



science and policy  
for a healthy future

HORIZON2020 Programme  
Contract No. 733032 HBM4EU

# Report on re-analysis of existing HBM data for priority substances to identify mixtures

## Deliverable Report

### D15.7

### WP15 - Mixtures, HBM and human health risks

**Deadline: February 2022**

**Upload by Coordinator: 29 April 2022**

Entity	Name of person responsible	Short name institution	Date [Received]
Coordinator	Marika Kolossa-Gehring	UBA	21/03/2022
Grant Signatory		RIVM	

Entity	Name of person responsible	Short name institution	Date [Approved]
Pillar Leader	Robert Barouki	INSERM	06/03/2022
Work Package Leader	Mirjam Luijten	RIVM	16/03/2022
Task Leader	Mirjam Luijten	RIVM	16/03/2022

Responsible author	Mirjam Luijten	E-mail	mirjam.luijten@rivm.nl
Short name of institution	RIVM	Phone	+31-30-2743628
Co-authors	Erik Lebret (IRAS/RIVM), Madeline Carsique (ANSES), Eva Govarts (VITO), Miranda Loh (IOM), Miguel Juliá Molina (ISCIII), Ilse Ottenbros (IRAS/RIVM), Laura Rodriguez Martin (VITO), Katarína Řiháčková (RECETOX), Philipp Schmidt (UBA), Jelle Vlaanderen (IRAS), Nina Vogel (UBA), Yvette Christopher-de Vries (IOM), Susana Pedraza (ISCIII), Pavel Cupr (RECETOX), Amélie Crépet (ANSES), Jos Bessems (VITO)		

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebet, Madeline Carsique, Eva Govarts, et al.	Page: 2

## Table of contents

1	Authors and acknowledgements .....	4
2	Glossary .....	5
3	Abstract .....	6
4	Introduction and background.....	8
4.1	Definition of mixtures for the purpose of this deliverable .....	8
4.2	Overall objectives and tasks of WP15.....	9
4.3	Questions underlying the tasks in 15.1 .....	9
4.4	Overall approach .....	10
5	Material and Methods .....	11
5.1	Selection of existing HBM data .....	11
5.2	Data selection and preparation .....	11
5.2.1	Network analysis.....	11
5.2.2	Network analysis with toxicity weighting.....	12
5.3	Existing HBM data .....	12
5.3.1	GerES V (Germany) .....	12
5.3.2	BIOAMBIENT.ES (Spain) .....	13
5.3.3	CELSPAC - FIREexpo (Czech Republic) .....	15
5.3.4	3xG (Belgium).....	16
5.3.5	FLEHS (Belgium).....	18
5.3.6	Toxicity weighting .....	20
5.4	Statistical analysis .....	21
5.4.1	Descriptive analysis .....	21
5.4.2	Network analysis.....	21
5.4.3	SNMU .....	22
5.4.4	Data selection for toxicity weighting .....	23
6	Results.....	25
6.1	Descriptive analysis .....	25
6.1.1	GerES V (Germany) .....	25
6.1.2	BIOAMBIENT.ES (Spain) .....	28
6.1.3	CELSPAC - FIREexpo (Czech Republic) .....	32
6.1.4	3xG (Belgium).....	35
6.1.5	FLEHS data SNMU.....	39
6.2	Network analysis.....	40
6.2.1	Generic example.....	40
6.2.2	GerES V (Germany) .....	40

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebrecht, Madeline Carsique, Eva Govarts, et al.	Page: 3

6.2.3	BIOAMBIENT.ES (Spain) .....	47
6.2.4	CELSPAC - FIREexpo (Czech Republic) .....	54
6.2.5	3xG (Belgium).....	56
6.3	SNMU analysis .....	62
6.3.1	FLEHS results from SNMU analyses .....	62
6.3.2	Discussion and conclusions on SNMU and SNMU – network analysis comparison .	66
6.4	Toxicity weighting .....	67
6.4.1	Results of UBA-IOM data application.....	67
6.4.2	Discussion and conclusions on toxicity weighting .....	69
7	Discussion and conclusions .....	70
7.1	Methodological aspects .....	70
7.1.1	Network analyses .....	70
7.2	Results discussion .....	71
7.2.1	Network analysis.....	71
7.2.2	SNMU .....	71
7.2.3	Toxicity weighting .....	72
7.3	Closing remarks.....	72
7.4	Conclusions .....	73
8	References .....	75
9	Annexes.....	78
9.1	Annex 1: Matrix U from SNMU method on FLEHS III .....	78
9.2	Annex 2: SNMU summary tables .....	79
9.3	Annex 3: List of substances used for toxicity weighting in network analyses .....	83
9.4	Annex 4: Health-Based Human Biomonitoring Guidance Values Description.....	86

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 4

# 1 Authors and acknowledgements

## Lead authors

Mirjam Luijten (RIVM)

Erik Lebret (IRAS/RIVM)

Madeline Carsique (ANSES)

Eva Govarts (VITO)

Miranda Loh (IOM)

Miguel Juliá Molina (ISCIII)

Ilse Ottenbros (IRAS/RIVM)

Laura Rodriguez Martin (VITO)

Katarína Řiháčková (RECETOX)

Phillipp Schmidt (UBA)

Jelle Vlaanderen (IRAS)

Nina Vogel (UBA)

Yvette Christopher-de Vries (IOM)

## Contributors

Susana Pedraza (ISCIII)

Pavel Cupr (RECETOX)

Amélie Crépet (ANSES)

Jos Bessems (VITO)

## Acknowledgements

The authors gratefully acknowledge the constructive contribution of the data owners of the HBM databases used, and the use of their computational facilities.



D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 5

## 2 Glossary

3xG	Gezondheid, Gemeenten, Geboorten (Health, Municipalities, Births)
ADME	Absorption, Distribution, Metabolism, and Excretion
BE	Biomonitoring Equivalent
BMI	Body Mass Index
BP	Benzophenones
CAG	Cumulative Assessment Group
CAN	Comparative Network Analysis
CRT	Creatinine
DINCH	1,2-Cyclohexane dicarboxylic acid diisononyl ester
FLEHS	Flemish Environment and Health Studies
GDPR	General Data Protection Regulation
GerES V	German Environmental Survey for children and adolescents, 2014-2017
HB-HBGV	Human Biomonitoring Health-Based Guidance Value
HBM	Human Biomonitoring
HI	Hazard Index
HQ	Hazard Quotient
ISCED	International Standard Classification of Education
LOD	Level of Detection
LOQ	Level of Quantification
NUTS	Nomenclature of territorial units for statistics
PAHs	Polycyclic Aromatic Hydrocarbons
RPF	Relative Potency Factors
SD	Standard deviation
SNMU	Sparse Non-negative Matrix Underapproximation

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebrecht, Madeline Carsique, Eva Govarts, et al.	Page: 6

### 3 Abstract

This report describes the work done under Task 15.1, which is focused on the ‘differential network analysis’ methodology and its future application for risk assessment of chemical mixtures. Network analyses applied to existing Human Biomonitoring data allow for the identification of exposure patterns, which also include coincidental exposures, and for a subsequent systematic comparison of the patterns observed. Following a successful proof-of-concept, described by Ottenbros and co-workers (2021), this methodology has been applied to datasets from four different countries across Europe, i.e., Germany, Belgium, Spain and the Czech Republic. The four studies involved are quite diverse, both in terms of the exposure biomarkers that have been measured and the population groups involved. For each of the four studies, results of the network analyses are shown, and findings are discussed. Focus was on the dependencies between HBM levels within the same individual; absolute HBM levels were largely ignored here.

Additionally, to network analysis, the usefulness of an alternative statistical method for mixture risk assessment is explored. This method, which involves a combination of the Sparse Non-negative Matrix Underestimation and a clustering algorithm, is explained in more detail in this report. Comparison of the forthcoming results with those obtained from the network analyses enabled the identification of the usefulness of both approaches for risk assessment of chemical mixtures. Finally, an approach to enrich the network analysis methodology with toxicity weighting was explored, with the aim to allow for prioritisation of identified mixtures. To this end, a database of health-based guidance values was developed. Information from that database was combined with the results obtained from the German network analysis to demonstrate proof-of-concept.

In conclusion:

- The co-occurrences of mixtures of substances using HBM data is feasible through network analysis and SNMU, each with its own strength and limitation. Combined application to explore and quantify co-occurrence is recommended.
- The SNMU analysis of the FLEHS data indicate that a substantial part of the variation in the HBM data (> 70 %) can be captured with a limited number of clusters. While this needs to be replicated in other data, there is no reason to believe that this will be very different in other HBM datasets.
- Existing databases of early HBM studies are useful for the first exploration of co-occurrence of substances in the human body.
- Existing studies typically have limited number of individuals in which the full range of chemical substances have been measured. This limits the ability to identify patterns of co-occurrence and even more so to study the role of determinant/covariates, with fewer observations per stratum. For future studies it is recommended to expand the number of observations with full range of chemicals substances, to improve the ability to study mixtures.
- The stability and consistency of identified networks and mixture communities deserves further study, particularly for high dimensional data when strata of covariates are being studied.
- Toxicity weighting of mixture communities/clusters is feasible, but severely limited by the shortage of available HBGV's or other indicators of toxic potency of the substances involved. More generic inroads need to be explored, the more so when suspect screening and untargeted screening is wider applied in HBM studies.

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebet, Madeline Carsique, Eva Govarts, et al.	Page: 7

- So far, four existing datasets were subjected to network analysis techniques, in one dataset network analysis and SNMU were compared, and toxicity weighting was explored in one dataset. The datasets varied widely in study population, study design, matrixes under study and chemicals analysed. It is therefore recommended to expand current analysis to a wider set of existing data and to screen existing samples collected within the HBM4EU WP8 framework and earlier relevant HBM studies on feasibility aspects for re-analysis through suspect screening and untargeted analysis.

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 8

## 4 Introduction and background

### 4.1 Definition of mixtures for the purpose of this deliverable

Humans are exposed to a myriad of concurrent and protracted environmental, occupational, dietary, lifestyle and consumer product exposures. Due to the large number of chemicals present in the environment, exposure and risk assessment of chemical mixtures is complex and poses several challenges for scientists, risk assessors and managers (EFSA, 2016; Drakvik et al., 2020). Increasing awareness that daily-life exposure involves exposure to an almost infinite number of different combinations of chemicals, needing a move beyond chemical-by-chemical assessments, has led to a prioritisation of this topic in policy and research. These mixtures of exposure can form an almost infinite number of different combinations of chemicals, which makes the exposure and risk assessment extremely challenging. The phenomenon of mixtures (in the context of HBM) refers to the common occurrence at the level of the individual of chemical xenobiotic substances.

There is no broadly accepted operational definition of mixtures. The European Commission communication on “*The combination effects of chemicals – Chemical mixtures*” (2012) was published in response to a request from the European Parliament for the Commission to consider the extent to which the existing legislation “adequately addresses risks from exposure to multiple chemicals from different sources and pathways, and on this basis considers appropriate modifications, guidelines and assessment methods”. In the communication, mixtures are differentiated as follows:

1. **Intentional mixtures:** these are manufactured formulated products that are marketed as such. The composition of such mixtures and the hazardous properties and classification of the constituents is (generally) known (e.g., pharmaceuticals, plant protection products).
2. **Mixtures originating from a single source:** also known as ‘unintentional mixtures’ (Kienzler et al., 2016; Rotter et al. 2018) these are the result of discharges to the environment during the production, transport, use or disposal of goods, often contain a mixture of chemical substances. The composition can either be known (for example an effluent) or it can be unknown (e.g., waste related).
3. **Mixtures of chemicals originating from multiple sources and through multiple pathways:** also known as “coincidental mixtures” these relate to multiple substances from multiple and varying sources. Their composition is unknown and can vary in both space and time (e.g., exposure to humans to multiple chemicals from food and drinking water).

Intentional, unintentional and coincidental mixtures can arise from combinations of ambient environments and indoor sources, food products or contamination, consumer products, cosmetics, occupational exposures, medication and medical implants and lifestyle (e.g., smoking, recreational drugs, tattoo ink). In principle, every single substance, once it enters the body, will exhibit its health effects in interaction with a person’s genetic makeup and acquired characteristics, and in concert with all other (xenobiotic) substances from previous and simultaneous exposures. These mixtures thus form a challenge to (experimental and observational) science, to mechanistic and casual assessment of risks and to regulation of substances and general risk management policies (Agier et al, 2016; Barrera-Gómez et al., 2017).

In this document the term mixture is used to describe any combination of exposure to chemical substances or of internal dose biomarkers that have been measured in one or more biological matrices of a person during a single time point. These biomarkers include the chemical substances themselves, and their metabolites. The mixture might include compounds from multiple substance

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebrecht, Madeline Carsique, Eva Govarts, et al.	Page: 9

groups (e.g., perfluorinated compounds, phthalates, etc.), but also multiple compounds from the same substance group (e.g., pesticides, mycotoxins, etc.).

## 4.2 Overall objectives and tasks of WP15

The **overall aim** of the HBM4EU work package “Mixtures, HBM and human health risks” is to improve the precision and efficacy of HBM to inform science, policy and regulatory actions with respect to dealing with mixtures. More specifically, WP15 has four main objectives:

1. To develop summary indicators to describe the exposure and body burdens of mixtures with an emphasis on defining priority mixtures and drivers of mixture toxicity.
2. To re-evaluate existing HBM mixture data and the collection of new HBM mixture data, to identify real-life exposure patterns to mixtures.
3. To further develop and apply practical approaches to identify and assess the potential health risks and impacts of mixtures.
4. To inform policy makers, stakeholders and the public at large about mixture exposures and associated health risks.

