Selecting mixtures on the basis of dietary exposure and hazard data: application to pesticide exposure in the European population in relation to steatosis

Amélie Crépet, Marie Vanacker, Corinne Sprong, Waldo de Boer, Urska Blaznik, Marc Kennedy, Chris Anagnostopoulos, Despo Louca Christodoulou, Jiří Ruprich, Irena Rehurkova, et al.

Selecting mixtures on the basis of dietary exposure and hazard data: application to pesticide exposure in the European population in relation to steatosis

Amélie Crépet[1], Marie Vanacker[1], Corinne Sprong[2], Waldo de Boer[3], Urska Blaznik[4], Marc Kennedy[5], Chris Anagnostopoulos[6], Despo Louca Christodoulou[7], Jiří Ruprich[8], Irena Rehurkova[8], José Luis Domingo[9], Bodil Hamborg Jensen[10], Francesca Metruccio[11], Angelo Moretto[11-12], Liesbeth Jacxsens[13], Pieter Spanoghe[14], David Senaeve[14], Hilko van der Voet[3], and Jacob van Klaveren[2]

[1] ANSES, French Agency for Food, Environmental and Occupational Health and Safety, Risk assessment department, Methodology and studies unit, 947001, Maisons-Alfort, France
[2] RIVM, National Institute for Public Health and the Environment, The Netherlands, PO Box 1, 3720 BA Bilthoven, The Netherlands
[3] Wageningen University & Research, Biometris, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands
[4] National Institute of Public Health, Environmental Health Centre, Trubarjeva 2, Ljubljana, Slovenia
[5] Fera Science Ltd, Sand Hutton, York, YO41 1LZ, United Kingdom
[6] Benaki Phytopathological Institute, Department of Pesticides Control and Phytopharmacy, Laboratory of Pesticide Residues, 8 Stefanou Delta Street, Kifissia, Athens, 14561, Greece.
[7] State General Laboratory, Ministry of, Nicosia, Cyprus
[8] National Institute of Public Health in Prague, Centre for Health, Nutrition and Food, Brno, Czech Republic
[9] Laboratory of Toxicology and Environmental Health, School of Medicine, IISPV, Universitat Rovira I Virgili, Reus, Catalonia, Spain
[10] Technical University of Denmark, National Food Institute, Division of Risk Assessment and Nutrition, Kemitorvet, Building 201, DK 2800 Lyngby, Denmark
[11] ICPS, International Centre for Pesticides and Health Risk Prevention, ASST Fatebenefratelli Sacco, Ospedale L. sacco via GB Grassi 74, 20157 Milano, Italy.
[12] Department of Biomedical and Clinical Sciences, Università degli Studi di Milano, Milan, Italy.
[13] Ugent university, Coupure links, 653b Ghent, Belgium
[14] Ugent, Laboratory for crop protection chemistry, Coupure links, 653b Ghent, Belgium

Abstract

Populations are exposed to mixtures of pesticides through their diet on a daily basis. The question of which substances should be assessed together remains a major challenge due to the complexity of the mixtures. In addition, the associated risk is difficult to characterise. The EuroMix project (European Test and Risk Assessment Strategies for Mixtures) has developed a strategy for mixture risk assessment. In particular, it has proposed a methodology that combines exposures and hazard information to identify relevant mixtures of chemicals belonging to any cumulative assessment group (CAG) to which the European population is exposed via food. For the purposes of this study, food consumption and pesticide residue data in food and drinking water were obtained from national surveys in nine European countries. Mixtures of pesticides were identified by a sparse non-negative matrix underestimation (SNMU) applied to the specific liver steatosis effect in children from 11 to 15 years of age, and in adults from 18 to 64 years of age in nine European countries. Exposures and mixtures of 144 pesticides were evaluated through four different scenarios: (1) chronic exposure with a merged concentration dataset in the adult population, (2) chronic exposure with country-specific concentration datasets in the adult population, (3) acute exposure with a merged concentration dataset in the adult population, and (4) chronic exposure with a merged concentration dataset in the paediatric population. The relative potency factors of each

substance were calculated to express their potency relative to flusilazole, which was chosen as the reference compound. The selection of mixtures and the evaluation of exposures for each country were carried out using the Monte Carlo Risk Assessment (MCRA) software.
Concerning chronic exposure, one mixture explained the largest proportion of the total variance for each country, while in acute exposure, several mixtures were often involved. The results showed that there were 15 main pesticides in the mixtures, with a high contribution of imazalil and dithiocarbamate. Since the concentrations provided by the different countries were merged in the scenario using merged concentration data, differences between countries result from differences in food consumption behaviours. These results support the approach that using merged concentration data to estimate exposures in Europe seems to be realistic, as foods are traded across European borders. The originality of the proposed approach was to start from a CAG and to integrate information from combined exposures to identify a refined list of mixtures with fewer components. As this approach was sensitive to the input data and required significant resources, efforts should continue regarding data collection and harmonisation among the different aspects within the pesticides regulatory framework, and to develop methods to group substances and mixtures to characterise the risk.

Keywords
Mixture prioritization; Cumulative assessment group; Sparse non-negative matrix underestimation; Relative Potency Factors; Dietary exposure and hazard

Highlights

- Mixtures were prioritized from dietary exposure and hazard data
- Acute and chronic exposure were estimated for 9 European countries
- One main pesticide mixture drives the risk related to chronic exposure and steatosis in Europe
- Imazalil and dithiocarbamate contribute the most to the European pesticide mixture exposure

1. Introduction

Through the environment and diet, on a daily basis populations are exposed to mixtures of chemicals that can interact and cause health effects. Due to the complexity of mixtures, the associated risk is difficult to characterise. Over the past decade, considerable efforts have been made to propose concepts, methods, guidance and applications for the risk assessment of mixtures (Boobis et al., 2008; EFSA, 2007, 2008; Fox et al., 2017; WHO, 2009). Given the multitude of possible combinations, the question of which substances to assess together remains a major challenge. One solution is to perform risk assessments for chemicals belonging to the same chemical family and/or having the same mode of action. In this way, the European Food Safety Authority (EFSA) proposed a hazard-wise method based on "common adverse outcomes" to group pesticides into "cumulative assessment groups" (CAGs) (EFSA, 2013b; Nielsen et al., 2012; RIVM et al., 2013). Four levels of criteria for grouping were defined, with each higher level being more refined: target organs (level 1), specific phenotypic effects (level 2), mode of action (level 3), and mechanism of action (level 4). Currently, level 1 and 2 CAGs have been identified in the nervous system and the thyroid

for pesticides. Preliminary work has been done on effects on the liver, adrenal glands, eyes, and developmental and reproductive systems (EFSA, 2012; RIVM et al., 2013). Dose addition is the default hypothesis to assess the risk of these CAGs, but the appropriateness of this assumption is hardly ever investigated experimentally. The difficulty in cumulative risk assessment is the lack of information on hazard and exposure of the substances classified into a certain CAG. Firstly, for several pesticides, grouping into a certain CAG can be based on a small number of observations, thereby introducing uncertainties regarding CAG membership and relative potency in comparison to other substances in a CAG. Secondly, the mode and mechanism of action is unknown for many substances, and this may not only hamper refinement into level 3 and level 4 CAGs, but also introduce uncertainties in addressing the combined effect. Because of this, there is a need for greater efforts to study the modes and mechanisms of action of pesticides. However, as a certain CAG can contain a high number of components, it is necessary to prioritise the substances to be assessed in mixture testing. Like all risk assessments, combined risk assessments to chemicals should not be based on the hazard (toxicological properties) alone, but also on population exposure. Combined exposure can be estimated by combining food consumption patterns of individuals in a population with occurrence levels of chemicals in food. The number of combinations of compounds to which an individual in a population is exposed can be large. Therefore, it is essential to develop a strategy that considers actual exposure to extract the most relevant mixtures to which the population is exposed (Crépet et al., 2013) as a prioritisation tool for further studies.

The present study is part of the EuroMix project (No. 633172, H2020-SFS-2014-2) which has developed a strategy for mixture risk assessment. It proposes a prioritisation methodology combining both exposure and hazard information to identify the most relevant mixtures of chemicals belonging to any CAG to which European populations are exposed chronically and acutely via food. The proposed approach starts from the list of substances in a defined CAG, and reduces this list by using risk-based identification of co-occurring pesticides in the diet for a given time frame. The mixture selection approach is based on sparse non-negative matrix underapproximation (SNMU) (Gillis and Plemmons, 2013), which is a statistical method making it possible to select the main mixtures. SNMU is a modified version of non-negative matrix factorisation (NMF) (Lee and Seung, 2001), recently used to identify the main mixtures associated with the diet (Béchaux et al., 2013; Traoré et al., 2016; Traoré et al., 2018). The proposed approach was implemented using the web-based Monte Carlo Risk Assessment (MCRA) platform, version 8.2 (Boon et al., 2015). It was applied to the level 2 CAG for liver steatosis defined by EFSA (Nielsen et al., 2012; RIVM et al., 2013, 2016) and on exposure data from several European countries. If needed, the identified substances in the mixtures and their individual components will be further studied using several *in vitro* and *in vivo* tests. The results of these additional tests may provide a more precise picture of the potency and the mode of action of each substance. The mixture of substances will also be tested *in vitro* and *in vivo* to refine the assumptions made on the dose- and/or response addition. The aim of our study was to describe the mixture selection procedure and the identified priority mixtures for further testing. The results we obtained aim to facilitate a cost-effective test procedure.

2. Materials and methods

2.1 Exposure and hazard data to identify mixtures

The proposed method is based on a combination of exposure and hazard information to define mixtures. In practice it consists in 1) selecting a CAG and its level of grouping, 2) calculating the exposures for each pesticide belonging to the selected CAG by combining quantities of consumed food with the substance concentrations in those foods, 3) converting the exposure of each substance to the toxicity equivalent value of the substance of reference for the selected CAG, and 4) applying statistical methods to the converted exposures to determine the main mixtures to which the studied population is exposed.

2.2 Data

2.2.1 Hazard data

The CAG for liver effects was chosen for the specific steatosis effect (second level of liver toxicity). The list of pesticides in this CAG with their corresponding NOAEL and/or LOAEL was established from three reports supported by EFSA (Nielsen et al., 2012; RIVM et al., 2013, 2016) and their associated database. The underlying studies were critically evaluated regarding the following criteria, which yielded a total of 155 substances:

- All repeated-dose (short-term and long-term) toxicology studies based on oral administration (diet, gavage, capsule) at the NOAELs/LOAELs were taken into consideration.
- Inhalation studies were considered only for pesticides that are gasses and that could therefore not be toxicologically tested via the oral route.
- Studies by the dermal route were not reported, except for substances for which no data were available concerning the oral route.
- Acute LD_{50} studies were not considered.
- *In vitro* studies were considered for information on the mechanism/mode of action only.
- Studies performed with metabolites were not included, except when the metabolite itself was used in the toxicity studies instead of the parent compound due to its high instability.
- In the particular case where the active substance consists of isomer mixtures, the studies performed with the racemic mixture and those carried out with the different isomers were reported.
- When different isomers and/or variants were considered to be toxicologically equivalent, the same specific effect was applied and the studies were reported only once.

