(7):3293. [131 pp.]. doi: 10.2903/j.efsa.2013.3293 (EFSA, Parma, Italy).

EFSA (2014). Guidance on the EU Menu methodology. EFSA Journal; 12(12. [3944-n/a pp.]. doi: 10.2903/j.efsa.2014.3944.

Etchevers, A., Bretin, P., Lecoffre, C., Bidondo, M. L., Le Strat, Y., Glorennec, P., and Le Tertre, A. (2014). Blood lead levels and risk factors in young children in France, 2008-2009. [Article]. *International Journal of Hygiene and Environmental Health, 217*(4-5), 528-537. doi: 10.1016/j.ijheh.2013.10.002

Gazan, R., Béchaux, C., Crépet, A., Sirot, V., Drouillet-Pinard, P., Dubuisson, C., and Havard, S. (2016). Dietary patterns in the French adult population: a study from the second French national cross-sectional dietary survey (INCA2) (2006–2007). [Article in Press]. *British Journal of Nutrition*, 1-16. doi: 10.1017/S0007114516001549

GEMS/Food-EURO (2013). Addendum 2013 - Second Workshop on Reliable Evaluation of Low-Level Contamination of Food. Workshop in the frame of GEMS/Food-EURO. Technical report. [pp.]. (WHO, Kulmbach, Germany).

Gillis, N., and Glineur, F. (2010). Using underapproximations for sparse nonnegative matrix factorization. *Pattern Recognition, 43*(4), 1676-1687. doi: http://dx.doi.org/10.1016/j.patcog.2009.11.013

Gillis, N., and Plemmons, R. J. (2013). Sparse nonnegative matrix underapproximation and its application to hyperspectral image analysis. *Linear Algebra and its Applications, 438*(10), 3991-4007. doi: http://dx.doi.org/10.1016/j.laa.2012.04.033

Hercberg, S. (2000). *Portions alimentaires : manuel photos pour l'estimation des quantités*. Paris: Economica.

Hercberg, S., Deheeger, M., and Preziosi, P. (2002). SU.VI.MAX. Portions alimentaires : Manuel photos pour l'estimation des quantités: Paris: BMJ Publishing Group.

Heude, B., Forhan, A., Slama, R., Douhaud, L., Bedel, S., Saurel-Cubizolles, M. J., Hankard, R., Thiebaugeorges, O., De Agostini, M., Annesi-Maesano, I., Kaminski, M., Charles, M. A., Annesi-Maesano, I., Bernard, J. Y., Botton, J., Charles, M. A., Dargent-Molina, P., de Lauzon-Guillain, B., Ducimetière, P., de Agostini, M., Foliguet, B., Forhan, A., Fritel, X., Germa, A., Goua, V., Hankard, R., Heude, B., Kaminski, M., Larroque, B., Lelong, N., Lepeule, J., Magnin, G., Marchand, L., Nabet, C., Pierre, F., Slama, R., Saurel-Cubizolles, M. J., Schweitzer, M., and Thiebaugeorges, O. (2016). Cohort Profile: The EDEN mother-child cohort on the prenatal and early postnatal determinants of child health and development. [Article]. *International Journal of Epidemiology, 45*(2), 353-363. doi: 10.1093/ije/dyv151

Hulin, M., Bemrah, N., Nougadère, A., Volatier, J. L., Sirot, V., and Leblanc, J. C. (2014). Assessment of infant exposure to food chemicals: The French Total Diet Study design. *Food Additives and Contaminants - Part A Chemistry, Analysis, Control, Exposure and Risk Assessment, 31*(7), 1226-1239. doi: 10.1080/19440049.2014.921937

Kim, D. J., and Holowaty, E. J. (2003). Brief, validated survey instruments for the measurement of fruit and vegetable intakes in adults: A review. [Article]. *Preventive Medicine, 36*(4), 440-447. doi: 10.1016/S0091-7435(02)00040-3

Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: An R package for multivariate analysis. [Article]. *Journal of Statistical Software, 25*(1), 1-18.