In order to achieve these objectives, various tasks have been conducted in WP15. Task 15.1 is focused on the re-analysis of existing data on mixtures from earlier HBM studies using a key set of summary indicators. This work comprises the following:

1. Statistical analyses on existing data with focus on exposure profiles that take into account coincidental exposures.
2. Toxicity-based approaches, e.g., develop cumulative assessment groups (CAG) based on common working mechanisms.
3. Hybrid approaches, using Hazard Index (HI) and/or Biomonitoring Equivalents (BE)

This deliverable D15.7, entitled “Report on re-analysis of existing HBM data for priority substances to identify mixtures”, summarizes the work done, presents the results and interprets and discusses the findings.

## 4.3 Questions underlying the tasks in 15.1

The statistical analysis of existing HBM data is explorative in nature, with the primary question “What are the HBM mixture levels in the European population?”. More specific (research) questions that are addressed in task 15.1 include:

1. How can we rank order individuals based on body burdens to mixtures?
2. What patterns can we observe amongst body burdens of different substances within individuals? Or are people with high exposure levels for some substances more likely to have high exposure levels on other substances as well?
3. Are such patterns indicative for specific sources or pathways of exposure?
4. Can we identify hotspots or risk groups with high body burdens of mixtures?
5. Can we develop aggregate/hybrid indicators that encapsulate toxicity of the mixture in a meaningful way, e.g., by using HI or Cumulative Assessment Groups (CAGs)?
6. Can we define mixture levels of excess risk, based on toxicity-based aggregation of HBM mixture data?

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 10

These questions are the starting point for the statistical analysis carried out. The proposed approach, data preparation and statistical analysis and scripts have been described in Deliverable AD15.3, entitled: “HBM mixture database description and proposal statistical analysis plan”.

## 4.4 Overall approach

The research questions from the previous section were the starting point for the overall approach regarding statistical analysis. The analysis strategy aims to describe the distribution of (patterns in) biomarkers of exposure, to identify determinants that explain observed variation of patterns in biomarkers of exposure, as well as to rank individuals based on their cumulative burden of exposure. The correlation patterns in HBM mixture data can be presented graphically using heatmaps, circos plots (circular correlation globes) and correlation networks. These allow for the identification of groups of exposure biomarkers that are more closely related than others. Network analysis, a graphical method to represent the relationships between groups or communities in the data, is an intuitive interpretation of patterns in the data without any prior assumptions. Further insights can be generated using comparative network analysis (CNA), which is an analytical procedure that allows for the comparison of two or more networks based on (dis)similarities. CNA can be used to assess the impact of covariates on observed networks.

In brief, the approach taken for the network analyses involved development of statistical scripts for alternative avenues to analysis the co-occurrence of substances in blood and urine samples, data preparation steps, descriptive statistical analysis and network analysis as well as so-called SNMU (Sparse Non-negative Matrix Underapproximation) analysis. Both network analysis and SNMU are approaches to explore the possible patterns in co-occurrence of HBM biomarkers in samples collected from the same individual at the same point in time.

Parallel to network analysis, to address research questions 5 and 6, an approach to enrich the mixture HBM data with toxicity weighting was explored. To this end, a database of health-based guidance values was developed by IOM with other partners. As a ‘proof of principle’ exercise, the information from that database was used and combined with the German network analysis results.

The approach, statistical analysis and scripts are described in more detail in Deliverable Report AD 15.3 “HBM mixture database description and proposal statistical analysis plan” (2017), Deliverable Report D15.3 “Report real-life exposure profiles from re-analysis of existing HBM mixture data” (2019) and in Ottenbros et al (2021). D15.3 provides initial statistical scripts and first application on a simulated dataset, since GDPR delayed the access to actual HBM data. In the publication Ottenbros et al (2021) we describe the first application of the network analysis on the HBM data from the Flemish Environment and Health Studies (FLEHS) studies. Subsequent analysis of the approach on other datasets from Germany, Spain, the Czech Republic and Belgium was performed, and these are reported in this deliverable. Also, a comparison between the network analysis and the SNMU approach was made using the FLEHS data, which is also reported in this deliverable. In particular, we want to compare the mixtures obtained between both methods by applying them to the same data set, and to highlight the strengths and limitations of each of the approaches.



D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 11

## 5 Material and Methods

### 5.1 Selection of existing HBM data

After the initial development of statistical scripts, testing on simulated data and a first proof of principle on the FLEHS data, the scope of the network analysis was expanded to other existing HBM databases from four other studies participating in HBM4EU, with the aim to further explore the added value of network analysis. These studies originate from Germany, Spain, the Czech Republic and Belgium. The selection of the studies was based on data availability as well as availability of appropriate statistical expertise at the respective institutes: GDPR prevents the sharing of HBM data for re-analysis unless the data are fully anonymised. Therefore, the four studies are treated separately, and a combined analysis of the studies is not possible. In the following sections, data selection and preparation as well as the backgrounds of the four studies are described in more detail.

### 5.2 Data selection and preparation

#### 5.2.1 Network analysis

Data selection and preparation steps were performed with the subsequent network analysis in mind. Hence, for each of the studies the most data-rich subset was chosen in terms of the maximum number of biomarkers measured. The data preparation steps, described in more detail in D15.3 and Ottenbros et al. (2021), were harmonised with WP10 and involved:

- Checking the distribution of the variables
- Transforming the data if needed
- Imputation of the data points below the LOD or LOQ (Limit of Detection; Limit of Quantification)
- Correction for outliers
- Standardisation around zero
- Scaling of the data

The statistical approach taken is described in more detail in section 5.4. Concentrations of biomarkers were natural log transformed because HBM distributions are typically skewed. The actual network analysis is performed on a correlation matrix. Therefore, a strategy for dealing with missing data is required. Thus, an (arbitrary) cut-off of a maximum of 40% of HBM levels below the LOD/LOQ (level of detection/level of quantification) was applied. Substances with more than 40% of the measured HBM values below LOD/LOQ were excluded from further analysis. For the included substances, missing values below LOD/LOQ were imputed based on a maximum likelihood estimation via single conditional imputation, dependent on observed values for the other biomarkers (Lubin et al., 2004). Missing values in biomarkers (completely missing, e.g., due to insufficient sample volume) and determinants were imputed by using a single imputation strategy using the R package mice. Determinants/covariates were imputed first, using linear regression for continuous variables and logistic regression for the binary variables. The determinants and observed values were then used as prediction matrix for single imputation of the biomarkers (completely missing, e.g., due to insufficient blood volume), using linear regression.

For several substances, notably metals, different species were measured, e.g., for arsenic total arsenic, organic and inorganic arsenic values were available. Also, in some studies the same parent compound was measured in urine and in blood, e.g., lead or cadmium in urine and in blood.

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 12

This would lead to relatively high correlations between the different indicators for the same substance. In terms of mixture exposures such correlations do not provide relevant information; it may also affect the partial correlations structure with other substances. Therefore, only a single HBM indicator was selected for inclusion to the network analysis, preferably the indicator that best reflects the long-term exposure of the individual. Furthermore, metabolites of the large group of phthalates were not summed up to their diesters but included in their monoester concentration.

For substances measured in urine a correction for creatinine level was performed, to take into account the dilution level of spot or morning urine samples; the dilution level could affect the correlation structure with other substances measured in urine. For relevant substances measured in blood, blood lipid levels were used to adjust measured blood levels. This approach was preferred over adding creatinine or blood lipids as explanatory variable to the model, because in some of the studies a combination of matrices was used.

## 5.2.2 Network analysis with toxicity weighting

The same data preparation steps and dataset as described above were used prior to toxicity weighting. Depending on the unit of the HBGV either the imputed biomarker concentrations or the creatinine corrected data was used for the calculation of the hazard quotients. Since hazard quotients are unitless they can easily be combined even if the underlying HBGVs have different units. The hazard indices were calculated per community if two or more HBGVs were available for the present biomarkers.

## 5.3 Existing HBM data

### 5.3.1 GerES V (Germany)

#### 5.3.1.1 Characteristics

The German Environmental Survey for children and adolescents 2014-2017 (GerES V) is a population-representative cross-sectional study carried out in order to determine the exposure to pollutants of the general population in Germany. GerES V investigated children and adolescents by determining, on a representative basis, the body burden of environmental pollutants and the exposure to pollutants at home. It was performed in a stratified-randomly selected sample design consisting of 2,294 children and adolescents aged 3 to 17 years and living in 167 different sampling locations in Germany (Schwedler et al., 2020).

In GerES V, different substances were measured in subsets of participants. To avoid high proportions of missing data, urinary biomarkers available for all participants and urinary biomarkers in a subgroup of GerES V participants (n = 515) were included in this analysis. All urinary biomarkers were measured in first-morning void urine samples. For the current report weighted (i.e., nationally representative) data from a subsample of 515 participants aged 3 to 17 years were used.

#### 5.3.1.2 Data used for the analysis

##### 5.3.1.2.1 Substance groups and biomarkers of exposure

The following HBM4EU priority substances were available in first-morning void urine samples in the selected subsample of 515 participants: cadmium, chromium, mercury, phthalates, DINCH, bisphenol A, polyaromatic hydrocarbons (PAHs), acrylamide, pesticides, aprotic solvents (n-ethyl-pyrrolidone; n-methyl-pyrrolidone), UV-filters (benzophenones (BP)). In addition, the following non-HBM4EU substances were included: antimony (Sb), selenium (Se), parabens, lysmeral and CIT/MIT (methylchloroisothiazolinone/methylisothiazolinone).



D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebrecht, Madeline Carsique, Eva Govarts, et al.	Page: 13

Ten biomarkers were excluded from the analyses because over 40% of the measurements were below the Level of Quantification (LOQ): phthalate metabolites MnOP, MnPeP, MCHP, OH-MPHP, cx-MPHP; aprotic solvents (NEP) metabolite 5-HNEP; pesticide glyphosate and its metabolite AMPA; and the UV-filter metabolites of BP-1 and BP-3. As a result, a total of 51 substances was included in the analyses (see Table 6.1.1 in section 6.1 for details). On average, there were less than 1% missing data (ranging from 0 to 6.2% across biomarkers), which were imputed as described in section 5.2.

### 5.3.1.2.2 Determinants

Determinants that were made ready to be included in the differential network analysis are age (in years), sex, Body Mass Index (BMI), and smoking status of the participant, education of the household (ISCEDhh), and creatinine for imputation of missing data and as covariates in regression analyses. In addition, to be used as stratifying factors, binary variables of BMI, ISCEDhh, and age were created (see Table 5.3.1). Missing data in determinants (mean=0.3%; range = 0–1.9%) were imputed as described in section 5.2.

**Table 5.3.1: Selected determinants for GerES V**

Variable	Categories	N (%)
Sex	Female	250 (49%)
	Male	265 (51 %)
Smoking status	Smoking	12 (2 %)
	Non-smoking	503 (98 %)
Age at median	10 years and younger	263 (51 %)
	Older than 10 years	252 (49%)
ISCED of household	Low & medium (ISCED 0-4)	222 (43 %)
	High (ISCED >=5)	293 (57%)
Body Mass Index	Normal (<25 kg/m <sup>2</sup> )	443 (86%)
	Overweight (>=25 kg/m <sup>2</sup> )	73 (14 %)
Covariate	Unit	Mean (SD) Range
Creatinine	g/L	1.30 (0.65)
		0.15 – 5.02

### 5.3.1.3 Data preparation

Data preparation steps are described in section 5.2. In addition, in the last step of data preparation, population weights were included into regression analyses to make the dataset more population representative with respect to age, sex, community size, socioeconomic status, former East/West Germany, and migration background of the German population.

## 5.3.2 BIOAMBIENT.ES (Spain)

### 5.3.2.1 Characteristics

The BIOAMBIENT.ES study was designed as a population-based cross-sectional epidemiological study representative of the Spanish workforce, with self-administered questionnaires, medical examinations and collection of biological samples throughout the Spanish territory (Pérez-Gómez *et al.*, 2013). The study participants were selected through a stratified sample by conglomerates, to

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 14

guarantee the inclusion of all the geographical areas of the territory, both sexes and different sectors of activity (services sector and others). The study population includes subjects aged 16 or over, who were resident in Spain for at least the 5 years prior to the start of the study, and who attended the occupational medical examinations during 2009. The fieldwork was conducted between March 2009 and July 2010.

Of the 1,892 participants who constitute the population sample of the BIOAMBIENT.ES project, 1,880 subjects provided samples with sufficient whole blood volume, while 1,770 subjects provided valid morning void urine samples (defined by having creatinine levels between 0.3 and 3 g/L). Hair samples were provided by 577 participants.

The epidemiological questionnaire was designed to collect basic individual information on sociodemographic data, lifestyle, environmental conditions and some personal characteristics. Questions about the frequency of food consumption were also included to record habitual diet, as well as about recent illnesses and the use of medications.

### 5.3.2.2 Data used for the analysis

#### 5.3.2.2.1 Substance groups and biomarkers of exposure

For the purpose of this analysis, and as not all substances were measured in every participant, 3 different subsets were evaluated for the network analysis. Of the three subgroups analysed, the one with the highest number of substances was finally selected, despite reducing the sample size. As explained in section 5.2, only data for biomarkers with at least 40 % values >LOQ were used. The final subset includes 163 participants for whom heavy metals (mercury, cadmium, lead, thallium and cobalt), phthalates (DMP, DEP, BBzP, DiBP, DnBP, DEHP, DiNP and DiDP), DINCH and PFAs (PFHxS, PFOA, PFOS, PFNA and PFDA) were measured. This results in a total of 31 biomarkers measured in at least 60% of the participants in the subset.

#### 5.3.2.2.2 Determinants

The determinants included in the network analyses are sex, age, smoking status, ISCED, BMI, and fish consumption. Summary statistics for these determinates are provided in the table below.