Substances were coded using the ParamCodes from the harmonised European Standard Sample Description 1 format SDD1 (EFSA, 2010). Substances were removed if no ParamCode coding for pesticides, no NOAL or no LOAEL (copper compounds) were available. Some substances sharing the same residue definition (benalaxyl-M and benalaxyl, cypermethrin and alpha-cypermethrin, metam and dazomet, metalaxyl-M and metalaxyl,

triadimefon and triadimenol) were presented together in the database. This approach resulted in a total of 144 pesticides.

Relative potency factors (RPFs) were calculated to express the potency of each substance in the CAG relative to a selected reference compound chosen based on the following criteria:

- Considering that longer-term studies (i.e.12, 18 and 24 months) were generally performed using lower concentrations compared to shorter-term studies (i.e. 28 or 90 days), priority was given to long-term studies.
- Compounds characterised by an NOAEL causing fatty changes (steatosis) between 0 and 1 mg/kg bw/day, were first selected (to avoid the selection of an index compound eliciting other organ and/or different liver effects at doses lower that those eliciting fatty changes).
- The second step in selection was made on the basis of the LOAEL/NOAEL ratio (between 1 and 5) to avoid dose-spacing uncertainties.
- The third step in selection was made taking into consideration only those compounds also causing cell degeneration/cell alteration or cell death at similar or higher doses.
- As a final step, the compound with more studies showing liver effects was chosen as the reference compound.

The minimum required data set for calculation of potency was a well-performed chronic study with a dose-range that could provide a LOAEL for steatosis. The more studies available, the extent to which the above mentioned criteria could be applied to select the NOAEL or LOAEL of a particular substance to calculate its potency.

Flusilazole complying with the above criteria was selected as the reference compound. Data came from 4 long term studies where liver effects were evident and LOAEL/NOAEL ratio for fatty changes spaced between 2 and 5. Its NOAEL for fatty changes was of 0.53 mg/kg bw/day.

For each compound, the NOAEL of flusilazole was divided by the NOAEL of the particular compound, which yielded the RPF. When no NOAEL was available, the LOAEL divided by three was used as an assumption of the NOAEL. The RPFs make it possible to convert exposure to the substances into the "toxicity unit" of the reference compound, and thus to compare the exposure levels between substances within a CAG.

2.2.2 Consumption data

Food consumption data from the different countries were coded according to the harmonised FoodEx1 coding system (EFSA, 2011). FoodEx1 is a hierarchical system based on 20 main food categories divided into subgroups up to a maximum of 4 levels. For example, chocolate cake is given a numerical code responding to 'grain and grain-based products' at level 1, to 'fine bakery wares' at level 2, to 'pastries and cakes' at level 3, and to 'chocolate cake' at level 4. The age and body weight were also available for each individual. It was decided to focus on the adult population aged between 18 and 64 years, and for countries where data were available, on the paediatric population aged between 11 and 15 years, as these were the

age ranges shared by the largest number of different country surveys. A summary of consumption data is shown in Table 1.

Belgium: Consumption data were provided by the National Institute of Public Health from a consumption study conducted in 2004 by De Vriese et al. (2005). The study included 3,214 participants over 15 years of age who were interviewed about their consumption in a 2 x 24-hour period (repeated non-consecutive 24h recall), and asked to fill in a questionnaire about food frequency.

Cyprus: Consumption data were provided from a national study evaluating the frequency of eating disorder cases (Cyprus study on eating disorders among high school students called "Child Health"), which was conducted in 2003. In this study, food consumption data were collected for 303 children, aged between 11 to 15 years, using a 3-day estimated dietary record. No data were collected in the adult population. Most, but not all, dietary records were collected over consecutive days. Amounts consumed were estimated using food package sizes and household measures (e.g. cups and spoons). The consumed quantities of 1,043 food items were collected.

Czech Republic: Consumption data were provided by the National Institute of Public Health. They are from the national food consumption survey named SISP04 (Ruprich et al., 2006). Food consumption data were collected in 2003 and 2004 for 2,590 individuals representing the entire country, both genders and ages 4 to 90 years. This study used a 2 x 24h recall design (with non-correlated days D1 and D2 separated by more than 14 days). The face-to-face method was used for data collection. Reported data on food types were aggregated into 514 groups.

Denmark: Consumption data were provided by the Division of Risk Assessment and Nutrition at the National Food Institute. The data were collected as part of DANSDA (DAnish National Survey of Diet and physical Activity) 2005–2008, and constitute a subset of the data reported in "Dietary habits in Denmark 2003–2008" (Pedersen et al., 2010). Food consumption data were recorded concerning 2,700 Danish consumers aged 4 to 75 years. The dataset records food and beverages consumed over 7 consecutive days. The individuals were drawn as a simple random sample from the general population registration system. DANSDA used a 7-day pre-coded (semi-closed) food diary with answering categories for the most commonly consumed foods and drinks in the Danish diet. Data on a total of 414 food items were collected.

France: Consumption data were drawn from the second "Individual and National Study on Food Consumption" (INCA2) carried out by the French Agency for Food, Environmental and Occupational Health and Safety between late 2005 and April 2007. Two independent random samples were included in this survey: 1,455 children aged between 3 to 17 years (Lioret et al., 2010) and 2,624 adults aged between 18 to 79 years (Dubuisson et al., 2010). Participants were selected using a three-stage random design stratified by region of residence, size of urban area, and population group (adults and children). Subjects completed a 7-day food record diary and portion sizes were estimated through photographs compiled in a manual adapted from the Su-Vi-Max photographic booklet (Hercberg et al., 1994). The consumed quantities of 1,280 food items per day were collected.

Greece: Food consumption data were obtained from 10 surveys (Crete Region) conducted by the University of Crete, Faculty of Medicine, Department of Preventive Medicine and

Nutrition between 1988 and 2004 (Bertsias et al., 2003; Kafatos et al., 1991; Linardakis et al., 2008; Moschandreas and Kafatos, 1999; University of Crete, March 2016; Xatzis et al., 2004). In total, the surveys covered the dietary habits of 1,640 adults aged between 18 to 94 years and 528 children aged between 11 and 15 years living in Crete. The consumed quantities of approximately 72 food items per day were collected. Dietary consumption was measured using the 24-h recall method.

Netherlands: Food consumption data were obtained from two surveys: the Dutch National Food Consumption Survey (DNFCS)-Young children (Ocké et al., 2008), and the DNFCS 2007–2010 (van Rossum et al., 2011). The DNFCS-Young children survey covered the dietary habits of 1,279 young children aged 2 to 6 years representatively selected from the Dutch population, and was conducted in 2005 and 2006. The DNFCS 2007–2010 includes the dietary habits of 3,819 people aged 7 to 69 years representatively selected form the Dutch population. Dietary consumption was measured using the 24-h recall method on two non-consecutive days. The survey included 1,599 food items. Results of the consumption surveys were weighted for small deviations in socio-demographic characteristics in order to obtain results that are representative of the Dutch population.

Slovenia: Food consumption data were obtained from the National Food Consumption Survey (CRP 2008), provided by the National Institute of Public Health Slovenia. The survey covered the period 2007–2008 with data on the individual level for 407 persons, both genders, aged between 18 to 65 years. The participants were selected from the Central Register of Population in Slovenia with a two-stage, stratified sample design. Dietary consumption was measured using the 24-h recall method for one survey day. Consumed amounts of foods were estimated using a national picture book, complemented with household measures and portions indicated in standard recipes. A total of 283 food groups were recorded.

Spain: Food consumption data were provided from the Encuesta ENIDE survey (AESAN, 2011). Data were collected in 2011 for 3,386 individuals aged between 18 to 71 years. The consumed quantities of approximately 72 food items per day were collected. Dietary consumption was measured using the 24-h recall method.

United Kingdom: Food consumption data were extracted from the National Diet and Nutrition Survey (NDNS). The survey covered the period from July 2000 to June 2001 and included 1,724 adult respondents aged 19 to 64 years. After an initial face-to-face interview (CAPI method), the participants recorded dietary consumption in a 7-day consecutive diary (Henderson et al., 2002). A total of 490 food items were recorded.

2.2.3 Concentration data

Concentration data in food and drinking water were obtained from annual control and monitoring programmes between 2010 and 2014 for the countries for which this was available (Table 1). Data comprise pesticides levels measured in raw agricultural commodities and/or food as consumed (e.g. juices). Samples obtained by objective or selective sampling were included, whereas samples obtained by less formal sampling strategies were excluded since they are not representative of the market. A zero value was attributed to analytical results reported as below the limit of detection (LOD), following the optimistic basic scenario

included in guidance from the European Food Safety Authority (EFSA, 2012). A merged dataset was created by combining data from all countries. The merged data set contained 127 pesticides in the steatosis CAG, of which 93 pesticides had at least one sample above the LOD. This resulted in 3,161,615 analyses applied to 204 raw agricultural food commodities, from which 0.72% of measurements were quantified. For two countries, Spain and the United Kingdom, access to specific national monitoring programmes for concentrations of substances was not available.

Belgium: Concentration data on pesticides were collected between 2011 and 2014, as per the national monitoring programme on pesticides. The monitoring was carried out by the National Institute for Food Safety (FAVV/AFSCA). The datasets contain a total of n = 101,319 samples, of which 1,141 (1.12%) were positive detections of 135 different compounds in 112 raw agricultural commodities. 115 pesticides were classified in the steatosis CAG and out of these, 39 had at least one sample above the LOD. 0.87% of pesticides were quantified in the CAG.

Cyprus: Concentration data were collected between 2011 and 2014 as part of the national monitoring programmes. The dataset contained analytical results for up to 346 pesticides out of which 81 were classified in the steatosis CAG. A total of 48 of these pesticides had at least one sample above the LOD. This resulted in 124,599 analyses, of which 0.72% quantified values in 68 raw agricultural commodities.

Czech Republic: Concentration data generated between 2011 and 2014 were obtained from the national database of analytical results for food monitoring. From the 58 substances analysed, 42 pesticides were selected as relevant for the steatosis CAG, and 37 pesticides had at least one sample above the LOD. This resulted in 153,696 measurements in 114 raw agricultural commodities, for which 1.35% were quantified.

Denmark: Data were collected between 2011 and 2014 by the Danish Veterinary and Food Administration and represented commodities sold on the Danish market. The dataset contained analytical results for up to 280 pesticides. Among them, 95 were included in the steatosis CAG, and 58 pesticides had at least one sample above the LOD. In total, 503,879 measurements were recorded in 190 raw agricultural food commodities, and 0.62% of them contained quantified values.