Leblanc, J. C., Guérin, T., Noël, L., Calamassi-Tran, G., Volatier, J. L., and Verger, P. (2005). Dietary exposure estimates of 18 elements from the 1st French Total Diet Study. [Article]. *Food Additives and Contaminants, 22*(7), 624-641. doi: 10.1080/02652030500135367

Lee, D. D., and Seung, H. S. (2001). *Algorithms for non-negative matrix factorization.* Paper presented at the Advances in Neural Information Processing Systems.

Lioret, S., Dubuisson, C., Dufour, A., Touvier, M., Calamassi-Tran, G., Maire, B., Volatier, J. L., and Lafay, L. (2010). Trends in food intake in French children from 1999 to 2007: Results from the INCA (étude Individuelle Nationale des Consommations Alimentaires) dietary surveys. *British Journal of Nutrition, 103*(4), 585-601. doi: 10.1017/S0007114509992078

Mendez, M. A., Garcia-Esteban, R., Guxens, M., Vrijheid, M., Kogevinas, M., Goñi, F., Fochs, S., and Sunyer, J. (2011). Prenatal organochlorine compound exposure, rapid weight gain, and overweight in infancy. *Environmental Health Perspectives, 119*(2), 272-278. doi: 10.1289/ehp.1002169

Mørup, M., and Hansen, L. K. (2009). Automatic relevance determination for multi-way models. *Journal of Chemometrics, 23*(7-8), 352-363. doi: 10.1002/cem.1223

Nougadère, A., Sirot, V., Kadar, A., Fastier, A., Truchot, E., Vergnet, C., Hommet, F., Baylé, J., Gros, P., and Leblanc, J. C. (2012). Total diet study on pesticide residues in France: Levels in food as consumed and chronic dietary risk to consumers. [Article]. *Environment International, 45*(1), 135-150. doi: 10.1016/j.envint.2012.02.001

Pedersen, M., Von Stedingk, H., Botsivali, M., Agramunt, S., Alexander, J., Brunborg, G., Chatzi, L., Fleming, S., Fthenou, E., Granum, B., Gutzkow, K. B., Hardie, L. J., Knudsen, L. E., Kyrtopoulos, S. A., Mendez, M. A., Merlo, D. F., Nielsen, J. K., Rydberg, P., Segerbäck, D., Sunyer, J., Wright, J., Törnqvist, M., Kleinjans, J. C., and Kogevinas, M. (2012). Birth weight, head circumference, and prenatal exposure to acrylamide from maternal diet: The European prospective mother-child study (NewGeneris). [Article]. *Environmental Health Perspectives, 120*(12), 1739-1745. doi: 10.1289/ehp.1205327

Pinto, E., Barros, H., and Santos Silva, I. D. (2009). Dietary intake and nutritional adequacy prior to conception and during pregnancy: A follow-up study in the north of Portugal. [Article]. *Public Health Nutrition, 12*(7), 922-931. doi: 10.1017/S1368980008003595

Rivière, G., Sirot, V., Tard, A., Jean, J., Marchand, P., Veyrand, B., Le Bizec, B., and Leblanc, J. C. (2014). Food risk assessment for perfluoroalkyl acids and brominated flame retardants in the French population: Results from the second French total diet study. [Article]. *Science of the Total Environment, 491-492*, 176-183. doi: 10.1016/j.scitotenv.2014.01.104

Rothenberg, E. M. (2009). Experience of dietary assessment and validation from three swedish studies in the elderly. [Article]. *European Journal of Clinical Nutrition, 63*, S64-S68. doi: 10.1038/ejcn.2008.67

Schmidt, M. N., Winther, O., and Kaihansen, L. (2009) Bayesian non-negative matrix factorization. *Vol. 5441. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (pp. 540-547).