**Table 5.3.2: Selected determinants for BIOAMBIENT.ES**

Variable	Categories	N (%)
Sex	Female	86 (52.8%)
	Male	77 (47.2 %)
Age	22-39	92 (56.4%)
	40-60	71 (43.6%)
Smoking status	Smoking	67 (41.1 %)
	Non-smoking	96 (58.9 %)
ISCED	Low-Medium	103 (63.2%)
	High	49 (30.1%)
	n.a.	11 (6.7%)
Body Mass Index	Normal (<25 kg/m <sup>2</sup> )	78 (47.9%)
	Overweight (>25 kg/m <sup>2</sup> )	83 (36.8%)
	n.a.	2 (1.2 %)
Fish consumption	Low (<2 /week)	67 (41.1%)
	High (>2 /week)	94 (57.7%)
	n.a.	2 (1.2%)

n.a.: not available

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 15

### 5.3.3 CELSPAC - FIREexpo (Czech Republic)

#### 5.3.3.1 Characteristics

The CELSPAC - FIREexpo study, conducted in the Czech Republic, aims to determine the health risks resulting from the occupational exposure of Czech firefighters, and to implement measures to minimise such risks.

For the purposes of the network analysis and the scope of the work described in this report, data from two population groups were selected: i) the occupationally exposed professional firefighters involved in combating the fire and using the firefighting foams (n = 52); ii) and the control group (n = 55). All participants were males in the age of 18 - 35 years, to better represent the population characteristics of the active firefighters in the Czech Republic. Smoking was an exclusion criterion as it could lead to higher levels of PAHs and other markers.

The sampling took place from January 2019 till July 2020. Morning void urine and venous blood samples were sampled and transported to the CELSPAC laboratories at RECETOX, where the samples were processed and analysed for biomarkers of interest.

#### 5.3.3.2 Data used for the analysis

##### 5.3.3.2.1 Substance groups and biomarkers of exposure

A dataset containing the serum levels of 12 perfluorinated compounds (PFAS) and urine levels of 10 hydroxylated polycyclic aromatic hydrocarbons (OH-PAHs), markers of exposure to PAHs (Table 5.3.3) of all study participants has been used. The data on these HBM4EU priority substances were available at the time of the analyses carried out within the scope of the work reported here. The dataset was divided into two subsets, according to studied group (professional firefighters / control). There were no missing data in these subsets. For further analysis, only biomarkers with at least 60 % values >LOQ were used (see section 5.1. for detailed explanation).

**Table 5.3.3: List of serum PFAS and urine OH-PAHs used for the analysis**

Substance/group	Measured biomarker	Acronym
PFAS	Perfluoropentanoic acid	PFPA
	Perfluorohexanoic acid	PFHXA
	Perfluoroheptanoic acid	PFHPA
	Perfluorooctanoic acid	PFOA
	Perfluorononanoic acid	PFNA
	Perfluorodecanoic acid	PFDA
	Perfluoroundecanoic acid	PFUDA
	Perfluorododecanoic acid	PFDDA
	Perfluorobutanoic acid	PFBS
	Perfluorohexane sulfonic acid	PFHXS
	Perfluoroheptane sulfonic acid	PFHpS
	Perfluorooctane sulfonate	PFOS
Naphthalene	1-OH_naphthalene	1-OH-Nap
	2-OH-naphthalene	2-OH-Nap

Substance/group	Measured biomarker	Acronym
Fluorene	2-OH-fluorene	2-OH-Flu
	3-OH-fluorene	3-OH-Flu
Phenanthrene	2/3-OH-phenanthrene	2/3-OH-Phe
	9-OH-phenanthrene	9-OH-Phe
	1-OH-phenanthrene	1-OH-Phe
	4-OH-phenanthrene	4-OH-Phe
Pyrene	1-OH-pyrene	1-OH-Pyr
Benzo[a]pyrene	3-OH-benzo[a]pyrene	3-OH-BaP

### 5.3.3.2.2 Determinants

The data on education level were not collected in this study. All participants were non-smoking males. Therefore, stratification for education, sex and smoking status was not relevant for this study. Age and BMI were included as covariates in the analyses.

### 5.3.4 3xG (Belgium)

#### 5.3.4.1 Characteristics

The 3xG (Gezondheid, Gemeenten, Geboorten) study is a scientific study that monitors the health of the inhabitants of three bordering rural communities (Dessel, Mol and Retie) in Flanders, Belgium, over time. This study focuses on the effect of the environment and lifestyle on health. This is done by researching 301 growing children from the region and by processing the disease and mortality registers of the three municipalities. The 3xG study is carried out to safeguard the health of the population in the region in relation to environmental pressures in the study area. The aim of the 3xG study is to follow-up the health and development of growing children as a sentinel population and to study the influence of environmental exposures via biomonitoring. It is one of the initiatives in the region to positively impact well-being and welfare of the population.

All pregnant women in the region that fulfilled the inclusion criteria and were expected to give birth between 2010 and 2015 were invited to participate. In total 301 mother-newborn pairs were obtained from 3xG. All participants signed an informed consent. Inclusion criteria were to be able to fill out a Dutch questionnaire and to live in the recruitment area (Govarts et al. 2020).

All participants agreed to fill in questionnaires during pregnancy and after delivery. Socioeconomic characteristics, smoking habits, information on consumption of local food and the course of pregnancy were collected. A urine sample was also collected during pregnancy. Birth weight, length and head circumference of the baby at birth were collected in agreement of the mothers. A blood sample of the mother and umbilical cord blood were collected at delivery and mother's breast milk between 1-4 weeks after delivery and a questionnaire was filled in by the mothers at the same time points.

More information is publicly available on the website of the study, <https://studie3xg.be/nl>.

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebrecht, Madeline Carsique, Eva Govarts, et al.	Page: 17

### 5.3.4.2 Data used for the analysis

#### 5.3.4.2.1 Substance groups and biomarkers of exposure

The following substances and substance groups were available in a selected subsample of 125 participants: metals including cadmium (Cd), nickel (Ni), chromium (Cr), antimony (Sb), copper (Cu), tellurium (Te) and lead (Pb), total arsenic (As), hydroxypyrene (HPYR), trans-muconic acid (TTMA), phthalates including mono(2-ethylhexyl) phthalate (MEHP), mono(2-ethyl-5-hydroxy-hexyl) phthalate (HMEHP), mono(2-ethyl-5-oxo-hexyl) phthalate (OMEHP), mono-n-butyl phthalate (MNBP), mono-benzyl phthalate (BMZP), mono-ethyl phthalate (MEP) and mono-isobutyl phthalate (MIBP), and bisphenol A (BPA) were available in morning urine (UM) samples of the pregnant mother; musks including tonalide (AHTN) and galaxolide (HHCb) were available in the blood samples (MB) of the mother after delivery; metals (cadmium, nickel, chromium, antimony, copper, tellurium and lead) and arsenic were available in cord blood (PB) samples of the newborn and organochlorine compounds (OCs) including polychlorinated biphenyl 138 (PCB128), polychlorinated biphenyl 153 (PCB153), polychlorinated biphenyl 180 (PCB180), dichlorodiphenyldichloroethylene (DDE) and hexachlorobenzene (HCB), and PFASs including Perfluorooctane sulfonic acid (PFOS), Perfluorooctanoic acid (PFOA) and Perfluorohexane sulfonic acid (PFHXS) were available in cord blood plasma (PP) samples of the newborn.

Biomarkers were selected based on two factors. Firstly, the frequency of detection of the biomarker should be at least 60%, this is, at least 60% of the values had to be above the detection limit or quantification limit. Secondly, not all biomarkers were measured in the same participants, some were only measured in a subset. Therefore, we selected the biomarkers that ensure a subset with enough participants. For this reason, biomarkers such as, for example, arsenic metabolites, were not included as they were only measured in a small subset of participants. A subset of 125 mother-child pairs were consequently included in the network analysis.

It is important to address that the definition of mixtures was loosely applied to the 3xG cohort for the sole purpose of proof of concept. This loose definition of mixtures implies that biomarkers are measured in what we consider two participants (mother and newborn) and within the mother's samples, measurement times differ: urine samples were taken during pregnancy and blood samples were taken at the time of delivery.

#### 5.3.4.2.2 Determinants

For our data, stratification and network comparison by sex was not ideal, as then only those biomarkers measured in the children would be included in the network. Due to this issue, we have decided to use the following determinants (Table 5.3.4): BMI of the mother before pregnancy, categorised in normal (<25 kg/m<sup>2</sup>) and overweight (≥25 kg/m<sup>2</sup>), international standard classification of education (ISCED) of the mother, categorised as low (ISCED 0-4) and high (ISCED ≥5), smoking of the mother during pregnancy (yes/no), and fish consumption of the mother during pregnancy, categorised in low (< 2 times/week) and high (≥ 2 times/week).

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebrecht, Madeline Carsique, Eva Govarts, et al.	Page: 18

**Table 5.3.4: Selected determinants for 3xG**

Variable	Categories	N (%)
<b>Child, N = 125</b>		
Sex		
	Male	62 (49.6%)
	Female	63 (50.4%)
<b>Mother, N=125</b>		
BMI before pregnancy (kg/m <sup>2</sup> )		
	Normal (<25 kg/m <sup>2</sup> )	88 (70.4%)
	Overweight (>=25 kg/m <sup>2</sup> )	37 (29.6%)
ISCED		
	Low (ISCED 0-4)	36 (28.8%)
	High (ISCED >=5)	89 (71.2%)
Smoking during pregnancy		
	Yes	12 (9.6%)
	No	113 (90.4%)
Fish consumption		
	Low (< 2 times/week)	51 (40.8%)
	High (>= 2 times/week)	74 (59.2%)

### 5.3.5 FLEHS (Belgium)

#### 5.3.5.1 Characteristics

The Flemish Environmental Health Study (FLEHS) is a large Human Biomonitoring (HBM) project conducted in Flanders, an industrialised region in the north of Belgium with 6 million inhabitants, since the year 2002, covering 3 campaigns and with the fourth one ongoing, to assess human exposure to environmental chemicals by measuring biomarkers of exposure in combination with biomarkers of effect in three age categories (newborns, adolescents and adults). The population samples for each campaign are representative for the geographical distribution and the population density in Flanders.

The Human Biomonitoring component of FLEHS III includes measurements of more than 20 biomarkers of exposure and 30 biomarkers of effect. In the newborn campaign, cord blood samples have been collected between 2013 and 2014 for a total of 281 participants. As we aimed to identify mixtures among biomarkers, we decided to focus on the newborn campaign of FLEHS III which has a larger variety of substances compared to the other two campaigns.

#### 5.3.5.2 Data used for the analysis

##### 5.3.5.2.1 Substance groups and biomarkers of exposure

In FLEHS III, the 19 biomarkers among the 42 biomarkers analysed were selected: metals: arsenic (As), cadmium (Cd), copper (Cu), Manganese (Mn), lead (Pb), thallium (Tl); persistent chlorinated pollutants: p,p'-dichlorodiphenyldichloroethylene (p,p'-DDE), hexachlorobenzene (HCB), polychlorinated biphenyls (PCB118, PCB138, PCB146, PCB153, PCB170, PCB180, PCB187);

per- and polyfluoroalkyl substances (PFHXS, PFNA, PFOA, PFOS). All the selected exposure biomarkers were measured from cord blood and they were included for analysis because more than 60% of the measurements was above the LOD.

### 5.3.5.2.2 Determinants

The determinants available for the 281 pairs of mother-newborn of FLEHS III are given in Table 5.3.5.

**Table 5.3.5: Selected determinants for FLEHS III**

Variable	Categories	N (%)	Mean	Min	Max
<b>Child, N = 281</b>					
Sex	Male	144 (51.2%)			
	Female	137 (48.8%)			
Birth year			2013.95	2013	2014
<b>Mother, N=281</b>					
BMI before pregnancy (kg/m2)			23.85	15.24	45.17
Age at birth (years)			30.18	18.93	44.84
Smoking during pregnancy	Yes	33 (11.7%)			
	No	137 (88.3%)			
Parity	1st child	126 (44.8%)			
	2nd child	100 (35.6%)			
	3 or more children	55 (19.6%)			
Height (cm)			167	150	181
Gestational age (weeks)			39.35	35	42
Size for gestational age	Small	21 (7.5%)			
	Normal	228 (81.1%)			
	Large	32 (11.4%)			
Fish consumption	Low (less than weekly)	109 (38.8%)			
	Medium (weekly)	71 (25.3%)			
	High (more than weekly)	101 (35.9%)			
Lipid (g/L)			1.86	1.19	3.73
Cholesterol (mg/dL)			0.67	0.38	1.54
Triglycerides (mg/dL)			0.36	0.13	0.93

### 5.3.5.3 Data preparation

The same data preparation process described in Ottenbros et al. (2021) was applied to FLEHS III data. The SNMU method is a method based on non-negativity constraints. Therefore, an exponential transformation is applied on the residualised biomarker concentrations obtained with above-described preparation to get a non-negative matrix as dataset.



D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebrecht, Madeline Carsique, Eva Govarts, et al.	Page: 20

### 5.3.6 Toxicity weighting

#### 5.3.6.1 Background

Various approaches for estimating the joint toxic action of chemical components of mixtures exist. The simplest of these is the hazard index approach that assumes dose-additivity (Meek et al. 2011; US DoHHS, 2004). To derive the Hazard Index of a chemical mixture, a Hazard Quotient of each chemical component ( $HQ_i$ ) is firstly calculated by scaling the estimated exposure or dose of the component in the population ( $C_i$ ) by a level of exposure considered safe or acceptable; for the purposes of this project this is the case. Calculation of hazard indices requires data on typical population exposure levels and Health-Based Human Biomonitoring Guidance Values (HB-HBGV<sub>*i*</sub>, where *i* indicates the *ith* component of a mixture)

$$HQ_i = C_i / HB-HBGV_i$$

The Hazard Index (HI) of the chemical mixture is then calculated by summing the respective hazard quotients:

$$HI = \sum_{i=1}^n HQ_i$$

An HQ or HI >1 indicates that exposure to the substance is greater than a threshold level of concern and warrants further investigation. This approach assumes that the cumulative effect of the doses of each mixture component is additive. The HB-HBGV may be derived based on different effects for different chemicals, but in the HI approach treats all guidance values agnostically and does not take into account common modes of action. This approach should be relatively conservative, although it is possible that if the effects are synergistic, there could be an underestimation of toxicity. The HI approach is recommended for screening mixtures to determine if further investigation is needed. If a mixture HQ is greater than one, then evaluation of the mixture in more detail, considering the mode of action of each component, can be done. Calculation of hazard indices requires data on typical population exposure levels and Health-Based Human Biomonitoring Guidance Values (HB-HBGV) for each component of the mixture.