France: Concentration data were collected between 2010 and 2014 by the French ministries in charge of consumer affairs, agriculture and health. The monitoring programmes provided analytical results for up to 194 pesticides. Among them, 120 were in the steatosis CAG, and 70 substances had at least one sample above the LOD. This represented 907,565 measurements in 153 raw agricultural food commodities, of which 0.53% were quantified.

Greece: Pesticide residue data were provided by the Hellenic Ministry of Rural Development and Food (Department of Plant Protection Products & Biocides) for the period between 2010 and 2014. Among the analysed pesticides, 91 pesticides were relevant for the steatosis CAG, and 56 pesticides had at least one sample above the LOD. This represented 324,561 measurements and 0.65% were quantified in 68 raw agricultural food commodities.

Netherlands: Concentration data were collected between 2010 and 2013. The dataset contained analytical results for 665 pesticides, of which 110 were included in the steatosis CAG. In all, 67 pesticides had at least one sample above the LOD. This resulted in 643,538 analyses with 0.89% quantified values in 131 raw agricultural food commodities.

Slovenia: Slovenian concentration data were collected between 2011 and 2014 by the Ministry of Agriculture, Forestry and Food. Among the 109 pesticides analysed, 87 belonged to the steatosis CAG, and 40 pesticides had at least one sample above the LOD. The dataset contained 109,810 analyses with 0.49% quantified values in 70 raw agricultural food commodities.

2.2.4 Data matching

Matching concentration and consumption data: All data were uploaded into the MCRA software. To match food consumption data with concentration data in raw agricultural products, a conversion table was used (Boon et al., 2015). This conversion table is based on Dutch recipes and contains conversion factors to convert foods classified according to FoodEx1 to their edible raw agricultural commodity (RAC) ingredients (e.g. an apple pie is broken down in its mass percentage of apple, flour, butter, sugar and eggs, or the mass percentage of raw spinach to obtain 100 g of cooked spinach) The conversion table included information on important processing steps, such as cooking, milling and juicing. Processing factors from the German *Bundesinstitut für Risikobewertung* (BfR; accessed on 1 September 2015) were used to account for the effect of these processing steps on exposure levels. For 46 out of the 144 pesticides, processing factors were available.

Matching hazard and exposure data: Pesticides in the CAG lists from EFSA and DTU are given as parent compounds rather than residues, whereas concentration data were mostly expressed as residue definitions for enforcement, which can be a single parent compound, one or more metabolites (i.e. pesticide metabolites in plants or animals), or a combination of the parent compound and metabolites. To match the parent compounds in the CAG to the concentration data, the SSD1 ParamCodes for current residue definitions were obtained from the pesticides database of the European Commission; these are the residue definitions for enforcement. It should be noted that according to the EFSA Opinion of 2012, residue definitions for risk assessment should be used rather than residue definitions for enforcement. The residue definition for risk assessment can be obtained by applying conversion factors to concentrations obtained from the residue definition for enforcement. For simplicity, these conversion factors were assumed to be 1.

2.3 Exposure calculation and scenarios

The optimistic basic approach of EFSA (2012) implemented in the MCRA software was followed to calculate both chronic (long-term) and acute (short-term) exposure. Under this approach, values lower than the LOD as well as missing values were set to 0. The empirical distributions were used for concentration data and processing factors were applied to integrate the effect of process on concentration levels. No between-lot and sample variability factors were considered. In the chronic scenario, the mean of available concentration values per pesticide/food combination was multiplied by the mean of consumed food quantity on the different recorded days for each individual, which is the simple Observed Individual Means (OIM) model (EFSA, 2012). In the acute scenario, concentration values and individual-days

of consumption were randomly selected by Monte Carlo simulations in their empirical distributions to produce individual-day exposure to each pesticide.

Therefore, exposure per day was calculated by multiplying the consumed quantities per food for each individual by the concentrations of the different substances in this food, following the chronic and acute scenarios. Then, the exposures from the different foods for each substance were summed, divided by the body weight of each individual, and multiplied by the relative potency factors RPF:

$$E_{ijs} = \frac{\sum_{f=1}^{F} q_{ijf}\ c_{ijfs}}{bw_i} \text{ x } \text{RPF}_s$$

where E_{ijs} is the exposure to substance *s* by individual *i* on day *j* (in microgram substance per kg body weight), q_{ijf} is the consumed quantity of food f (in g) by the individual *i* on day *j*, c_{ijfs} is the concentration of substance *s* in food *f* eaten by individual *i* on day *j* (in mg/kg), and bw_i is the body weight of individual *i* (in kg). F is the number of foods in which the substance is present. Note that all exposures are zero or positive values.

Four exposure scenarios were tested and compared:

1. Chronic exposure calculated with the merged concentration dataset for the adult population (18-64 years).
2. Chronic exposure calculated with the country-specific concentration datasets for the adult population (18-64 years).
3. Acute exposure calculated with the merged concentration dataset for the adult population (18-64 years).
4. Chronic exposure calculated with the merged concentration dataset for children aged between 11 and 15 years.

2.4 Mixture selection method

The method used to extract the mixtures from the matrix of exposures E is based on the sparse non-negative matrix underestimation (SNMU) model (Gillis and Plemmons, 2013). The SNMU can be described as a method that finds a representation of the data in a lower dimension. The SNMU solution approximates the non-negative input matrix (i.e. the exposure matrix E) by two non-negative matrices (U and V) with lower dimension k, such that the product of the two is as close as possible to the original input matrix (Figure 1). k represents the pre-set number of mixtures. The matrix U contains weights (SNMU weight) of pesticides per mixture, the matrix V contains the coefficients of the presence of the mixture per individual or exposure day, and Ɛ is the matrix of residuals due to the approximation. The matrices U, V and Ɛ were obtained by minimising the criterion: $\|E - UV\|^2$ such that $U \geq 0$ and $V \geq 0$.

Figure 1: SNMU decomposition of exposure data. The exposure matrix *E* with dimensions *s* (number of pesticides) and *n* (number of individuals for chronic or exposure days for acute exposure) is approximated by matrix *U* and *V* with dimensions (*s x k*) and (*k x n*) respectively, where *k* represents the number of mixtures.

The non-zero entries in each column of U indicate the components of the selected mixtures. The higher the SNMU weight, the higher the participation of the substance to the mixture. In a technical sense, a mixture, as defined from the non-zero elements of a column of matrix U, could be composed of just one substance. In order to avoid solutions with only or mostly single-substance 'mixtures', the method was adapted by first using the maximum cumulative ratio (MCR, Price and Han (2011)) to restrict the columns of E to only cases where mixtures are important, in order to focus on the individuals (or the individual-days for acute cases) with exposure profiles composed of multiple substances. The MCR is defined as the ratio of the cumulative exposure received by an individual to the largest exposure contribution from a single compound:

MCR = cumulative exposure/maximum exposure from a single compound

If the MCR is large, it is important to consider cumulative effects, if the MCR is close to 1, the individual exposure (or individual-days) will not differ extensively from a single-compound assessment. Only individuals (or individual-days) with an MCR above a chosen threshold were used for the SNMU mixture selection. It was decided to work on the 5% exposures with the highest MCR values. The SMNU and MCR methods were implemented in MCRA software.

3. Results

Selection of pesticide mixtures was carried out for each of the nine countries following the four exposure scenarios and considering at most three mixtures (k=3). For acute exposure, it was necessary to select highly co-exposed individuals. For chronic exposure, the three mixtures explained between 95% and 100% of the total variance in each of the countries and exposure scenarios. For acute exposure, the variance explained by the three mixtures ranged between 41% and 75%. Irrespective of the exposure scenario and the country, the first mixture was the one that explained the higher percentage of variance: at least 55.1% for the chronic scenarios, and 16.2% for the acute scenario. Results are detailed below for this first main mixture.

3.1. Mixture components across the scenarios and countries

Looking at all countries, the main pesticides in the first selected mixture that contributed to population exposure were similar across scenarios (Table 2). In particular, seven compounds were observed in almost all scenarios: imazalil, dithiocarbamates, carbendazim and benomyl, cypermethrin, thiacloprid and deltamethrin, and triadimefon and triadimenol. Among these compounds, two pesticides, imazalil and dithiocarbamates, were observed in almost all countries and contributed the most to the mixture in comparison to the other substances. For the first scenario (adult, chronic, merged data), imazalil and dithiocarbamates were observed with an SNMU weight of 85% and 13% for Belgium and the Netherlands, 72% and 23% for Denmark, and 72% and 24% for France, respectively. Imazalil and dithiocarbamates were also observed as major components for the scenario in "children, chronic, merged data". Regarding the scenarios with country-specific data, imazalil was found to be the main pesticide, followed by dithiocarbamates for Belgium, Denmark, France, and the Netherlands.

The seven compounds with the highest participation to the mixture were confirmed by high contributions of these substances to the total exposure (Figure 2). Imazalil contributed most to the mixture for all countries and scenarios, and may lead to 75% of the total exposure for the adult population with chronic exposure and merged concentration data in the Czech Republic. In fact, regarding exposure levels, imazalil was the compound with the highest exposure levels. The highest value of P95 exposure to imazalil was observed for the Netherlands, in the scenario on chronic exposure in adults using country-specific data with a value of 7.25 µg/kg bw/day contributing to 57% of the total exposure. Another high P95 exposure of 7.15 µg/kg bw/day was observed in the paediatric population for Cyprus, which contributed 67% to the total exposure. For dithiocarbamates, the second major contributor to the mixture, the highest values of P95 exposure were also observed for the Netherlands, in the scenario on chronic exposure for adults with specific concentration data at 0.77 µg/kg bw/day, contributing 33% of the total exposure, followed by the P95 exposure of Slovenia and Spain, in the scenario on chronic exposure in adults with merged concentration data (e.g. 0.76 and 0.72 µg/kg bw/day respectively, contributing 48% and 34% of the total exposure).

Greece had slightly different results. Imazalil was not observed in the mixture found for the chronic adult exposure scenario with merged and specific data. The substances that contributed the most to the mixture were dithiocarbamates, with an SNMU weight of 95% and a contribution to total exposure of 56% for merged data in adults, and 90% and 64% for specific data in adults. For the children scenario, dithiocarbamates were in the first position (78%) followed by cypermethrin (9%) and imazalil (6%).

Looking at the different scenarios, the contributions of compounds for the whole population were generally lower for the acute scenario. Thus, except for Greece, where imazalil highly contributed with a SNMU weight of 92% and a contribution to total exposure of 19%, imazalil contributed less to the mixture in acute exposure. Furthermore, the SNMU weights of triadimefon and triadimenol were significantly higher in the mixture with acute exposure and reached an SNMU weight of 42% in Slovenia.