Sirot, V., Fremy, J.-M., and Leblanc, J.-C. (2013). Dietary exposure to mycotoxins and health risk assessment in the second French total diet study. *Food and Chemical Toxicology, 52*, 1-11. doi: http://dx.doi.org/10.1016/j.fct.2012.10.036

Sirot, V., Hommet, F., Tard, A., and Leblanc, J. C. (2012a). Dietary acrylamide exposure of the French population: Results of the second French Total Diet Study. [Article]. *Food and Chemical Toxicology, 50*(3-4), 889-894. doi: 10.1016/j.fct.2011.12.033

Sirot, V., Tard, A., Venisseau, A., Brosseaud, A., Marchand, P., Le Bizec, B., and Leblanc, J. C. (2012b). Dietary exposure to polychlorinated dibenzo-p-dioxins, polychlorinated dibenzofurans and polychlorinated biphenyls of the French population: Results of the second French Total Diet Study. [Article]. *Chemosphere, 88*(4), 492-500. doi: 10.1016/j.chemosphere.2012.03.004

Sirot, V., Volatier, J. L., Calamassi-Tran, G., Dubuisson, C., Ménard, C., Dufour, A., and Leblanc, J. C. (2009). Core food of the French food supply: Second Total Diet Study. *Food Additives and Contaminants - Part A Chemistry, Analysis, Control, Exposure and Risk Assessment, 26*(5), 623-639. doi: 10.1080/02652030802695506

Sly, P. D., and Flack, F. (2008) Susceptibility of children to environmental pollutants. *Vol. 1140. Annals of the New York Academy of Sciences* (pp. 163-183).

Sy, M. M., Feinberg, M., Verger, P., Barré, T., Clémençon, S., and Crépet, A. (2013). New approach for the assessment of cluster diets. *Food and Chemical Toxicology, 52*, 180-187. doi: 10.1016/j.fct.2012.11.005

Tan, V. Y. F., and Févotte, C. (2013). Automatic relevance determination in nonnegative matrix factorization with the (β)-divergence. [Article]. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 35*(7), 1592-1605. doi: 10.1109/TPAMI.2012.240

Team, R. C. (2016). R: A Language and Environment for Statistical Computing (Version 3.3.2): R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Traoré, T., Béchaux, C., Sirot, V., and Crépet, A. (2016). To which chemical mixtures is the French population exposed? Mixture identification from the second French Total Diet Study. [Article]. *Food and Chemical Toxicology, 98*, 179-188. doi: 10.1016/j.fct.2016.10.028

Valvi, D., Mendez, M. A., Martinez, D., Grimalt, J. O., Torrent, M., Sunyer, J., and Vrijheid, M. (2012). Prenatal concentrations of polychlorinated biphenyls, DDE, and DDT and overweight in children: A prospective birth cohort study. *Environmental Health Perspectives, 120*(3), 451-457. doi: 10.1289/ehp.1103862

Vandentorren, S., Bois, C., Pirus, C., Sarter, H., Salines, G., and Leridon, H. (2009). Rationales, design and recruitment for the Elfe longitudinal study. [Article]. *BMC Pediatrics, 9*, 58. doi: 10.1186/1471-2431-9-58

Veyrand, B., Sirot, V., Durand, S., Pollono, C., Marchand, P., Dervilly-Pinel, G., Tard, A., Leblanc, J.-C., and Le Bizec, B. (2013). Human dietary exposure to polycyclic aromatic hydrocarbons: Results of the second French Total Diet Study. *Environment International, 54*, 11-17. doi: http://dx.doi.org/10.1016/j.envint.2012.12.011

Wang, G., Kossenkov, A. V., and Ochs, M. F. (2006). LS-NMF: A modified non-negative matrix factorization algorithm utilizing uncertainty estimates. *BMC Bioinformatics, 7*. doi: 10.1186/1471-2105-7-175

Yang, Z., Zhu, Z., and Oja, E. (2010). Automatic Rank Determination in Projective Nonnegative Matrix Factorization. In V. Vigneron, V. Zarzoso, E. Moreau, R. Gribonval & E. Vincent (Eds.), *Latent Variable Analysis and Signal Separation: 9th International Conference, LVA/ICA 2010, St. Malo, France, September 27-30, 2010. Proceedings* (pp. 514-521). Berlin, Heidelberg: Springer Berlin Heidelberg.

Zetlaoui, M., Feinberg, M., Verger, P., and Clémençon, S. (2011). Extraction of food consumption systems by nonnegative matrix factorization (NMF) for the assessment of food choices. *Biometrics, 67*(4), 1647-1658. doi: 10.1111/j.1541-0420.2011.01588.x