#### 5.3.6.2 Human Biomonitoring Health-Based Guidance Values

We established a database of HB-HBGV values for the HBM4EU priority substances listed by Schoeters et al. (2020). A scoping review was conducted. This targeted biological guidance or regulatory values preferably established by (but not limited to) national or European agencies, organizations or commissions, such as the German Human Biomonitoring Commission, ANSES (France), or Health and Safety Executive (UK). Guidance values for substances that were not available through agencies in Europe, or that were considered interesting for comparison purposes, were sourced from similar type organisations outside the region, e.g., US Environmental Protection Agency (US EPA), US Agency for Toxic Substances and Disease Registry (ASTDR), or New Zealand Government - WorkSafe.

Descriptions of the different types of guidance values are described in Annex 4. Briefly, HBM-GV<sub>GenPop</sub> and HBM-GV<sub>worker</sub> are derived within HBM4EU for substances identified by the HBM4EU chemicals prioritisation strategy based on existing needs to answer policy relevant questions as raised by national and EU policy makers (Apel et al. 2020). They are equivalent to the HBM-I values from the German Human Biomonitoring Commission and refer to the concentrations in biological media at which no health risk is expected to occur. This is unlike HBM-II values that represent concentrations above which there is an increased risk for adverse health effects.

Where an HBM-GV<sub>GenPop</sub>, HBM-GV<sub>worker</sub>, HBM-I or HBM-II was not available for a substance, biomonitoring equivalent (BE) values were sourced from the peer-reviewed literature. These values



D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebrecht, Madeline Carsique, Eva Govarts, et al.	Page: 21

are derived from pharmacokinetic data to estimate the concentration of a chemical or its metabolite in biological media that is consistent with an existing health-based exposure guidance value (e.g., tolerable daily intake). They are intended for use as screening tools to provide an assessment of which chemical biomarkers are present at levels well below, near, or at or above health risk-based exposure concentrations (Kirman et al. 2011).

The methods used for establishing workplace guidance values vary considerably between countries but most incorporate information on patterns of occupational exposure across the working life in their derivations such as repeated exposures and 8-hour exposure durations (Annex 4; Table A4.2).

## 5.4 Statistical analysis

### 5.4.1 Descriptive analysis

The descriptive analysis of the data used for network analysis largely follows the conventions developed in WP10. Thus, central tendency and distributional measures are provided to allow an assessment of the HBM levels observed. Common scripts were used to generate the tables presenting the descriptive results. Where there is overlap in measured substances, a comparison between studies is possible, keeping in mind that sampling strategies, chemical analytical procedures and study populations vary substantially between studies.

Descriptive statistics were calculated using the software tool RStudio (developed using version 2021.09.2). Number of values and missing values, percentage below LOD and LOQ, mean, standard deviation, standard error and geometric mean were calculated using standard R functions. Percentiles (P05 to P95) were calculated by means of the quantile function (package 'stats', version 3.6.2). Descriptive statistics were calculated on the imputed values and corrected for creatinine or lipid, in those cases where it proceeded (biomarker measured in urine in case of creatinine correction or measured in blood in case of blood correction for specific biomarkers). Percentiles below LOD and LOQ were specified instead of providing the percentile obtained after imputation.

In preparation of the network analysis, also the Pearson correlation structures in the datasets is displayed using heatmaps and circos plots analogous to the scripts described in AD15.3.

### 5.4.2 Network analysis

Network analyses were performed as described in AD15.3 and in the publication on the first application of the approach on the FLEHS data (Ottenbros et al. 2021). Unweighted network analysis is used to describe the conditional independence between multiple variables, making use of the packages *huge* and *igraph*, using R (v3.5.0 or higher) (Csárdi et al., 2006; Zhao et al., 2012). Within these networks, a node or dot represents a biomarker, and an edge or line between two nodes reflects the conditional dependency between these two biomarkers given all other variables. For comparison purposes, weighted network analysis was applied as well, making use of the package EGAnet (v0.9.6) (Golino et al., 2020). The weighted network shows the strength of the edge (absolute correlation) by thickness of the line, and direction of the correlation by color of the line (green for a positive correlation, red for a negative correlation).

The graph estimation was conducted using the graphical lasso, which involves penalised maximum likelihood estimation (Friedman et al., 2008). This method is a simple and fast algorithm for estimation of a sparse inverse covariance matrix using an L1 penalty. The graphical lasso cycles through the variables, fitting a modified lasso regression to each variable in turn. Regularisation of the graph was conducted along a sequence of 10 equally spaced lambdas ranging from the maximum lambda (resulting in an empty graph) to the minimum lambda set at 10% of the

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 22

maximum lambda. Optimal lambda selection was conducted using the stability approach to regularisation selection method (StARS), which selects the optimal lambda by variability across subsamples (Liu et al., 2010). Variability (or instability) across subsamples is defined as the fraction of times (range: 0–0.5) that two graphs disagree on the presence of an edge, averaged over all edges in the graphs. We used the default variability threshold of 0.1. Within the selected network, the walktrap clustering algorithm from the igraph package was used, which performs random walks (in default of 4 steps) across the network to merge separate communities in a bottom-up manner (Orman and Labatut, 2009; Pons and Latapy, 2005). Nodes were colored according to the community they were assigned to; edges linking different communities were colored in red, edges within a community were colored in black.

For the weighted correlation networks, the graphical lasso algorithm was used as well, with an EBIC tuning parameter setting to 0.5. A parametric bootstrap (1,000 iterations) was used to estimate the median network structure, which was then plotted as the final result.

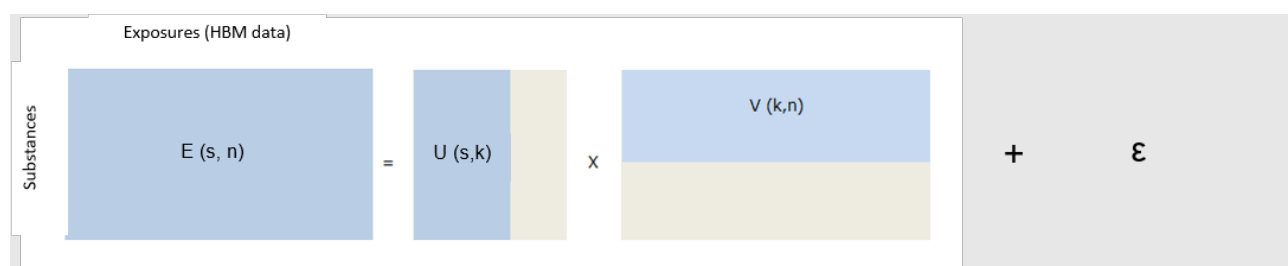
The systematic comparison of the unweighted networks can be done based on the presence or absence of edges. Also called exact graph matching, between two or more graphs with the exact same set of nodes. With this comparison the impact of covariates can be assessed, by comparing networks of strata of the data. Similarities and/or dissimilarities will be presented in another network, pointing out the edges which are similar or different.

### 5.4.3 SNMU

The proposed approach is a combination of a dimension reduction method based on the Sparse Non-negative Matrix Under-approximation (SNMU) model (Gillis et Plemmons 2013, Gillis and Glineur, 2010) to identify mixtures and a hierarchical classification to cluster individuals regarding their co-exposure to the identified mixtures. It also integrates an analysis of exposure determinants. This method was applied several times to prioritise mixtures of chemicals from dietary co-exposures (Béchaux et al. 2013, Traoré et al. 2016, 2018) and breastmilk contamination (Crépet et al. 2022).

First, the SNMU aims to find a representation of the non-negative data input matrix in a lower dimension. In other words, the exposure matrix  $E$  (input data) is approximated by two non-negative matrices ( $U$  and  $V$ ) such as the product of the two should be as close as possible to matrix  $E$ . The model of this approximation is (Figure 5.4.1). The matrix  $U$  represents the composition of each identified mixtures and contains the weight of each substance per mixture. The substance weight represents the participation of each substance to the overall exposure of one mixture. Each column of the matrix represents the composition of one mixture where a non-zero entry in this column indicates that the corresponding substance is part of the mixture, and if it is equal to 0 the substance is not contributing to the mixture in return. Then, a high substance weight indicates a high participation of the substance to the mixture. The other matrix  $V$  represents the exposure of the  $n$  individuals to the identified mixtures and contains the coefficient of the presence of the mixture per individual. The dimension of matrices  $U$  and  $V$  are  $(s,k)$  and  $(k,n)$  respectively, with  $s$  the number of substances,  $n$  the number of individuals, and  $k$  the candidate number of mixtures manually pre-set. How to obtain the final  $k$  is described below. The matrices  $U$  and  $V$  are obtained by minimising the criterion:  $\|E - UV\|^2$  such that  $U \geq 0$  and  $V \geq 0$ .

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebrecht, Madeline Carsique, Eva Govarts, et al.	Page: 23



**Figure 5.4.1: SNMU model**

Then, the SNMU method is combined with a classification by performing hierarchical clustering of individuals with similar exposure profiles. A hierarchical cluster analysis (HCA) with Ward's method is applied to the  $V$  matrix obtained for a  $k$  candidate values, to group individuals with similar profiles of exposure to the  $k$  mixtures into clusters. Results of the clustering are presented in a dendrogram. The optimal number of population clusters is chosen by searching the longest jump between two levels in the dendrogram. In addition, we examine how exposure to the  $k$  mixtures characterises the clusters to choose it, in order to compare the optimal number of cluster choice with the biological reality.

Generally, the selection of the optimal  $k$  mixtures is performed by examining results of the SNMU combined with hierarchical clustering calculated with different  $k$  values (Traoré et al., 2016, Crepet et al., 2022). In other words, the SMNU algorithm is run for several values of  $k$  comprised between 1 and  $n$ , and the associated residual sum of squares between the exposure matrix  $E$  and the product of the estimated matrices  $U$  and  $V$  is calculated for each trial.

The proposals for the optimal  $k$  can be ranked in descending order of their residual sum of squares, the highest ones being considered as optimal. Then, this statistical criterion is balanced in regards of the context of the study, of biological relevance and quality of the interpretation of results in terms of clustering using the first optimal  $k$  values. For example, the  $k$  values for which at least one mixture is not used to characterise a cluster, or only concerns a small part of the population is rejected. In our specific case where we would like to compare the SNMU mixtures with the ones obtained with the network analysis, therefore the first optimal  $k$  values obtained with the above-proposed process is challenged with the one corresponding to the number of mixtures identified with the network analysis.

Once the final number of mixtures  $k$  is defined, the clusters of individuals obtained are described by their exposure to the mixtures and the individual characteristics. By applying a Fisher test between the characteristics of the individuals in a cluster and the total population, we can highlight particularises of the clusters.

#### 5.4.4 Data selection for toxicity weighting

To develop the database on HBGVs, the list of substances measured in the four studies involved in the network analyses and reported in D15.9 were taken as a starting point. A first subset of 51 biomarkers were prioritised, based on the 60% cut-off of substances above LOQ in the German data set that was selected for this proof-of-principle exercise. HB-HBGVs were identified for 26 of the selected biomarkers. A more detailed list can be found in Annex 3 along with an indication of the biomarkers for which an HB-HBGV exist. As some substances included more than one HB-HBGV, they were prioritised as follows:

1. HB-HBGVs for the general population were preferred, as this was not an occupational study.
2. HBM-GVs derived from HBM4EU or HBM-I values were preferred, with values derived from European agencies, before other international agencies.

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 24

3. The most recently derived values were preferred.
4. If the above values were not available, then BEs were used.

The structure of the database is shown in Figure 5.4.2 For each biomarker of interest, the following information was captured in the database:

- HBM4EU priority substance group to which the biomarker belonged (PrioritySubstGp)
- HB-HBGV and its type, value, unit, biological matrix
- Biomarker (name, acronym, CAS-number (where no CAS-number was found its acronym was used in the CAS-number field))
- parent compound of the biomarker (name, acronym, CAS number)
- country where the BGV was established (BGV\_origin)
- year the BGV was established (BGV\_year)
- for BE's the guidance value used in deriving it (e.g., RfD, TDI, MRL) was also noted
- description of the study population (e.g., sex, adult, children, ages) (StudyPop\_ages; StudyPopDesc)
- source of the HB-HBGV

For the substances in the network, HBGVs were matched on medium, i.e., urine, dimension and where appropriate creatinine normalisation. Where appropriate, matching on age-group was performed. The network analysis as described in Section 5.3.3 was run on the GerES V data and the hazard index per individual was calculated. Also, a hazard index was calculated at the level of the identified communities in the network. Analyses were done using HBM-GV's where these were available. Where no HBM-GVs existed, alternative indicators e.g., Biological Equivalent (BE) were used.