Some compounds were observed only in one scenario for Greece and the Czech Republic. Abamectin and ethoprophos were observed in Greece only for the chronic scenario with specific national concentration data in the adult population, but the contribution of these compounds to the mixture was relatively low (e.g. SNMU weight of 1%). Furthermore, for

this country, in the child population, the mixture contained metalaxyl and metalaxyl-M, which were only observed in this case (e.g. SNMU weight of 1%). The compound flufenoxuron was observed in Greece only for chronic exposure in the adult population, with country-specific concentration data and in the paediatric population with merged data. For the Czech Republic, fluazinam was observed only in the chronic adult exposure scenario with specific concentration data (e.g. SNMU weight of 0.05%) and iprodione in the acute exposure scenario (SNMU weight of 1%). These compounds in combination contribute less than 10% to total exposure.

Concerning other mixtures and considering the first scenario (adult chronic and merged data) for France, Spain and Greece, mixtures 2 and 3 were composed of the same 7 compounds found for the first mixture but with a different order of importance. For example, in France mixture 2 compounds and their SNMU weights were: dithiocarbamates (93%), cypermethrin (3%), carbendazim and benomyl (2%), triadimefon and triadimenol (1%), deltamethrin (1%), thiacloprid (1%). The last two compounds were not present in the first French mixture. Imazalil was not present in the second mixture but found alone (SNMU weight of 100%) in the third mixture.

For Denmark, the Netherlands, the United Kingdom, Slovenia and the Czech Republic considering the first scenario (adult chronic merged data), new compounds were found in addition to those found in the first mixture. Their SNMU weights were equal to 1% each: dicofol, acetamiprid, iprodione, tebuconazole, fenbuconazole, flufenoxuron, deltamethrin, dithiocarbamates, fipronil, and iprovalicarb. Similar results were found for the other scenarios.

3.2 Contribution of food pesticides to the total population exposure

Table 3 shows the proportion of the different food/pesticide combinations where the SNMU weight of the mixture was relatively high (higher than 5%) contributing the most to the mixture in chronic and acute cases. Imazalil and dithiocarbamates are the major compounds found in food for both chronic and acute exposure. For chronic exposure, imazalil was mainly recorded in oranges and grapefruits in many countries, and at a lower level in mandarins for Belgium and the Czech Republic. For acute exposure, imazalil was also mainly observed in oranges, mandarins, grapefruit, but also in bananas, lemons, limes and pears.

However, dithiocarbamates were not observed in the same foods following the different exposure scenarios. For chronic exposure, dithiocarbamates were mainly observed in cultivated mushrooms in Belgium, the Czech Republic, France, the Netherlands, Spain and the United Kingdom, but not in Greece where cucumbers formed an important part of exposure (e.g. 21.4% of the total measurements), and wine grapes for the Czech Republic (e.g. 4.2%). For acute exposure, dithiocarbamates were mainly observed in lettuce, apples, wine grapes, tomatoes, and pears in several countries, but also in cucumbers for Greece (e.g. 15.9%).

A high contribution of triadimefon and triadimenol to exposure was also recorded for pineapples for acute exposure and especially in Spain (e.g. 12.5% in acute exposure).

Cypermethrin was mainly recorded in wheat in many countries, but in Greece, cocoa (fermented beans) was the main source of exposure to cypermethrin (e.g. 19.9%).

Figure 2. Cumulative contribution (%) of the different substances in each country for the four scenarios: 1. Adults, chronic exposure and merged concentration data; 2. Adults, chronic exposure and specific concentration data; 3. Adults, acute exposure and merged concentration data; 4. Children, chronic exposure and merged concentration data.

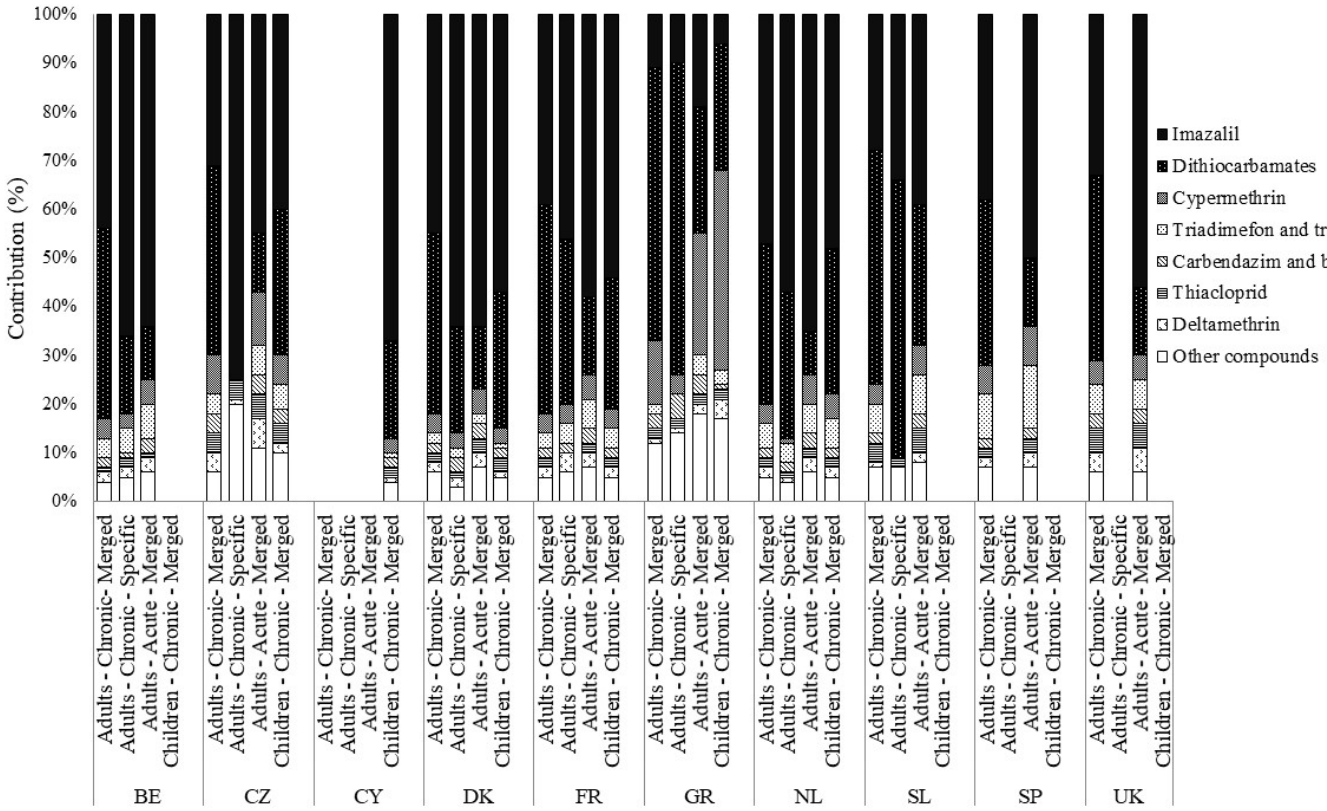

Figure 2. Cumulative contribution (%) of the different substances in each country for the four scenarios: 1. Adults, chronic exposure and merged concentration data; 2. Adults, chronic exposure and specific concentration data; 3. Adults, acute exposure and merged concentration data; 4. Children, chronic exposure and merged concentration data.

Table 1. Description of consumption and concentration data for nine different European countries. n total = number of individuals in the overall consumption survey, n=number of individuals included in this study (adults 18–64 years old, children 11–15 years old), N= number of substances in steatosis CAG after matching with contamination data; the number in brackets indicates the number of substances with measurements ≥ LOD. No national monitoring data was available for Spain (SP) and the United Kingdom (UK)

Country	Consumption survey					Consumption data used for the study			National concentration survey			
	Method	Years	Name	Population	n total	Mean age (min / max values)	Weight mean (min /max values)	n	Years	N	Number of measurements	Percent of measurements ≥ LOD in total measurements
Belgium (BE)	2 x 24 h recall	2004	Diet_National _2004	Adults (14-105 years)	3214	40 (18-64)	71.4 (39-133)	1356	2011-2014	115 (39)	393 967	0.87
Cyprus (CY)	3-d record	2003	Child Health	Children (11-15 years)	303	12.8 (11-15)	54.0 (27-144)	303	2011-2014	81 (48)	124 599	0.72%
Czech Republic (CZ)	2 x 24 h recall	2003-2004	SISP04	Adults (18-64 years)	1666	43.0 (18-64)	75.8 (43-183)	1666	2011-2014	42 (37)	153 696	1.35%
				Children (11-14 years)	109	12.3 (11-14)	46.1 (27-83)	109				
Denmark (DK)	7-d record	2003–2008	DANSDA 2005-08	Adults (18-79 years)	1990	43.0 (18-64)	75.8 (43-183)	1710	2011-2014	95 (58)	503 879	0.62%
				Children (4-17 years)	710	12.7 (11-15)	52.3 (28-100)	234				
France (FR)	7-d record	2005–2007	INCA2	Adults (18-79 years)	2624	40.6 (18-64)	70.6 (35-171)	2276	2010-2014	120 (70)	907 565	0.53%
				Children (3-17 years)	1455	13.1 (11-15)	49.5 (25-128)	585				
Greece (GR)	3-d record	1988-2004	Regional Crete	Adults (18-94 years)	1640	33.2 (18-64)	72.9 (40-141)	1585	2010-2014	91 (56)	324 561	0.65%
				Children (11-15 years)	528	13.4 (11-15)	55.3 (26-109)	528				
Netherland (NL)	2 x 24 h recall	2007-2010	VCP-Basic	Adults (18-69 years)	2230	41.5 (18-64)	80.3 (39-192)	2056	2010-2013	110 (67)	643 538	0.89%
				Children (7-17 years)	1589	13.0 (11-15)	52.5 (27-116)	727				
		2005-2006	VCP-Kids	Children (2-6 years)	1279							
Slovenia (SI)	24 h recall	2007-2008	CRP 2008	Adults (18-65 years)	407	41.4 (18-64)	74.5 (44-125)	400	2012 - 2014	87 (40)	109 810	0.49%
Spain (SP)	3-d record	2011	Encuesta ENIDE	Adults (18-71 years)	3386	39.4 (18-64)	68.5 (41-140)	3371				
United Kingdom (UK)	7-d record	2000-2001	NDNS	Adults (19-64 years)	1724	40.6 (19-64)	76.3 (39-200)	1724				

Table 2. Characteristics of the exposure estimates (mean, median, P5 and P95 in µg/kg bw/day), SNMU weights and contributions to the total exposure for the main mixture following the four scenarios in each country.