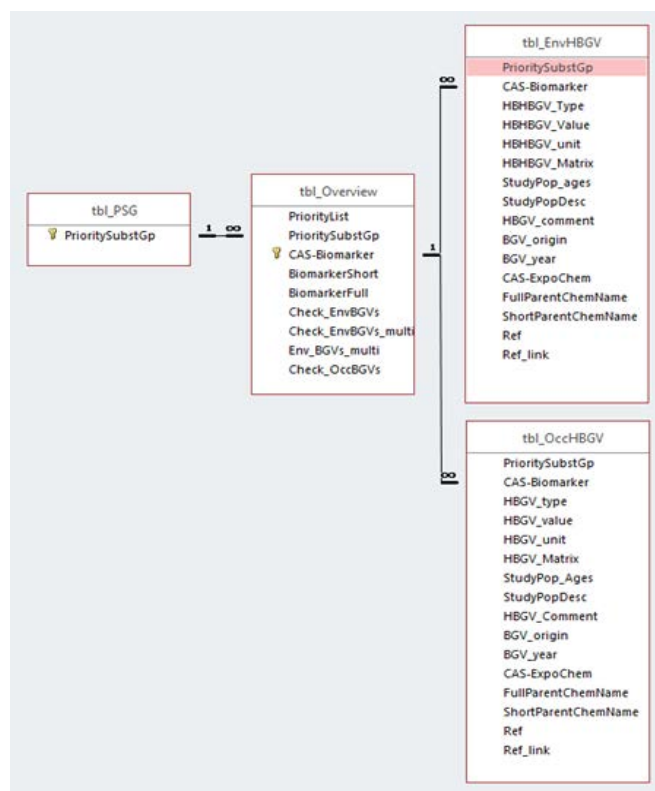


Figure 5.4.2: Structure of the database of HB-HBGVs in MS Access



## 6 Results

### 6.1 Descriptive analysis

#### 6.1.1 GerES V (Germany)

Table 6.1.1 shows all substances included for network analyses in GerES V, their proportions below LOQ, and percentiles and GM of the creatine-normalised biomarker concentrations. Results with data corrected for creatinine and major exposure determinants (see section 5.3.1) from heatmaps (Figure 6.1.1) indicate mostly positive, small to medium correlations. For example, chromium and NMMA show correlations around .3 with several metabolites from other substance groups such as acrylamide, aprotic solvents, and some phthalates whereas the lowest correlations observed with other substance groups ( $r = \sim 0.027$ ) are observed for phthalate substitute DINCH, arsenic, mercury, and the parabens. In contrast, correlations between metabolites of the same substance showed the highest correlations (up to  $r = \sim 0.95$ ), for example acrylamide and phthalates and their substitute DINCH and DEHTP.

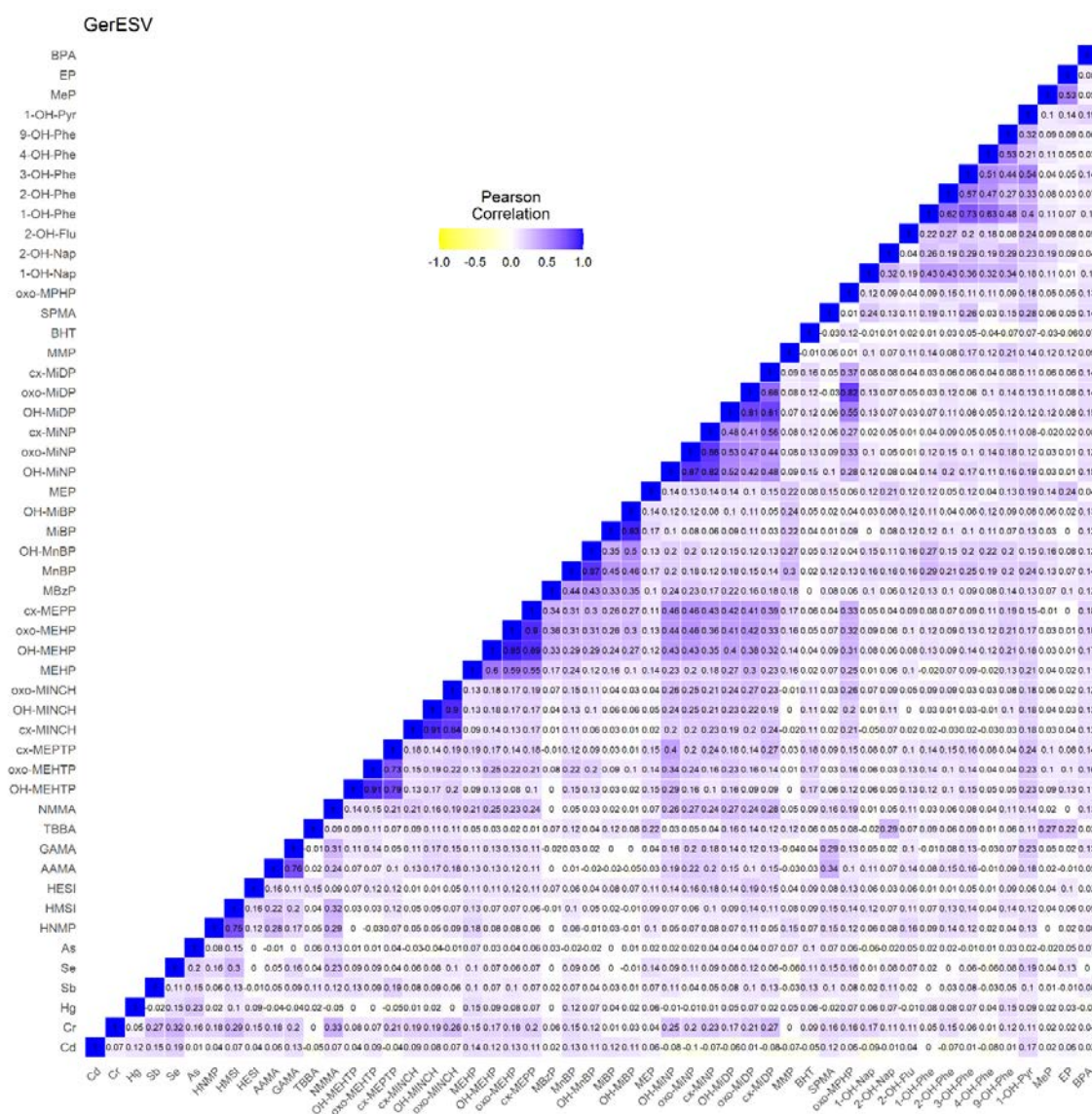


Figure 6.1.1: Heatmap of Pearson correlations of 51 urinary biomarkers in the GerES V subsample of 515 participants

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 26

**Table 6.1.1: Descriptive statistics for substances included in the network analysis of the GerES V subsample (unweighted) from 515 participants**

Substance group	Substance	Biomarker	N < LOQ	% < LOQ	LOQ	P05	P10	P25	P50	P75	P90	P95	GM
Elements	Cadmium	Cd	134	26.02 %	0.05	< LOQ	< LOQ	< LOQ	0.06	0.09	0.12	0.15	0.06
	Chromium	Cr	40	7.77 %	0.2	< LOQ	0.2	0.26	0.34	0.49	0.62	0.77	0.36
	Mercury	Hg	26	5.05 %	0.02	< LOQ	0.02	0.04	0.06	0.1	0.19	0.26	0.06
	Antimony	Sb	108	20.97 %	0.04	< LOQ	< LOQ	0.03	0.05	0.07	0.1	0.13	0.05
	Selenium	Se	0	0 %	0.5	15.09	16.81	21.17	27.8	37.95	47.93	57.08	28.34
	Arsenic	As	0	0 %	0.1	2.45	2.93	4.35	6.89	14.17	30.5	55.21	8.42
Aprotic solvents		HNMP	0	0 %	2.5	17.78	22.73	31.79	48.6	73.76	107.36	152.65	49.71
		HMSI	0	0 %	2	15.6	19.62	26.81	37.15	57.12	84.41	104.4	39.32
		HESI	66	12.82 %	2	< LOQ	< LOQ	2.58	4.68	10.2	39.19	70.35	5.87
Acrylamide		AAMA	0	0 %	1	26.84	32.62	43.56	60.5	84.27	125.31	189.67	63.22
		GAMA	0	0 %	1	5.52	6.79	8.96	12.53	17.66	24.62	28.84	12.74
Phthalate substitutes	DEHTP	OH-MEHTP	171	33.2 %	0.3	< LOQ	< LOQ	< LOQ	0.41	1.13	2.59	4.23	0.48
		oxo-MEHTP	103	20 %	0.2	< LOQ	< LOQ	0.19	0.46	1.06	2.28	3.83	0.47
		cx-MEPTP	0	0 %	0.2	1.1	1.51	2.85	6.22	15.45	36.64	54.09	6.76
	DINCH	cx-MINCH	1	0.19 %	0.05	0.21	0.29	0.49	1.02	2.11	4.91	7.8	1.08
		OH-MINCH	1	0.19 %	0.05	0.41	0.54	0.98	2.13	4.66	9.48	14.72	2.22
		oxo-MINCH	8	1.55 %	0.05	0.15	0.21	0.39	0.93	2.03	4.6	7.16	0.94
Phthalates	DEHP	MEHP	67	13.01 %	0.5	< LOQ	< LOQ	0.71	1.22	2.04	3.35	4.19	1.2
		OH-MEHP	0	0 %	0.2	3.12	3.94	5.87	8.98	13.94	21.69	28.84	9.25
		oxo-MEHP	0	0 %	0.2	1.96	2.52	4.08	6.42	10.49	15.89	21.63	6.47
		cx-MEPP	0	0 %	0.2	3.45	4.03	6.1	9.92	16.9	26.26	35.81	10.17
	BBzP	MBzP	2	0.39 %	0.2	0.69	0.9	1.45	2.38	4.75	10.36	17.46	2.77
	DnBP	MnBP	0	0 %	1	6.05	7.81	12.04	18.18	28.67	40.34	54.79	18.39
		OH-MnBP	4	0.78 %	0.25	0.6	0.78	1.25	2.12	3.49	5.24	7.27	2.11
	DiBP	MiBP	0	0 %	1	7.29	9.04	13.54	21.36	33.58	58.2	87.22	22.34
		OH-MiBP	0	0 %	0.25	2.28	3.02	4.7	7.52	12.17	21.02	30.15	7.78

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 27

Substance group	Substance	Biomarker	N < LOQ	% < LOQ	LOQ	P05	P10	P25	P50	P75	P90	P95	GM
	DEP	MEP	0	0 %	0.5	5.09	6.83	10.96	17.76	32.05	65.95	113.45	19.75
	DiNP	OH-MiNP	0	0 %	0.2	1.88	2.28	3.35	5.27	8.73	15.12	24.62	5.76
		oxo-MiNP	0	0 %	0.2	0.73	0.92	1.39	2.17	3.66	6.35	9.65	2.36
		cx-MiNP	0	0 %	0.2	1.52	1.82	2.88	4.55	7.5	12.48	19.47	4.87
	DiDP	OH-MiDP	4	0.78 %	0.2	0.37	0.47	0.75	1.19	2.06	3.54	5.9	1.28
		oxo-MiDP	54	10.49 %	0.2	< LOQ	< LOQ	0.29	0.54	0.89	1.55	2.56	0.54
		cx-MiDP	11	2.14 %	0.2	0.24	0.3	0.41	0.7	1.19	2.2	3.62	0.76
	DPHP	oxo-MPHP	184	35.73 %	0.25	< LOQ	< LOQ	< LOQ	0.27	0.54	1.01	1.57	0.29
	DMP	MMP	8	1.55 %	1	1.49	1.93	3.21	5.07	10.44	21.45	36	6.02
PAHs		1-OH-Nap	18	3.5 %	0.05	0.11	0.19	0.36	0.68	1.41	3.42	4.88	0.7
		2-OH-Nap	1	0.19 %	0.05	0.92	1.15	1.86	3.15	5.89	11.06	15.89	3.38
		2-OH-Flu	54	10.49 %	0.05	< LOQ	< LOQ	0.23	0.43	0.69	1.27	2.19	0.36
		1-OH-Phe	0	0 %	0.005	0.04	0.05	0.08	0.12	0.2	0.34	0.46	0.13
		2-OH-Phe	4	0.78 %	0.005	0.03	0.03	0.05	0.07	0.11	0.18	0.28	0.08
		3-OH-Phe	0	0 %	0.005	0.05	0.05	0.08	0.11	0.18	0.3	0.4	0.12
		4-OH-Phe	2	0.39 %	0.001	0.01	0.01	0.02	0.04	0.08	0.18	0.26	0.04
		9-OH-Phe	14	2.72 %	0.005	0.01	0.02	0.03	0.05	0.09	0.19	0.28	0.05
		1-OH-Pyr	7	1.36 %	0.01	0.03	0.04	0.06	0.09	0.14	0.22	0.29	0.09
Parabens	Methylparaben	MeP	13	2.52 %	0.5	0.8	1.04	1.9	4.37	19.61	122.99	321.93	7.02
	Ethylparaben	EP	164	31.84 %	0.5	< LOQ	< LOQ	< LOQ	0.62	1.42	4.38	10.39	0.72
Bisphenols	Bisphenol A	BPA	19	3.69 %	0.5	0.52	0.67	1.03	1.6	2.88	4.8	6.91	1.77
Other	Lysmeral	TBBA	0	0 %	0.2	2.12	2.87	4.49	8.16	15.45	24.19	35.56	8.49
	CIT/MIT	NMMA	0	0 %	0.5	2.48	2.99	3.92	5.31	7.53	10.24	12.2	5.47
	Butylhydroxytoluol	BHT	1	0.19 %	0.2	0.58	0.73	1.23	1.98	3.45	6.22	9.52	2.1
	Benzene	SPMA	12	2.33 %	0.02	0.02	0.03	0.05	0.08	0.14	0.26	0.44	0.09

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 28

Security: Public

Version: 1.1

Page: 28

**Figure 6.1.2: Circos plot of 51 urinary biomarkers in the GerES V subsample of 515 participants.**

Table 6.1.3 shows the descriptive statistics for biomarkers included for network analyses in the subsample of the BIOAMBIENT.ES study.



D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 29

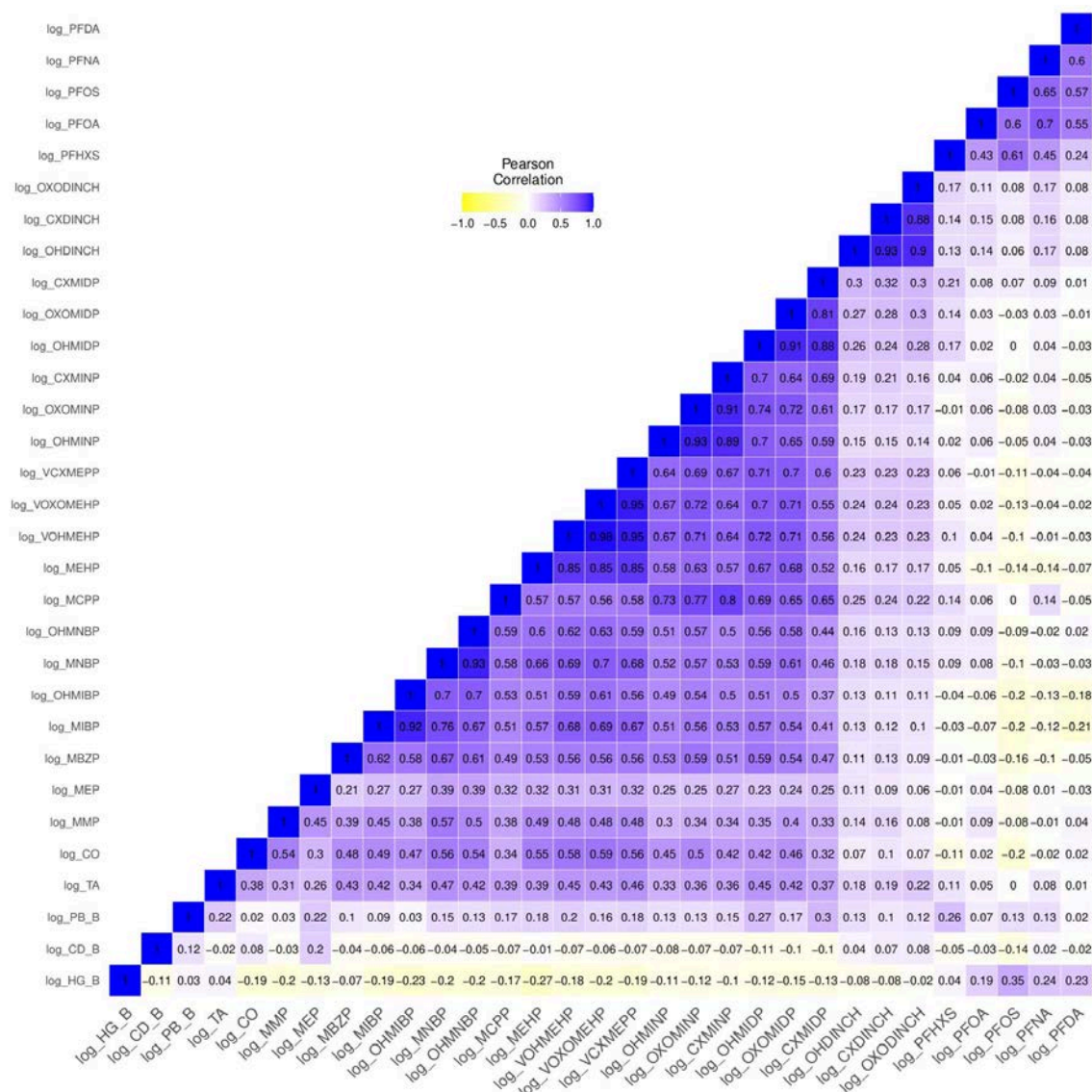
**Table 6.1.2: Descriptive statistics for substances included in the network analysis for BIOAMBIENT.ES**

Substance group	Substance	Biomarker	N	N < LOQ	% < LOQ	LOQ	P05	P10	P25	P50	P75	P90	P95	GM
Metals	Mercury	Hg	163	1	0,61	0,10 µg/L	1,91	2,75	5,30	6,94	10,30	13,94	15,48	6,53
	Cadmium	Cg	163	1	0,61	0,05 µg/L	0,15	0,21	0,30	0,46	0,96	1,50	2,16	0,52
	Lead	Pb	163	0	0	0,10 µg/L	11,02	13,50	15,95	21,60	28,75	44,84	56,98	22,82
	Thallium	Tl	141	16	11,35	0,05 µg/L	0,04	0,04	0,10	0,17	0,24	0,30	0,33	0,14
	Cobalt	C	137	2	1,46	0,05 µg/L	0,23	0,34	0,49	0,74	1,27	1,79	2,83	0,75
Phthalates	DMP	MMP	163	8	4,91	1 µg/L	1,10	1,60	2,40	3,60	5,85	11,74	18,43	4,01
	DEP	MEP	163	0	0	0,5 µg/L	30,51	42,08	113,00	262,00	477,00	1120,00	1944,00	241,49
	BBzP	MBzP	163	2	1,23	0,2 µg/L	1,21	2,14	4,00	6,60	13,65	25,90	38,81	7,26
	DiBP	MiBP	163	0	0	1 µg/L	7,24	11,34	20,45	32,00	51,15	80,02	101,58	30,68
		OH-MiBP	163	0	0	0,25 µg/L	3,01	5,34	7,95	12,90	20,35	28,84	33,04	12,23
	DnBP	MnBP	163	1	0,61	1 µg/L	4,54	6,70	11,10	19,70	35,00	47,68	73,73	18,56
		OH-MnBP	163	4	2,45	0,25 µg/L	0,50	0,72	1,30	2,30	3,55	5,50	7,59	2,11
		MCPP	163	22	13,5	0,50 µg/L	0,40	0,40	0,80	1,30	2,20	3,70	5,47	1,36
	DEHP	MEHP	163	6	3,68	0,50 µg/L	0,81	1,42	3,00	5,40	9,25	17,44	29,66	5,07
		OH-MEHP	163	0	0	0,20 µg/L	5,41	8,42	13,05	24,40	40,40	69,02	92,47	22,65
		oxo-MEHP	163	1	0,61	0,20 µg/L	3,90	4,80	9,50	15,40	26,05	43,60	58,42	14,75
		cx-MEPP	163	0	0	0,20 µg/L	4,84	10,06	14,45	26,70	40,90	66,98	97,85	24,49
	DiNP	OH-MiNP	163	3	1,84	0,20 µg/L	0,82	1,20	2,50	4,80	8,50	19,06	40,42	4,62
		oxo-MiNP	163	5	3,07	0,20 µg/L	0,50	0,70	1,50	3,00	5,40	10,72	23,63	2,80
		cx-MiNP	163	1	0,61	0,20 µg/L	1,62	2,54	4,85	8,80	13,90	27,20	73,10	8,78
	DiDP	OH-MiDP	163	3	1,84	0,20 µg/L	0,50	0,72	1,40	2,40	4,05	7,58	9,49	2,35
		oxo-MiDP	163	17	10,43	0,20 µg/L	0,10	0,22	0,50	0,90	1,40	2,30	3,29	0,82
		cx-MiDP	163	1	0,61	0,20 µg/L	0,51	0,80	1,25	2,00	3,05	5,08	8,23	2,03
DINCH	DINCH	OH-DINCH	163	8	4,91	0,05 µg/L	0,05	0,19	0,37	0,91	2,54	10,14	26,45	1,09
		cx-MINCH	163	6	3,68	0,05 µg/L	0,08	0,15	0,33	0,65	1,88	6,35	14,35	0,81
		oxo-DINCH	163	22	13,5	0,05 µg/L	0,04	0,04	0,14	0,40	1,64	7,52	14,07	0,51

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 30

Substance group	Substance	Biomarker	N	N < LOQ	% < LOQ	LOQ	P05	P10	P25	P50	P75	P90	P95	GM
PFAS	PFHxS	PFHXS	163	34	20,86	0,34 µg/L	0,24	0,24	0,39	0,68	1,13	1,99	2,39	0,70
	PFOA	PFOA	163	0	0	0,16 µg/L	0,81	0,96	1,39	2,03	2,94	3,92	5,09	1,98
	PFOS	PFOS	163	2	1,23	0,33 µg/L	2,48	3,53	5,30	8,09	11,02	15,18	17,11	7,25
	PFNA	PFNA	163	1	0,61	0,16 µg/L	0,48	0,59	0,70	0,95	1,39	1,74	2,14	0,98
	PFDA	PFDA	163	18	11,04	0,2 µg/L	0,14	0,14	0,26	0,37	0,53	0,75	0,84	0,37

Biomarker correlation analysis showed high overall correlation among substances in the same group, except for the metals. Cobalt and thallium showed higher correlations with the phthalates group than with mercury, cadmium and lead. No correlations were observed among cadmium and the rest of biomarkers. Some negative correlations but with very low absolute values were also present. These were observed mostly among mercury and all other biomarkers except PFAS, and among PFAS and most phthalates.



**Figure 6.1.3: Heatmap of 31 biomarkers in the BIOAMBIENT.ES subsample of 163 participants.**

The visualisation of the results in circo plots confirms these findings.

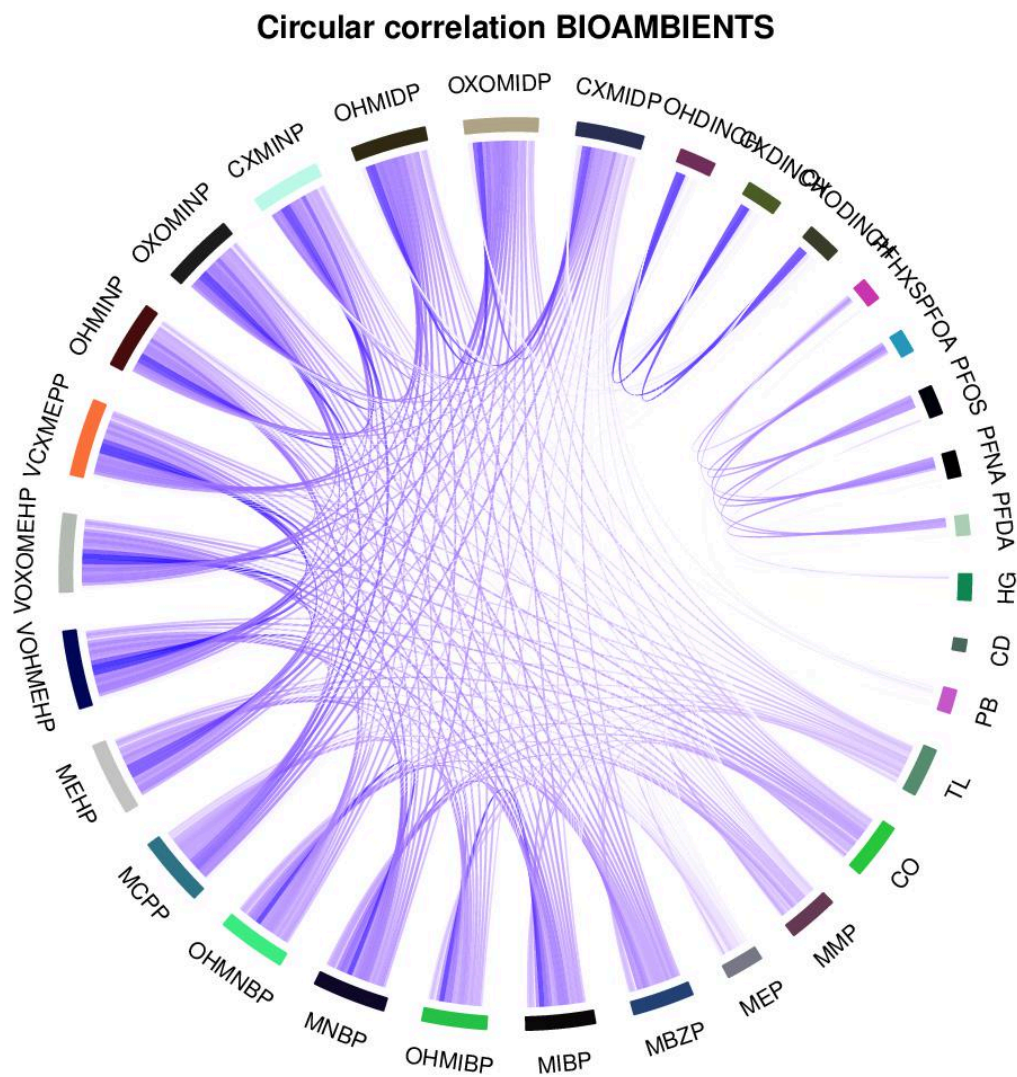


Figure 6.1.4: Circos plot of 31 biomarkers in the BIOAMBIENT.ES subsample of 163 participants

6.1.3 CELSPAC - FIREexpo (Czech Republic)

Table 6.1.3 shows the descriptive statistics for the biomarkers with at least 60% values above LOQ, for professional firefighters and the control group. Because of the case-control study design, the analysis of the detection frequency and the imputation was done separately for the two datasets (groups). Therefore, the list of substances used for further statistical analyses slightly differs between the two groups. The descriptive statistics were calculated from the imputed dataset that was further used for the network analysis.

Figure 6.1.5 shows the correlation heatmap of the analysed substances. The analysed compounds are in general more correlated within the same substance group, than among compounds that belong to different groups. This trend is better visible in professional firefighters, where the intra-group correlation of some compounds slightly increased, while the inter-group correlations remained weak (except for one correlation between PFBS and 4-OH-Phe). In the control group, the heatmap was more heterogeneous, the intra-group correlations were slightly weaker compared to firefighters, but some moderate inter-group correlations were observed.