	Name compound	RPF	Belgium (BE)						Czech Republic (CZ)						Cyprus (CY)					
			SNMU weight	Contrib.	Mean	Median	P5	P95	SNMU weight	Contrib.	Mean	Median	P5	P95	SNMU weight	Contrib.	Mean	Median	P5	P95
Scenario 1 (Adults, chronic, merged)			1356 individuals. Variance: 75.6%						1666 individuals. Variance: 63.7%.											
	Imazalil	0.13	85%	44%	0.98	0.22	0	3.80	65%	31%	0.41	0.09	0.002	1.1						
	Dithiocarbamates	0.53	13%	39%	0.22	0.17	0.02	0.53	25%	39%	0.13	0.09	0.016	0.35						
	Carbendazim and benomyl	0.2	1%	2%	0.03	0.02	0.002	0.10	2%	4%	0.03	0.02	0.003	0.08						
	Cypermethrin	0.28	1%	4%	0.04	0.03	0.01	0.09	3%	8%	0.05	0.04	0.013	0.12						
	Triadimefon and triadimenol	0.59							2%	4%	0.01	0.002	0	0.06						
	Thiacloprid	0.44							2%	4%	0.01	0.005	0.001	0.06						
	Deltamethrin	0.53							1%	4%	0.01	0.008	0.002	0.04						
Scenario 2 (Adults, chronic, specific)			1356 individuals. Variance: 95.9%						756 individuals. Variance: 99.3%.											
	Imazalil	0.13	91%	66%	1.54	0.27	0	5.91	99%	75%	0.25	0.05	0.001	1.11						
	Dithiocarbamates	0.53	9%	16%	0.09	0.07	0.01	0.24												
	Carbendazim and benomyl	0.2																		
	Cypermethrin	0.28																		
	Thiacloprid	0.44							0.5%	4%	0.004	0.002	0	0.01						
	Abamectin	2.1																		
	Deltamethrin	0.53																		
	Ethoprophos	21																		
	Fluazinam	0.13							0.5%	5%	0.02	0.007	0	0.06						
	Flufenoxuron	2.3																		
	Triadimefon and triadimenol	0.59																		
Scenario 3 (Adults, acute, merged)			2445 exposure days. Variance: 35.5%. MCR cut-off at 5% of co-exposed population						1629 exposure days. Variance: 60.5%. MCR cut-off at 5% of co-exposed population											
	Imazalil	0.13	54%	64%	0.84	0.007	0	5.09	46%	22%	0.07	0.007	0	0.68						
	Dithiocarbamates	0.53							12%	12%	0.02	0.003	0	0.1506						

	Triadimefon and triadimenol	0.59	38%	7%	0.02	0	0	0.05	20%	13%	0.02	0.002	0	0.12								
	Cypermethrin	0.28	5%	5%	0.03	0	0	0.11	3%	8%	0.01	0.001	0	0.07								
	Carbendazim and benomyl	0.2	3%	3%	0.03	0	0	0.09	4%	7%	0.02	0.001	0	0.07								
	Thiacloprid	0.44							12%	11%	0.01	0.001	0	0.11								
	Tebuconazole	0.09							1%	2%	0.01	0.002	0	0.09								
	Deltamethrin	0.53							1%	3%	0.003	0	0	0.01								
	Iprodione	0.005							1%	1%	0.14	0.02	0	0.87								
									109 individuals. Variance: 83.8%						303 individuals. Variance: 96.9%							
Scenario 4 (Children, chronic, merged)	**Imazalil**	0.13							75%	40%	0.9	0.36	0.002	1.3	88%	67%	2.43	1.78	0.001	7.15		
	Dithiocarbamates	0.53							16%	30%	0.17	0.14	0.034	0.41	11%	20%	0.18	0.15	0.045	0.39		
	Cypermethrin	0.28							2%	6%	0.07	0.06	0.021	0.16	1%	3%	0.05	0.04	0.014	0.1		
	Thiacloprid	0.44							2%	4%	0.03	0.01	0.001	0.12								
	Carbendazim and benomyl	0.2							1%	3%	0.04	0.03	0.004	0.11								
	Triadimefon and triadimenol	0.59							2%	5%	0.02	0.01	0.0004	0.1								
	Metalaxyl and metalaxyl-M	0.06																				
	Deltamethrin	0.53																				
	Flufenoxuron	2.3																				

Table 2. Continuation of the table.

	Name compound	RPF	Denmark (DK)						France (FR)						Greece (GR)					
			SNMU weight	Contrib.	Mean	Median	P5	P95	SNMU weight	Contrib.	Mean	Median	P5	P95	SNMU weight	Contrib.	Mean	Median	P5	P95
Scenario 1 (Adults, chronic, merged)			1710 individuals. Variance: 83.5%						2276 individuals. Variance: 71.6%						1785 individuals. Variance: 83.0%					
	Imazalil	0.13	72%	45%	1.03	0.67	0.022	3.35	72%	39%	0.87	0.45	0.004	2.97						
	Dithiocarbamates	0.53	23%	37%	0.21	0.19	0.057	0.45	24%	43%	0.24	0.19	0.038	0.57	95%	56%	0.06	0.003	0	0.32
	Carbendazim and benomyl	0.20	1%	2%	0.03	0.03	0.007	0.07	1%	2%	0.03	0.02	0.004	0.08	2%	3%	0.007	0.002	0	0.03
	Cypermethrin	0.28	2%	4%	0.04	0.04	0.014	0.08	2%	4%	0.04	0.03	0.013	0.08	1%	13%	0.03	0.02	0.002	0.07
	Triadimefon and triadimenol	0.59							1%	3%	0.02	0.01	0.001	0.06	1%	2%	0.002	0.001	0	0.01
	Thiacloprid	0.44	1%	2%	0.02	0.01	0.002	0.04												
	Deltamethrin	0.53																		
Scenario 2 (Adults, chronic, specific)			1710 individuals. Variance: 95.8%						2276 individuals. Variance: 77.2%						1585 individuals. Variance: 93.6%					
	Imazalil	0.13	90%	64%	0.79	0.51	0.016	2.53	84%	46%	0.75	0.41	0.003	2.53						
	Dithiocarbamates	0.53	9%	22%	0.07	0.06	0.015	0.16	12%	34%	0.14	0.11	0.008	0.35	90%	64%	0.07	0.003	0	0.42
	Carbendazim and benomyl	0.20	1%	3%	0.02	0.02	0.005	0.05	1%	2%	0.02	0.02	0.003	0.05	2%	5%	0.01	0	0	0.07
	Cypermethrin	0.28							1%	4%	0.03	0.03	0.01	0.06	2%	4%	0.007	0	0	0.04
	Thiacloprid	0.44																		
	Abamectin	2.10													1%	2%	0.0004	0	0	0.004
	Deltamethrin	0.53							1%	4%	0.02	0.01	0.001	0.05						
	Ethoprophos	21.00													2%	3%	0.001	0	0	0.001
	Fluazinam	0.13																		
	Flufenoxuron	2.30													3%	4%	0.001	0.002	0	0.006
	Triadimefon and triadimenol	0.59							1%	4%	0.01	0.006	0	0.05						

			8917 exposure days. Variance: 33.8%. MCR cut-off at 5% of co-exposed population						9451 exposure days. Variance: 42.1%. MCR cut-off at 5% of co-exposed population						4635 exposure days. Variance: 16.2%. MCR cut-off at 5% of co-exposed population					
Scenario 3 (Adults, acute, merged)	**Imazalil**	0.13	37%	9%	0.03	0.008	0.001	0.06	49%	26%	0.15	0.006	0	0.78	92%	19%	0.05	0	0	0.08
	Dithiocarbamates	0.53	42%	3%	0.005	0.004	0.002	0.007	51%	14%	0.02	0	0	0.08						
	Triadimefon and triadimenol	0.59	3%	10%	0.01	0.004	0	0.04												
	Cypermethrin	0.28													5%	25%	0.03	0.01	0	0.09
	Carbendazim and benomyl	0.20																		
	Thiacloprid	0.44	18%	9%	0.006	0.002	0	0.04							1%	2%	0.002	0	0	0.001
	Tebuconazole	0.09																		
	Deltamethrin	0.53																		
	Iprodione	0.005																		
			234 individuals. Variance: 95.3%						585 individuals. Variance: 88%						5328 individuals. Variance : 56.3%					
Scenario 4 (Children, chronic, merged)	**Imazalil**	0.13	79%	57%	1.71	1.07	0.053	4.84	84%	54%	1.5	0.91	0.012	4.69	6%	6%	0.01	0	0	0.08
	Dithiocarbamates	0.53	17%	28%	0.21	0.18	0.060	0.39	13%	27%	0.19	0.15	0.025	0.47	78%	26%	0.01	0.001	0	0.07
	Cypermethrin	0.28	1%	3%	0.04	0.04	0.017	0.09	1%	4%	0.05	0.04	0.014	0.1	9%	41%	0.03	0.03	0.01	0.07
	Thiacloprid	0.44	1%	3%	0.03	0.02	0.004	0.09							3%	2%	0.001	0	0	0.001
	Carbendazim and benomyl	0.20	0.6%	2%	0.03	0.03	0.007	0.07	1%	2%	0.03	0.02	0.004	0.1	2%	1%	0.002	0	0	0.01
	Triadimefon and triadimenol	0.59							1%	4%	0.03	0.01	0.001	0.11						
	Metalaxyl and metalaxyl-M	0.06													1%	10%	0.04	0.03	0.01	0.08
	Deltamethrin	0.53													1%	4%	0.002	0.001	0	0.005
	Flufenoxuron	2.30													1%	1%	0.0001	0	0	0.001

Table 2. Continuation of the table.