**Table 6.1.3: Descriptive statistics for the imputed substances included in the network analysis for CELSPAC - FIREexpo**

Study population	Substance group	Biomarker	% below limit	LOQ	P10	P25	P50	P75	P90
Professional firefighters	Blood serum (ng/ml)								
	PFASs	PFOA	0%	0.07	0.663	0.92	1.205	1.5025	1.877
		PFNA	0%	0.004	0.231	0.29	0.4	0.5425	0.628
		PFDA	0%	0.004	0.101	0.1375	0.19	0.2525	0.299
		PFUDA	11.54%	0.012	<LOQ	0.04	0.05	0.07	0.1
		PFBS	26.92%	0.04	<LOQ	<LOQ	0.145	0.18	0.245
		PFHXS	0%	0.004	0.3	0.3775	0.485	0.665	0.759
		PFHPS	3.85%	0.005	0.04	0.0575	0.08	0.1	0.14
		PFOS	0%	0.03	1.659	2.3475	3.215	4.7825	6.357
	Morning urine (µg/g CRT)								
	PAHs	2-OH-Nap	0%	0.006	1.4284	2.8975	4.179	6.2071	9.5067
		2-Oh-Flu	0%	0.006	0.1408	0.1875	0.2581	0.3207	0.4751
		3-OH-Flu	5.77%	0.006	0.0214	0.0416	0.0646	0.1053	0.1554
		1-OH-Phe	38.46%	0.006	<LOQ	<LOQ	0.0186	0.0398	0.1038
		4-OH-Phe	1.92%	0.006	0.0446	0.2098	0.3241	0.5028	0.5893
		1-OH-Pyr	0%	0.006	0.0369	0.0468	0.0745	0.1024	0.1372
		2/3-OH-Phe	0%	0.006	0.0724	0.0955	0.1406	0.2004	0.2983
		1-OH-Nap	0%	0.006	0.5173	0.9887	1.6355	2.5602	3.6179
Control group	Blood serum (ng/ml)								
	PFASs	PFPA	16.36 %	0.036	<LOQ	0.18	0.22	0.255	0.306
		PFHXA	7.27%	0.04	0.05	0.07	0.08	0.095	0.11
		PFOA	1.82%	0.07	0.486	0.685	0.9	1.135	1.338
		PFNA	0%	0.004	0.184	0.23	0.3	0.36	0.412
		PFDA	0%	0.004	0.084	0.11	0.12	0.165	0.232
		PFUDA	3.64%	0.012	0.03	0.05	0.07	0.095	0.11
		PFHXS	0%	0.004	0.274	0.325	0.43	0.515	0.65
		PFHPS	36.36 %	0.005	<LOQ	<LOQ	0.04	0.07	0.09
		PFOS	0%	0.03	1.128	1.68	2.24	2.73	3.492
	Morning urine (µg/g CRT)								
	PAHs	2-OH-Nap	0%	0.006	1.2547	1.6589	2.7581	4.0497	5.2754
		2-OH-Flu	0%	0.006	0.0904	0.1264	0.1775	0.2451	0.313
		3-OH-Flu	3.64%	0.006	0.0144	0.0201	0.0303	0.0545	0.0721
		1-OH-Phe	5.45%	0.006	0.024	0.0508	0.0865	0.1179	0.1966
		4-OH-Phe	20%	0.006	<LOQ	0.0091	0.0204	0.0439	0.3543
		1-OH-Pyr	0%	0.006	0.0242	0.0301	0.0442	0.0624	0.0914
		2/3-OH-Phe	0%	0.006	0.0644	0.0918	0.1239	0.1988	0.2791
		1-OH-Nap	0%	0.006	0.3859	0.5763	0.9286	1.2848	2.2541





D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 35

### 6.1.4 3xG (Belgium)

Table 6.1.4 shows the descriptive statistics for all biomarkers previously mentioned that were detected for at least 60% in the subset of 125 mother-child pairs.

**Table 6.1.4: Substances included in the network analysis for 3xG. Matrix abbreviations used for the plots are also indicated in this table**

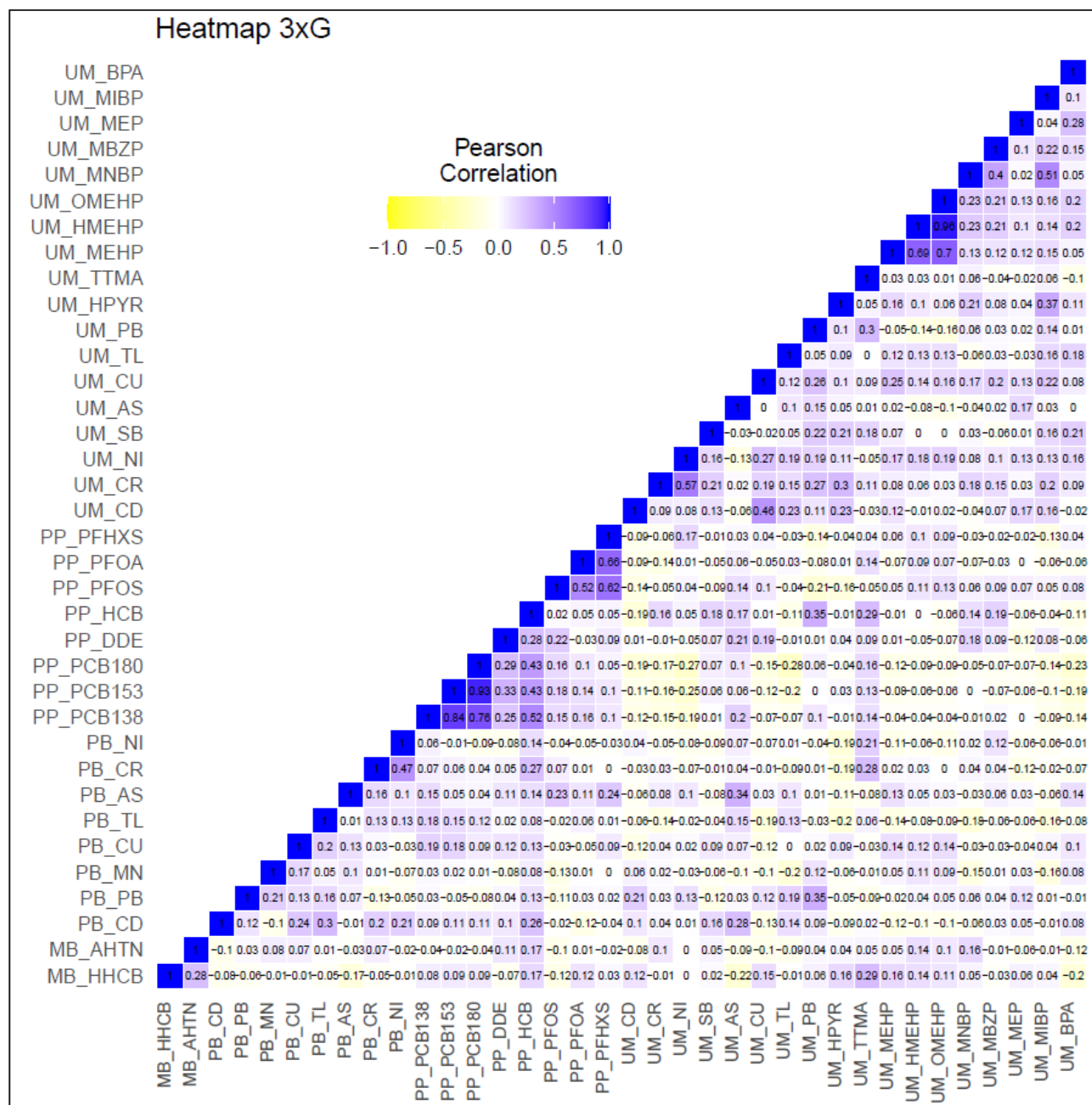
Substance group	Biomarker	Percentage below limit	LOD	LOQ	P05	P10	P25	P50	P75	P90	P95	Mean	SD	SEM	Geomean
Morning urine (UM) of the mother (µg/g CRT)															
<b>Metals</b>	CD	0 %	0.01 - 0.014		0.14	0.16	0.21	0.28	0.37	0.44	0.54	0.3	0.13	0.03	0.28
<b>Heavy metals</b>	CR	1.6 %	0.09		0.14	0.16	0.26	0.49	0.82	1.26	1.76	0.72	0.93	0.06	0.48
	NI	0 %	0.05 - 0.08		0.89	1.01	1.37	2.07	2.95	4.52	6.23	2.57	1.98	0.23	2.12
	SB	17.6 %	0.016 - 0.02		<LOD	<LOD	0.03	0.04	0.06	0.09	0.15	0.07	0.11	0.01	0.04
	CU	0 %	0.17 - 0.18		10.2	11.49	1	6	8	21.37	22.14	15.64	4.13	1.4	15.18
	TL	0 %	0.001 - 0.002		0.13	0.15	0.18	0.22	0.26	0.31	0.35	0.23	0.07	0.02	0.21
	PB	0 %	0.04 - 0.1		0.5	0.56	0.64	0.84	1.14	1.45	1.71	0.95	0.4	0.08	0.88
								13.8	28.7				90.2		
<b>Arsenic</b>	AS	0 %	0.04 - 0.074		4.07	4.5	6.72	9	9	51.59	81.22	33.89	1	3.03	15.33
<b>PAHs</b>	HPYR	1.6 %	0.0089	0.0179	0.06	0.08	0.11	0.15	0.24	0.34	0.49	0.28	0.73	0.03	0.16
<b>VOCs</b>	TTMA	0 %	0.29	0.59	28.91	38.41	6	5	229	2	7	159.02	72	14.22	104.58
<b>Phthalates</b>	MEHP	2.4 %	0.15	0.5	0.85	1.05	1.75	2.53	4.3	5.76	8.42	3.73	5.65	0.33	2.61
	HMEHP	0 %	0.03	0.1	4.46	5.3	6.67	8	5	25.59	38.04	15.09	6	1.35	10.7
	OMEHP	0 %	0.03	0.1	3.13	3.58	4.93	7.18	9.69	16.63	22.98	10.01	3	0.9	7.46
	MNBP	0 %	0.15	0.5	14.75	2	6	34	5	78.39	91.75	42.15	6	3.77	35.87
	MBZP	0 %	0.06	0.2	1.93	2.38	3.79	7.1	7	18.48	22.59	10.99	18.1	0.98	7.03
	MEP	0 %	0.15	0.5	6.42	9.23	6	8	8	6	7	68.5	2	6.13	39.65
						32.8	43.1				288.4		90.5		
	MIBP	0 %	0.15	0.5	25.7	8	6	60	94.4	173.4	6	90.05	5	8.05	67.36

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 36

Substance group	Biomarker	Percentage below limit	LOD	LOQ	P05	P10	P25	P50	P75	P90	P95	Mean	SD	SEM	Geomean
<b>Bisphenols</b>	BPA	2.4 %		0.2	0.59	0.67	0.9	1.29	2.3	3.73	4.61	2.07	2.84	0.19	1.46
<b>Blood (MB) of the mother (µg/L)</b>															
<b>Musks</b>	HHCB	0 %	20		148.4	175	230	282	359	497.6	542.8	307.51	126.62	27.5	284.29
	AHTN	0.8 %	20		32.6	38.4	50	63	81	132	186	77.57	51.18	6.94	66.69
<b>Cord blood (PB) of the newborn (µg/L)</b>															
<b>Metals</b>	CD	8 %	0.01 - 0.016,	<LOD	0.01	0.02	0.03	0.04	0.05	0.05	0.03	0.03	0	0.03	
	PB	0 %	0.04 - 0.2	3.89	4.1	5.07	6.09	7.29	9.13	9.8	6.48	2.25	0.58	6.17	
	MN	0 %	0.12 - 0.9	20.12	22.78	27.84	35.1	43.9	58.66	66.62	37.35	13.56	3.34	35.16	
<b>Heavy metals</b>	CU	0 %	0.17 - 1.2	487.2	503	540	576	626	664.8	697.8	587.76	87.55	52.57	584.06	
	TL	0 %	0.002 - 0.003	0.01	0.01	0.01	0.02	0.02	0.03	0.04	0.02	0.01	0	0.02	
	CR	18.4 %	0.09	<LOD	<LOD	0.14	0.24	0.57	0.86	1.23	0.43	0.53	0.04	0.25	
	NI	36.8 %	0.08	<LOD	<LOD	<LOD	0.12	0.2	0.34	1.19	0.33	0.87	0.03	0.13	
<b>Arsenic</b>	AS	0 %	0.04	0.24	0.29	0.5	0.89	1.55	2.62	3.27	1.27	1.32	0.11	0.89	
<b>Cord blood plasma (PP) of the newborn (µg/g lipid)</b>															
<b>OCs</b>	PCB138	6.4 %		0.01	<LOQ	0.01	0.01	0.01	0.02	0.03	0.03	0.01	0.01	0	0.01
	PCB153	1.6 %		0.01	0.01	0.01	0.02	0.02	0.03	0.04	0.05	0.02	0.01	0	0.02
	PCB180	5.6 %		0.01	<LOQ	0.01	0.01	0.01	0.02	0.03	0.03	0.02	0.01	0	0.01
	DDE	0 %		0.02	0.03	0.04	0.05	0.09	0.15	0.24	0.29	0.12	0.13	0.01	0.09
	HCB	24.8 %		0.01	<LOQ	<LOQ	0	0.01	0.01	0.01	0.02	0.01	0	0	0.01
<b>Cord blood plasma (PP) of the newborn (µg/L)</b>															
<b>PFASs</b>	PFOS	0 %		0.2	0.67	0.82	1.1	1.56	2.16	3.15	3.59	1.84	1.26	0.16	1.57
	PFOA	0 %		0.25	0.52	0.57	0.77	1.05	1.37	1.9	2.14	1.24	1.07	0.11	1.05
	PFHXS	20.8 %		0.2	<LOQ	<LOQ	0.21	0.34	0.46	0.57	0.76	0.35	0.22	0.03	0.29



The data available for the 3xG study were analysed in a subset of participants (see section 5.4 for details). For presentation purposes, results are shown for this subset of 125 participants. In Figure 6.1.7 the correlations between the substances analysed are depicted as heatmaps. The results show high correlations for e.g., cadmium and copper, and for different PCBs.