	Name compound	RPF	Netherlands (NL)						Slovenia (SL)						Spain (SP)					
			SNMU weight	Contrib.	Mean	Median	P5	P95	SNMU weight	Contrib.	Mean	Median	P5	P95	SNMU weight	Contrib.	Mean	Median	P5	P95
Scenario 1 (Adults, chronic, merged)			2056 individuals. Variance: 79.8%						400 individuals. Variance: 61.8%						3371 individuals. Variance: 55.1%					
	Imazalil	0.13	85%	47%	0.99	0.31	0	4.04	82%	28%	0.76	0.1	0.0002	3.43	78%	38%	1.26	0.69	0.002	4.3
	Dithiocarbamates	0.53	13%	33%	0.18	0.14	0.02	0.46	15%	48%	0.33	0.29	0.03	0.76	19%	34%	0.29	0.23	0.04	0.72
	Carbendazim and benomyl	0.2	1%	2%	0.03	0.02	0.003	0.08	1%	2%	0.04	0.02	0.002	0.12	1%	2%	0.04	0.03	0.006	0.1
	Cypermethrin	0.28	1%	4%	0.04	0.03	0.02	0.09	1%	4%	0.05	0.04	0.01	0.13	2%	6%	0.1	0.07	0.014	0.26
	Triadimefon and triadimenol	0.59							1%	6%	0.04	0.004	0.0005	0.16						
	Thiacloprid	0.44							1%	4%	0.03	0.01	0.001	0.11						
	Deltamethrin	0.53																		
Scenario 2 (Adults, chronic, specific)			2056 individuals. Variance: 87.9%						400 individuals. Variance: 66.9%											
	Imazalil	0.13	85%	57%	1.74	0.48	0	7.25	99%	34%	0.54	0.06	0	2.32						
	Dithiocarbamates	0.53	13%	30%	0.23	0.16	0.005	0.77												
	Carbendazim and benomyl	0.2																		
	Cypermethrin	0.28																		
	Thiacloprid	0.44							1%	2%	0.01	0.008	0	0.03						
	Abamectin	2.1																		
	Deltamethrin	0.53																		
	Ethoprophos	21																		
	Fluazinam	0.13																		
	Flufenoxuron	2.3																		
	Triadimefon and triadimenol	0.59																		
Scenario 3 (Adults, acute, merged)			2878 exposure days. Variance: 45%. MCR cut-off at 5% of co-exposed population						1576 exposure days. Variance: 44.8%. MCR cut-off at 5% of co-exposed population						3367 exposure days. Variance: 38.7%. MCR cut-off at 5% of co-exposed population					
	Imazalil	0.13	45%	17%	0.04	0.004	0	0.24	44%	22%	0.18	0.005	0	0.90	75%	22%	0.17	0.002	0	0.092

	Dithiocarbamates	0.53	8%	10%	0.006	0	0	0.03	6%	12%	0.02	0	0	0.14	6%	11%	0.02	0	0	0.13
	Triadimefon and triadimenol	0.59	34%	11%	0.006	0	0	0.03	42%	15%	0.03	0	0	0.16	8%	11%	0.02	0	0	0.11
	Cypermethrin	0.28	5%	11%	0.01	0.002	0	0.05	1%	9%	0.03	0.003	0	0.19	2%	9%	0.03	0.001	0	0.18
	Carbendazim and benomyl	0.2	2%	9%	0.01	0.001	0	0.06	2%	7%	0.04	0.002	0	0.18	2%	7%	0.04	0.001	0	0.16
	Thiacloprid	0.44							6%	13%	0.03	0	0	0.17	5%	12%	0.03	0	0	0.16
	Tebuconazole	0.09	1%	3%	0.01	0.001	0	0.04							1%	4%	0.04	0.001	0	0.16
	Deltamethrin	0.53																		
	Iprodione	0.005																		
			727 individuals. Variance: 84.2%																	
	Imazalil	0.13	87%	48%	0.95	0.15	0.001	4.35												
	Dithiocarbamates	0.53	11%	30%	0.15	0.12	0.02	0.39												
	Cypermethrin	0.28	1%	5%	0.04	0.04	0.01	0.09												
Scenario 4 (Children, chronic, merged)	**Thiacloprid**	0.44																		
	Carbendazim and benomyl	0.2																		
	Triadimefon and triadimenol	0.59																		
	Metalaxyl and metalaxyl-M	0.06																		
	Deltamethrin	0.53																		
	Flufenoxuron	2.3																		

Table 2. Continuation of the table.

	Name compound	RPF	United Kingdom					
			SNMU weight	Contrib.	Mean	Median	P5	P95
			1724 individuals. Variance: 71.6%					
Scenario 1 (Adults, chronic, merged)	Imazalil	0.13	76%	33%	0.77	0.29	0.003	2.98
	Dithiocarbamates	0.53	20%	38%	0.19	0.16	0.022	0.48
	Carbendazim and benomyl	0.2	1%	3%	0.03	0.02	0.003	0.07
	Cypermethrin	0.28	1%	5%	0.04	0.03	0.01	0.08
	Triadimefon and triadimenol	0.59						
	Thiacloprid	0.44						
	Deltamethrin	0.53						
Scenario 2 (Adults, chronic, specific)	Imazalil	0.13						
	Dithiocarbamates	0.53						
	Carbendazim and benomyl	0.2						
	Cypermethrin	0.28						
	Thiacloprid	0.44						
	Abamectin	2.1						
	Deltamethrin	0.53						
	Ethoprophos	21						
	Fluazinam	0.13						
	Flufenoxuron	2.3						
	Triadimefon and triadimenol	0.59						

			10767 exposure days. Variance: 37.5%. MCR cut-off at 5% of co-exposed population					
Scenario 3 (Adults, acute, merged)	**Imazalil**	0.13	85%	25%	0.10	0.005	0	0.42
	Dithiocarbamates	0.53	9%	13%	0.01	0	0	0.05
	Triadimefon and triadimenol	0.59	2%	11%	0.01	0	0	0.03
	Cypermethrin	0.28	1%	9%	0.02	0.001	0	0.08
	Carbendazim and benomyl	0.2	2%	7%	0.02	0	0	0.08
	Thiacloprid	0.44						
	Tebuconazole	0.09						
	Deltamethrin	0.53						
	Iprodione	0.005						
Scenario 4 (Children, chronic, merged)	**Imazalil**	0.13						
	Dithiocarbamates	0.53						
	Cypermethrin	0.28						
	Thiacloprid	0.44						
	Carbendazim and benomyl	0.2						
	Triadimefon and triadimenol	0.59						
	Metalaxyl and metalaxyl-M	0.06						
	Deltamethrin	0.53						
	Flufenoxuron	2.3						

Table 3. Contribution of major compounds in specific foods to cumulative exposure for all individual-days which contribute the most to the mixture exposure (at least 5%) for the adult population (18-64 years) with merged data concentration in the case of chronic and acute exposure.

Name compound	Food composition	Belgium (BE)	Czech Republic (CZ)	Denmark (DK)	France (FR)	Greece (GR)	Netherlands (NL)	Slovenia (SI)	Spain (SP)	United Kingdom (UK)
Chronic exposure										
Imazalil	Oranges	31.5%		5.3%				5.4%		23.7%
	Grapefruit	3.0%				1.9%	1.1%	1.1%		
	Mandarins	4.7%	0.4%							
Dithiocarbamates	Cultivated mushrooms	12.6%	3.6%		9.7%				6.6%	10.2%
	Cucumbers					21.4%				
	Wine grapes	7.2%	4.2%							
	Lettuce	4.3%								
	Apple	1.9%								1.1%
Acute exposure										
Imazalil	Oranges	43.9%	19.8%	45.7%	40.8%	3.2%	49.6%	16.8%	36.1%	38.9%
	Mandarins	7.3%	9.0%	11.4%	6.8%		6.4%	7.2%	5.6%	4.5%
	Grapefruit	4.4%	4.3%	1.2%	2.7%	4.9%	4.3%	5.1%	2.3%	3.9%
	Bananas	1.8%	3.7%	3.9%	1.3%		1.9%	2.1%	1.8%	3.4%
	Lemon	3.8%	6.5%		4.2%	8.4%		4.6%	1.9%	2.3%
	Limes		1.3%					2.8%		
	Pear					1.9%				
Dithiocarbamates	Lettuce	3.2%	1.3%	4.2%	7.6%		2.1%	21.3%	8.9%	6.5%
	Apple	1.0%	2.0%	2.4%	0.8%			1.2%		
	Wine grapes	1.2%	1.1%	1.5%	1.5%	4.20%				
	Tomatoes	0.7%	1.0%	0.9%	0.8%				0.7%	
	Currants (red, black, white)		0.9%					0.8%		1.0%
	Pear			1.0%		2.8%				
	Cucumbers					15.9%				
Triadimefon and triadimenol	Pineapple	6.4%	5.0%	1.4%	4.6%	1.7%	5.1%	6.8%	12.5%	4.6%
	Cucumbers					1.3%				
Cypermethrin	Wheat (spelt, triticale)	1.7%	3.2%	0.7%	1.4%		1.3%	1.8%	4.4%	1.1%
	Table grapes		0.5%					2.0%		
	Barley		2.6%	0.6%						1.1%
	Cocoa (fermented beans)					19.9%				
	Cherries								1.5%	
Thiacloprid	Currants (red, black, white)		2.3%	1.3%	0.9%			3.4%		3.4%

4. Discussion

The proposed approach in combining exposure levels with CAG grouping makes it possible to prioritize mixtures from a large range of pesticides. Applying this method to 144 pesticides classified in the steatosis CAG, and following several exposure scenarios for 9 countries, enabled us to prioritize 15 pesticides.

Across the different scenarios and countries, one mixture explained the major part of the total exposure. This mixture is composed of two high contributors which are imazalil and dithiocarbamates. The relative potency factors (RPFs) of the two substances are relatively low compared to the other substances, especially for imazalil. This implies that their presence in the mixture is due to high co-exposures of the population to these pesticides, and thus to high concentrations in consumed foods. In fact, imazalil and dithiocarbamates have one of the highest percentages of quantified values in food (around 7%). Since the same residue concentrations are used in the scenario using the merged dataset, inconsistencies between countries result from variability in food consumption behaviours and/or differences between the designs, the methodology, the time and the size of the consumption surveys. For most countries, the principal mixtures were similar, leading to the supposition that the design of the surveys had not a significant impact on mixture selection. The difference with Greece mixture came from the fact that cucumbers are the main drivers of dithiocarbamates intake whereas in other countries the presence of imazalil and dithiocarbamates were due to the consumption of fruits and mushrooms. During the last years, EFSA tended to harmonize the design and the food coding used in the food consumption surveys between the Member States of the European Union (EFSA, 2014). For example in France, the dietary collection method was changed from the 7-consecutive-day food record previously used in the Individual and National food consumption surveys (INCA) to 3-non-consecutive day of 24-h dietary recall, completed by a food propensity questionnaire for the INCA3 survey. So in future, comparison of mixtures between European countries would be less impacted by methodological issues related to food consumption survey design.

Scenarios with country-specific data lead to similar mixtures with fewer components compared to the one with the merged dataset, which could be due to data gaps. These results support the idea that using a merged dataset to estimate European exposures seems to be realistic as foods are traded between European countries. Moreover, using merged datasets makes it possible to fill data gaps for countries with lower numbers of analyses. However, using merging datasets with different analytical methodologies and not weighted for representativeness may introduce uncertainties in concentration. This uncertainty could be reduced in future works in considering information provided in the SSD1 format regarding analytical methodology, the subsequent quality assurance measures and the coverage of sampled regions. Efforts must continue to harmonize and to combine data at the European level for different parts of the pesticide regulatory framework to improve efficiency. For example, there is a difference between pesticide residue definitions for enforcement (usually those present in concentration databases), residue definitions for risk assessment, and the substances in the CAG list, which are usually parent compounds. For example, dithiocarbamates comprise all substances measured as carbon disulfide, including maneb, mancozeb, metiram, propineb, thiram and ziram, whereas ziram is the only substance in the CAG. To combine both databases, conversion factors should be applied to obtain the concentration of the residue definition for risk assessment of the parent compound. Such conversion factors are described for example in EFSA and Joint FAO/WHO Meeting on Pesticide Residues (JMPR) opinions, but no harmonized database is available. Moreover, as different conversion factors may occur for product-pesticide combinations, this would result in many concentration conversions to be manually performed, which requires significant resources. As a pragmatic approach, the conversion factors were set to 1, but may have led to an underestimation or overestimation of exposure. A harmonized database with conversion factors or concentration data with a focus on individual compounds would be helpful for future calculations. Another point that impacts exposure is the time lag of concentration data upon regulatory changes such as new authorizations and bans. Thus, concentration data are missing for new pesticides, whereas exposures could be overestimated for banned pesticides. Moreover, currently, processing factors are not available for all pesticide/food/process combinations. In addition, extrapolation of processing factors (e.g. a processing factor available for peeling of mandarins used for peeling of lemons) is not common practice. This may lead to an overestimation in cases where processing lowers the pesticide concentration, e.g. peeling and juicing, or an underestimation in cases where processing increases the concentration (drying of fruit, making tomato paste). This is for example the case of imazalil which was mainly found in oranges, grapefruits, mandarins for which no processing factor for peeling was available. It was also found in lemons for which processing factors of juicing, washing and oiling were applied. More research is needed to either develop new processing factors or to extrapolate processing factors between food items. Matching processing factors as provided in the BfR database to the foods measured in the concentration database and to foods in the food conversion table was a laborious process. A harmonised table with processing factors linked to harmonise coding of SSD1 would facilitate mixture selection. Another solution, which reduces uncertainty, is to measure concentrations directly in food as consumed, as is the case in total diet studies (Sirot et al., 2009). Running chronic exposure scenarios for adults for France and Netherlands did not affect the main composition of the

mixtures. There is also a need to collect information on substances other than pesticides. We decided to focus on pesticides in this study because these are the substances for which there are the most data regarding concentration values and CAG information. However, other substances present in food such as dioxins, polychlorinated biphenyls, bromated compounds, etc. could have a steatosis effect. This could lead to an underestimation of the total risk related to this CAG. The originality of the proposed approach is to combine information for hazard for a CAG with that on combined exposure to define mixture. Under the assumption of dose-addition, the RPFs make it possible to convert the exposure of all substances into the "unit toxicity" of the index compound. Although there is a consensus that in most cases, dose addition is the best conservative effect estimation for chemicals with exposure at low doses (Backhaus and Faust, 2012; EFSA, 2013a; Kamo and Yokomizo, 2015; Kortenkamp et al., 2009). In some cases, for examples for chemicals with dissimilar modes of action, this hypothesis could lead to underestimate mixture effect (Altenburger et al., 2013; Borgert et al., 2012; Gregorio et al., 2013). In the absence of detailed information, EFSA CAGs are currently defined on the basis of specific effects and not on their mechanism or mode of action. Thus, there is uncertainty on the membership of a pesticide in a CAG and on the validity of applying dose addition. Specific work related to hazard uncertainty is in progress in the Euromix project to analyse the impact of CAG membership on cumulative risk assessment. A probability is attributed to each substance in the CAG, and integrated in calculations. Moreover, RPF values are estimated from NOAELs or LOAELs sourced from bibliographic data. The BMD approach was not applied because several details on quantitative data were not or only partially available from the databases (e.g.: end-points incidences in each dose-groups, number of animals in each dose groups, etc.). There is a high level of uncertainty around the NOAEL and LOAEL values due to the diversity of the surveys from which they were collected. Thus, survey design, species, and duration of treatment could be different, and lead to different level of uncertainty and to results that are difficult to compare. For liver effects, 100% were repeated dose studies, more than 80% were from long-term studies, and 100% were from *in vivo* studies. Therefore, the liver data package can be considered homogeneous. The extrapolation of NOAELs from LOAEL values for 9% of the substances can also be a source of uncertainty. A ratio of three was used as it is generally used in toxicology studies dose spacing regime an as it was recommended in the first version of the WHO Guidance Uncertainty in Hazard Assessment, the available version at the time we made the calculations. In the second version (WHO, 2018) it is also proposed a ratio of 10 which can be used for future work. There is also a need to define a reference compound to convert toxicity, as none of the CAG lists of the DTU and EFSA indicate such index compounds. The choice of pesticide to serve as a reference compound has mathematically no impact on final results. This could lead to bias if there were high uncertainty on the NOAEL of the reference compound. In the present study, to minimize errors, it was decided to use a well-known compound with high quality criteria as listed under section 2.2.1. Modelling of uncertainty for RPFs remains research to be done in the future. The EuroMix project by developing *in vitro* and *in vivo* strategies for testing mixtures will contribute to greater knowledge on the toxicity of the CAG steatosis compounds. Some of the pesticides prioritized in this work are now being studied for their potency separately and in mixtures to test the dose-addition

assumption. The EuroMix project is also studying two other CAGs on developmental toxicity and endocrine disruptor.

During the last years, statistical developments have been proposed to identify combined exposures of concern through the diet. Crépet and Tressou (2011) used a Bayesian non-parametric model to determine the major mixtures classifying the population regarding their exposure profiles, and then studied correlations between pesticides. More recently, Béchaux et al. (2013) and Traoré et al. (2016) demonstrated the ability of the combination of non-negative matrix factorisation (NMF) (Lee and Seung, 2001) with a hierarchical clustering to identify principal mixtures connected with specific diets. This approach gave close results to the ones obtained with the Bayesian non-parametric model (Béchaux et al., 2013), but was found to produce more interpretable results in terms of mixtures and exposure systems combination using the two matrices U and V. The NMF and clustering methods have also been used to define dietary patterns and clusters of individual diets by Zetlaoui et al. (2011), Sy et al. (2013) and Gazan et al. (2016). In this study, a modified version of the NMF method, called sparse non-negative matrix under-approximation (SNMU) (Gillis and Plemmons, 2013), was used to determine the main mixtures from European exposure data. It was already applied with success in Traoré et al. (2018). This method is also based on the decomposition of the exposure matrix into two submatrices, but used a recursive algorithm which allows us to extract exposure systems one by one. From the original exposure matrix, the first rank one is extracted and therefore subtracted from this matrix. The same procedure is thus applied to the new obtained matrix. Thus, another rank one is extracted corresponding to the first rank for this matrix and to the second rank one for the original exposure matrix. At each step, a rank one is extracted from a new matrix and is identical, regardless of the number of exposure systems. Hence, this algorithm has the advantage that it produces stable mixtures for a selected number of mixtures. Moreover, the NMF and the SNMU are dedicated to positive and null values like exposures comparing to the principal component analyses which could also be used to reduce data dimension and to define mixtures.

As the goal of the approach is to prioritise mixtures to be assessed, the optimistic scenario proposed by EFSA was chosen (EFSA, 2012). This scenario, by considering a zero value for censored concentration data, makes it possible to focus on substances with quantified measurements. This is a way of selecting substances with observed values, and of removing the other substances, before applying the statistical method to extract mixtures. The fact that it is preferable to use a more realistic optimistic scenario to define mixtures was reinforced by the results obtained when using the EFSA pessimistic scenario for France as an example. New substances appeared in the mixture: dazomet, endrin, friponil, ethroprophos. The imazalil disappeared and the dithiocarbamates decreased. However, for dazomet for example no concentration data was available thus the MRL was used. Thus, the variability in the mixture is guided by the LOD and LOQ substitution and/or imputation of maximum residue limits and it is attributed to uncertainty on concentration data. Boon et al. (2015) also found that the pessimistic approach could lead to results far from reality, being dominated by LOD and LOQ substitution and imputation of missing data by MRLs.

As the steatosis effect appears with long-term exposure, it was decided to study chronic exposures. Acute exposure was also considered because repeated acute exposures could lead to chronic effects with time.

The purpose of this study was to identify mixtures that are relevant to study for their combined toxicological effects rather than identifying the main risk drivers. Thus, in the case of a single substance composing a mixture, it was decided to restrict the exposure matrix to

712 the exposure profiles which contain mixtures in using the MCR cut-off. This was the case for
713 all countries for the acute exposure scenario. Focusing on 5% of the population with high
714 combined exposure made it possible to extract mixtures containing several compounds. A test
715 was also done to focus on 30% of the population with high combined exposure, but it
716 produced similar results of a unique substance as for the whole population. It is important to
717 note that acute exposure values are lower than chronic exposure due to the fact that only
718 highly co-exposed individuals were considered. As a result, these individuals are highly co-
719 exposed but with lower doses than other people.
720

721 ## 5. Conclusions

722

723 To conclude, the proposed approach makes it possible to prioritise compounds in a given
724 CAG that need to be further studied. This may include performing further toxicological tests
725 to study modes and mechanisms of action, generating better relative potency factors and,
726 eventually, planning epidemiological surveys. As this approach is sensitive to the input data
727 and demands significant resources, it is important to continue efforts on data collection and
728 harmonisation among the different aspects within the pesticides regulatory framework, and to
729 develop methods to group substances in mixtures and to characterise the risk.

730 ## Acknowledgements

731 The authors would like to thank Ido Toxopeus and Gerda van Donkersgoed from the RIVM
732 for their highly valuable help in preparing the datasets used in this manuscript. This study is
733 part of the EuroMix project (No. 633172) funded in the framework of Horizon 2020 (H2020-
734 SFS-2014-21)

735 ## References

736 AESAN, 2011. ENIDE: Encuesta Nacional de Ingesta Dietética (2009-2010). Resultados sobre datos de
737 consumo.
738 Altenburger, B., Backhaus, T., Boedeker, W., Faust, M., Scholze, M., 2013. Simplifying complexicity:
739 Mixture toxicity assessment in the last 20 years. Environmental Toxicology and Chemistry 32, 1685-
740 1687.
741 Backhaus, T., Faust, M., 2012. Predictive environmental risk assessment of chemicals mixtures: A
742 conceptual framework. Environmental Science and Technology 46, 2564-2573.
743 Béchaux, C., Zetlaoui, M., Tressou, J., Leblanc, J.C., Héraud, F., Crépet, A., 2013. Identification of
744 pesticide mixtures and connection between combined exposure and diet. Food and Chemical
745 Toxicology 59, 191-198.
746 Bertsias, G., Mammas, I., Linardakis, M., Kafatos, A., 2003. Overweight and obesity in relation to
747 cardiovascular disease risk factors among medical students in Crete, Greece. BMC Public Health 3, 3.
748 Boobis, A.R., Ossendorp, B.C., Banasiak, U., Hamey, P.Y., Sebestyen, I., Moretto, A., 2008. Cumulative
749 risk assessment of pesticide residues in food. Toxicology Letters 180, 137-150.

Boon, P.E., van Donkersgoed, G., Christodoulou, D., Crépet, A., D'Addezio, L., Desvignes, V., Ericsson, B.-G., Galimberti, F., Eleni, I.-K., Hamborg Jensen, B., Rehurkova, I., Rety, J., Ruprich, J., Sand, S., Stephenson, C., Strömberg, A., Aida, T., van der Voet, H., Ziegler, P., Hamey, P., van Klaveren, J.D., 2015. Cumulative dietary exposure to a selected group of pesticides of the triazole group in different European countries according to the EFSA guidance on probabilistic modelling. Food and Chemical Toxicology 79, 13-31.

Borgert, C.J., Sargent, E.V., Casella, G., Dietrich, D.R., McCarty, L.S., Golden, R.J., 2012. The human relevant potency threshold: Reducing uncertaintity by human calibration of cumulative risk assessments. Regulatory Toxicology and Pharmacology 62, 313-328.

Crépet, A., Tressou, J., 2011. Bayesian nonparametric model for clustering individual co-exposure to pesticides found in the French diet. Bayesian Analysis 6, 127-144.

Crépet, A., Tressou, J., Graillot, V., Béchaux, C., Pierlot, S., Héraud, F., Leblanc, J.-C., 2013. Identification of the main pesticide residue mixtures to which the French population is exposed. Environmental Research 126, 125-133.

De Vriese, S., Huybrechts, I., Moreau, M., Van Oyen, H., 2005. Belgian Food Consumption Survey. Archives of Public Health 63, 1-16.

Dubuisson, C., Dufour, A., Carrillo, S., Drouillet-Pinard, P., Havard, S., Volatier, J.L., 2018. The Third French Individual and National Food Consumption (INCA3) Survey 2014–2015: method, design and participation rate in the framework of a European harmonization process. . Public Health Nutrition (in press)

Dubuisson, C., Lioret, S., Touvier, M., Dufour, A., Calamassi-Tran, G., Volatier, J.L., 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, 1035-1048.

EFSA, 2007. Scientific colloquium: cumulative risk assessment of pesticides to human health: the way foward, Summary report. 27-28 Novembrer 2006, Parma, Italy, p. 160p.

EFSA, 2008. Scientific opinion of the panel on plant protection products and their residues (PPR Panel) on a request from the EFSA evaluate the suitability of existing methodologies and, if appropriate, the identification of new approaches to assess cumulative and synergistic risks from pesticides to human health with a view to set MRLs for those pesticides in the frame of Regulation (EC) 396/2005. EFSA Journal 704, 84p.

EFSA, 2010. Standard sample description for food and feed. EFSA Journal 8, 54p.

EFSA, 2011. Evaluation of the FoodEx, the food classification system applied to the development of the EFSA Comprehensive European Food Consumption Database. EFSA Journal 9, 27p.

EFSA, 2012. Guidance on the use of probabilistic methodology for modelling dietary exposure to pesticide residues. EFSA Journal 10, 95p.

EFSA, 2013a. Scientifc Opinion on the relevance of dissimilar mode of action and its appropriate application for cumulative risk assessment of pesticides residues in food. EFSA Journal 11, 40p.

EFSA, 2013b. 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, 131p.

EFSA, 2014. Guidance on the EU Menu methodology. EFSA Journal 12, 3944.

Fox, M.A., Brewer, L.E., Martin, L., 2017. An overview of literature topics related to current concepts, methods, tools and applications for cumulative risk assessment (2007-2016). International Journal of Envrionmental Research and Public Health 14, 389.

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

Gillis, N., Plemmons, R.J., 2013. Sparse nonnegative matrix underapproximation and its application to hyperspectral image analysis. Linear Algebra and its Applications 438, 3991-4007.

Gregorio, V., Chèvre, N., Junghans, M., 2013. Critical issues in using the common mixture toxicity models concentration addition or response addition on species sensitivity distributions: A theoretical approach. Environmental Toxicology and Chemistry 32, 2387-2395.

802 Henderson, L., Gregory, J., Swan, G., 2002. The National Diet and Nutrition Survey: adults aged 19 to
803 64 years, Vol. 1: Types and quantities of foods consumed. . The Stationery Office, London.
804 Hercberg, S., Deheeger, M., Preziosi, P., 1994. Portions alimentaires: Manuel photos pour
805 l'estimation des quantités. (Food portions: Picture booklet for the estimation of quantities).
806 PolyTechnica, Paris.
807 Kafatos, A., Kouroumalis, I., C., T., Labadarios, D., 1991. Coronary-heart- disease risk-factor status of
808 the Cretan urban population in the 1980's. . The Amarican Journal of Clinical Nutrition 54, 591-598.
809 Kamo, M., Yokomizo, H., 2015. Explanation of non-additive effects in mixtures of similar mode of
810 action chemicals. Toxicology 335, 20-26.
811 Kortenkamp, A., Backhaus, T., Faust, M., 2009. State of the art report on mixture toxicity. Final report
812 of a project on mixture toxicology and ecotoxicology commissioned by the European Commission, DG
813 Environment.
814 Lee, D.D., Seung, H.S., 2001. Algorithms for non-negative matrix factorization, Advances in Neural
815 Information Processing Systems.
816 Linardakis, M., Bertsias, G., Sarri, K., Papadaki, A., Kafatos, A., 2008. Metabolic syndrome in children
817 and adolescents in Crete, Greece, and association with diet quality and physical fitness. Journal of
818 Public Health 16, 421-428.
819 Lioret, S., Dubuisson, C., Dufour, A., Touvier, M., Calamassi-Tran, G., Maire, B., Volatier, J.L., Lafay, L.,
820 2010. Trends in food intake in French children from 1999 to 2007: Results from the INCA (étude
821 Individuelle Nationale des Consommations Alimentaires) dietary surveys. British Journal of Nutrition
822 103, 585-601.
823 Moschandreas, J., Kafatos, A., 1999. Food and nutrient intakes of Greek (Cretan) adults. Recent data
824 for food-based dietary guidelines in Greece. British Journal of Nutrition 81, S71-76.
825 Nielsen, E., Norhede, P., Boberg, J., Krag Isling, L., Kroghsbo, S., Hadrup, N., Bredsdorff, L.,
826 Mortensen, A., Larsen, J.C., 2012. Identification of cumulative assessment groups of pesticides. EFSA
827 Supporting Publications 9, 303p.
828 Ocké, M.C., van Rossum, C.T.M., Fransen, H.P., Buurma, E.J.M., de Boer, E.J., Brants, H.A.M., Niekerk,
829 E.M., van der Laan, J.D., Drijvers, J.J.M.M., Ghameshlou, Z., 2008. Dutch National Food Consumption
830 Survey - Young children 2005/2006., RIVM report, Bilthoven, National Institue for Public Health and
831 the Environment (RIVM), p. 103p.
832 Pedersen, A.N., Fagt, S., Groth, M.V., Christensen, T., Biltoft-Jensen, A.P., Matthiessen, J., Andersen,
833 N.L., Kørup, K., Hartkopp, H.B., Ygil, K.H., Hinsch, H.-J., Saxholt, E., Trolle, E., 2010. Danskernes
834 kostvaner 2003 - 2008, Hovedresultater (Dietary Habits in Denmark 2003-2008, Main Results),
835 National Food Institute,Technical University of Denmark., p. 196p.
836 Price, P.S., Han, X., 2011. Maximum cumulative ratio (MCR) as a tool for assessing the value of
837 performing a cumulative risk assessment. International Journal of Envrionmental Research and Public
838 Health 8, 2212-2225.
839 RIVM, ICPS, ANSES, 2013. Toxicological data analysis to support grouping of pesticide active
840 substances for cumulative risk assessment of effects on liver, on the nervous system and on
841 reproduction and development, EFSA Supporting Publications, p. 88p.
842 RIVM, ICPS, ANSES, 2016. Toxicological data collection and analysis to support grouping of pesticide
843 active substances for cumulative risk assessment of effects on the nervous system, liver, adrenal,
844 eye, reproduction and development and thyroid system (GP/EFSA/PRAS/2013/02), EFSA Supporting
845 Publications, p. 184p.
846 Ruprich, J., Dofkova, M., Rehurkova, I., Slamenikova, E., Resova, D., 2006. Individual food
847 consumption - the national study SISP04. CHFCH NIPH in Prague.
848 Sirot, V., Volatier, J.L., Calamassi-Tran, G., Dubuisson, C., Ménard, C., Dufour, A., Leblanc, J.C., 2009.
849 Core food of the French food supply: Second Total Diet Study. . Food Additives and Contaminants -
850 Part A Chemistry, Analysis, Control, Exposure and Risk Assessment 26, 623-639.
851 Sy, M.M., Feinberg, M., Verger, P., Barré, T., Clémençon, S., Crépet, A., 2013. New approach for the
852 assessment of cluster diets. Food and Chemical Toxicology 52, 180-187.

Traoré, T., Béchaux, C., Sirot, V., Crépet, A., 2016. To which chemical mixtures is the French population exposed? Mixture identification from the second French Total Diet Study. Food and Chemical Toxicology 98, 179-188.
Traoré, T., Forhan, A., Sirot, V., Kadawathagedara, M., Heude, B., Hulin, M., de Lauzon-Guillain, B., Botton, J., Charles, M.A., Crépet, A., 2018. 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. Food and Chemical Toxicology 111, 310-328.
University of Crete, F.o.M., Departement of preventive Medecine and Nutrition, March 2016. Regional Crete food consumption studies 1988-2004, Unpublished data submitted to BPI for EuroMix project
van Rossum, C.T.M., Fransen, H.P., Verkaik-Kloosterman, J., Buurma - Retahns, E.J.M., Ocké, M.C., 2011. Dutch National Food Consumption Survey 2007-2010, Bilthoven, National Institute for Public Health and the Environment (RIVM). .
WHO, 2009. Assessment of combined exposures to multiple chemicals: report of a WHO/IPCS international workshop on aggregate/cumulative risk assessement. World Health Organization.
WHO, 2018. Guidance document on evaluating and expressing uncertainty in hazard characterization 2nd edition. Geneva: World Helath Oraganization
Xatzis, C., Berskias, J., Scott, J., Linardakis, M., Kafatos, A., 2004. Determination of levels of total homocysteine, folate and vitamin B12 in a healthy adult population in Crete Medecine 85, 293-304.
Zetlaoui, M., Feinberg, M., Verger, P., Clémençon, S., 2011. Extraction of food consumption systems by nonnegative matrix factorization (NMF) for the assessment of food choices. Biometrics 67, 1647-1658.

Concentrations
- ✓ Specific country data
- ✓ Merged data

Consumption
- ✓ 9 European countries

Cumulative assessment groups on steatosis effect (EFSA)

154 pesticides

NOAEL (no observed adverse effect)

RPFs (relative potency factors)

$$\frac{NOAELref}{NOAELp}$$

Combined Exposures

Pesticides

p=1,…,…,....,…S

Individus

i=1,……,……,……,……,……,n

$$E \times RPF$$

between 42 and 120 pesticides regarding countries

p : pesticide

ref : pesticide chosen as reference

Selection criteria
- ✓ High Exposures and Correlations
- ✓ Common effect and high toxicity

Statistical method
- ✓ SNMU (sparse non-negative matrix underaproximation)

Mixtures defined from hazard and co-exposure