**Figure 6.1.7: Heatmap for 3xG subset, showing the Pearson correlations. Matrices in which biomarkers were measured are morning urine of the mother (UM), whole blood of the mother (MB), cord blood whole blood of the newborn (PB) and cord blood plasma of the newborn (PP).**

Additionally, correlations were visualised as circo plots (Figure 6.1.8). These plots greatly facilitate the interpretation of the findings. For example, Figure 6.1.8 nicely shows that exposure to different PFAS is correlated; the same applies to DEHP metabolites and the PCBs. Exposure to the latter group, however, also correlates with exposure to HCB and DDE

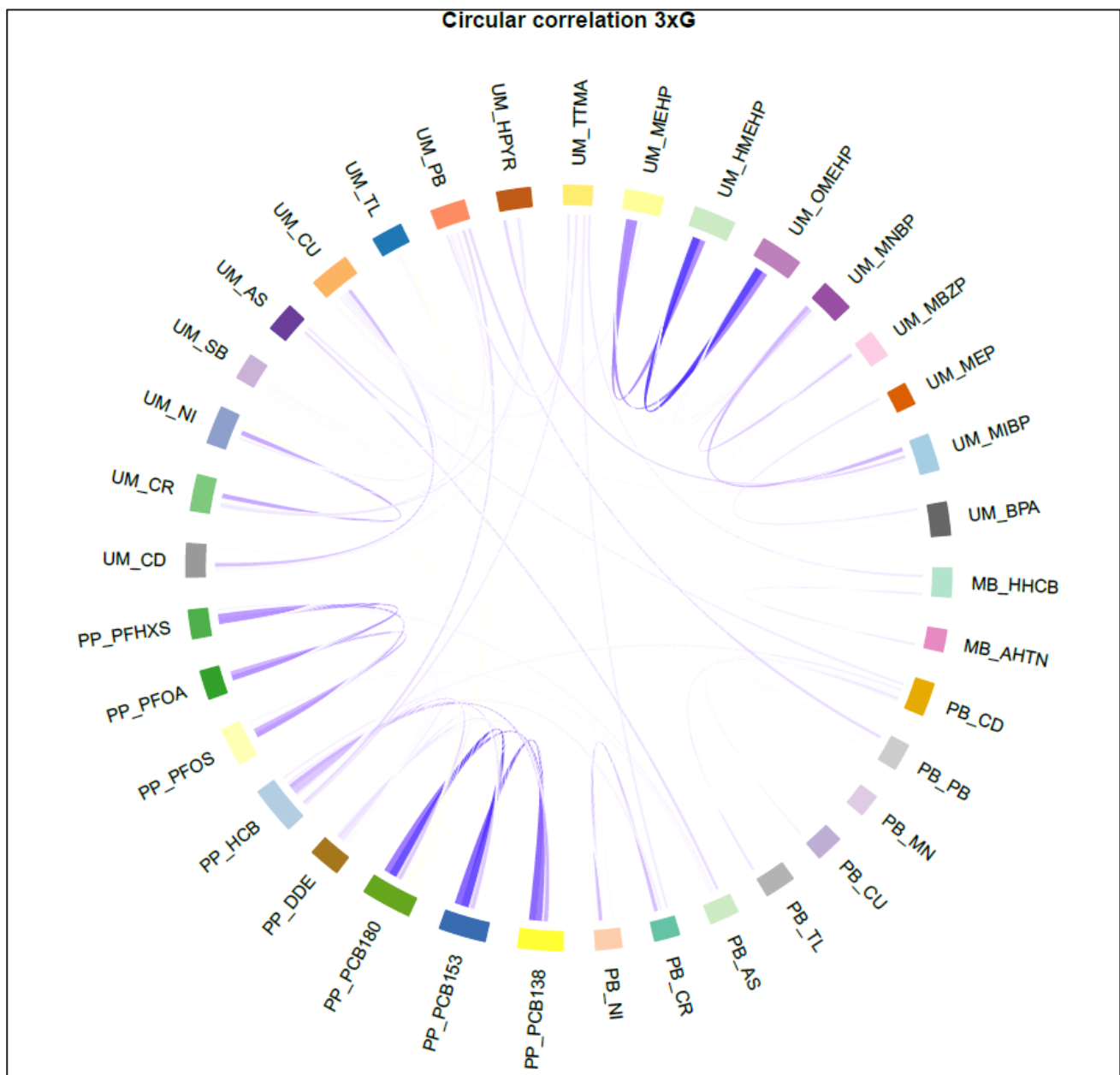


Figure 6.1.8: Circos plot for 3xG subset. Matrices in which biomarkers were measured are morning urine of the mother (UM), whole blood of the mother (MB), cord blood whole blood of the newborn (PB) and cord blood plasma of the newborn (PP).

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 39

### 6.1.5 FLEHS data SNMU

Table 6.1.5 shows the descriptive statistics for the 19 biomarkers selected from FLEHS III.

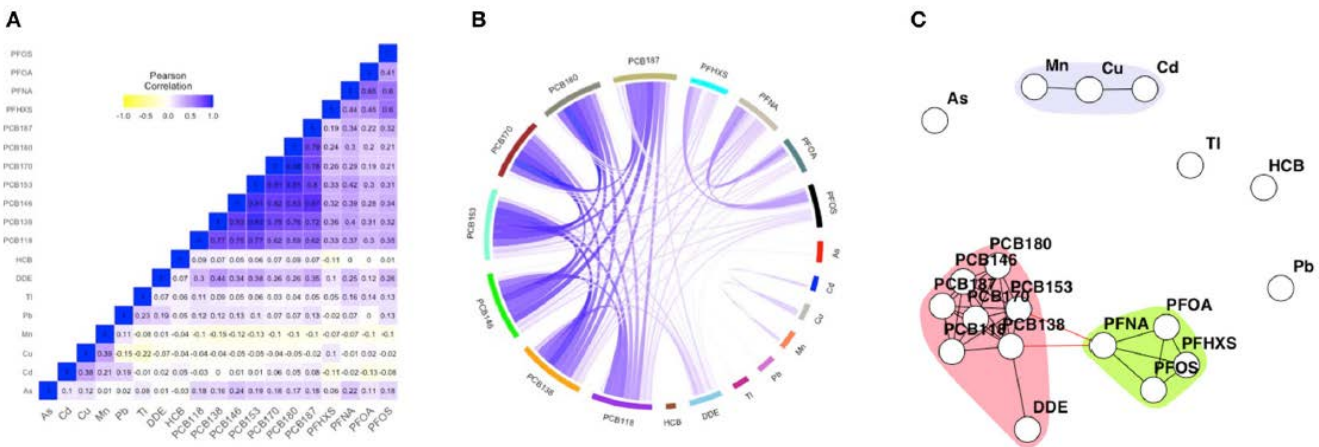
**Table 6.1.5. Substances included in the network analysis and SNMU for FLEHS III**

Substance group	Biomarker	Percentage below limit	FREQ.LOD	P05	P10	P25	P50	P75	P90	P95	Mean	SD	SEM	Geomean
Metals	As	0.36 %	0.09, N = 281	0.17	0.21	0.34	0.65	1.29	2.85	3.98	1.21	1.75	0.0721	0.7057
	Cd	1.42 %	0.0097, N = 281	0.011	0.013	0.016	0.021	0.027	0.034	0.039	0.0228	0.0109	0.0014	0.0209
	Cu	0 %	0.49, N = 281	399	424	507	571	624	686	730	565.8	99.72	33.75	556.8
	Mn	0 %	0.15, N = 281	16.69	19.24	24.03	30.15	38.03	46.25	53.04	31.92	11.28	1.90	30.10
	Pb	0 %	0.22, N = 281	3.16	3.5	4.88	6.07	8.57	11.5	12.73	7.206	4.04	0.43	6.45
	Tl	0 %	0.48, N = 281	12.36	13.59	15.66	18	22	26.31	30.52	19.32	5.77	1.15	18.59
Persistent chlorinated pollutants	DDE	0 %	20, N = 281	21.92	26.40	36.65	55.56	89.74	146.97	292.86	86.29	107.69	5.18	61.03
	HCB	21 %	10, N = 281	-1	-1	7.07	14.93	21.21	28.42	34.64	15.72	10.56	0.94	12.26
	PCB118	4.98 %	2, N = 281	1.21	1.46	2.26	3.39	5.05	7.14	8.61	3.97	2.45	0.24	3.31
	PCB138	0 %	2, N = 281	4.11	5.26	7.03	10.50	15.69	20.75	24.22	11.90	6.45	0.71	10.29
	PCB146	29.54 %	2, N = 281	-1	-1	-1	1.53	2.35	3.39	4.26	1.80	1.17	0.11	1.47
	PCB153	0 %	2, N = 281	6.41	8.33	11.47	16.93	24.02	32.08	40.11	19.08	10.48	1.14	16.47
	PCB170	1.78 %	2, N = 281	1.62	2.02	2.99	4.42	6.34	8.98	10.88	5.08	3.03	0.30	4.29
	PCB180	0 %	2, N = 281	2.99	4.09	5.63	8.47	13.07	18.18	21.48	10.23	6.66	0.61	8.51
	PCB187	11.03 %	2, N = 281	-1	-1	1.46	2.5	3.69	5.56	6.57	2.94	2.14	0.18	2.39
PFAS	PFHXS	15.3 %	0.2, N = 281	-1	-1	0.25	0.37	0.52	0.72	0.88	0.42	0.23	0.02	0.36
	PFNA	11.03 %	0.1, N = 281	-1	-1	0.14	0.21	0.31	0.43	0.52	0.25	0.17	0.01	0.21
	PFOA	0 %	0.2, N = 281	0.48	0.64	0.88	1.27	1.57	2.11	2.37	1.34	0.73	0.08	1.18
	PFOS	0.36 %	0.2, N = 281	0.42	0.5	0.73	1.12	1.68	2.55	3.13	1.3789	1.00	0.08	1.13

## 6.2 Network analysis

### 6.2.1 Generic example

As described in section 5.2, the network analysis describes the conditional dependence between multiple measured biomarkers, i.e. HBM levels. Where the heatmap and the circos plot show Pearson correlations, the network analysis is based on the partial correlation structure in the dataset, i.e., the direct associations between HBM levels corrected for the other markers present in the mixture. Figure 6.2.1 displays a heatmap, circos plot and network results, next to each other (from Ottenbros et al. 2021). Panel C displays the network results, where the circles represent the substance names or 'nodes' in the network and the lines, or 'edges' represent the connections between the HBM levels for those substances, based on their (partial) correlation structure. First, the network structure is estimated as described in section 5.2. Networks may consist of different subnetworks of connected nodes. Then, within the selected networks, so called 'communities' of more closely associated substances are identified and indicated by colors of grouped chemicals. Black lines indicate the network structure within a community, where red lines indicate associations across communities, in this case between PFNA and two PCBs. Thus, communities of biomarkers may display common sources of chemicals, common exposure routes, e.g., from environmental sources, lifestyle (diet, smoking, use of consumer products), occupation, or biochemical processes (absorption, distribution, metabolism, and excretion). When nodes are unconnected, there is little or no dependence between the substance HBM levels and that of other chemicals.



**Figure 6.2.1: Heatmap (A), circular correlation globe (B) and network including community detection (C) of FLEHS III, 19 biomarkers, n = 281.** Data is corrected for maternal age, smoking during pregnancy and maternal pre-pregnancy BMI. The heatmap is based on Pearson correlation between the biomarkers. Within the circular globe each biomarker is presented as a color-block on the circular axis. Within the network, each dot or node represents a biomarker, each edge represents a connection between the biomarkers, each different color represents a community within a subnetwork.

### 6.2.2 GerES V (Germany)

Network analyses were run on the subsample of 515 participants from GerES V with the methods as described in section 5.4.2. First, the role of different approaches of creatinine correction or lack thereof was checked (Figures 6.2.2-6.2.3). Weighted network analyses were performed as described in section 5.4.2 as well, using a parametric bootstrap of 1,000 iterations (Figure 6.2.5). Differential network analysis was applied to assess differences between stratifications by the covariates smoking status, sex, age, education (ISCED), and BMI (Figures 6.2.6 to 6.2.7).

To illustrate the effect of normalisation for creatinine, Figure 6.2.2 shows the resulting communities when adjusting raw concentrations for creatinine and controlling for creatinine in multivariate analyses in addition (recommended by HBM4EU, WP10). Several substances are not part of any community: for example, some elements (mercury, arsenic, cadmium), BPA, the phthalate MEP. A total of eight communities are found which contain three or more substances. The communities are grouped into DINCH metabolites (yellow), PAHs (green), parabens and TBBA (salmon), acrylamide and SPMA (plum), DEHTP metabolites (blue), aprotic solvent HNMP, NMMA, acrylamide (plum), selenium, chromium, antimony, aprotic solvent HMSI (lavender), and two communities of phthalate metabolites. Among the phthalate communities, MMP co-occurs together with BBzP, and DnBP and DiBP metabolites (grey); and DEHP metabolites co-occur with DPHP, and DiNP and DiDP metabolites (blue).

### Network GerES V (subsample, adjusted and controlled for crt, n = 515)

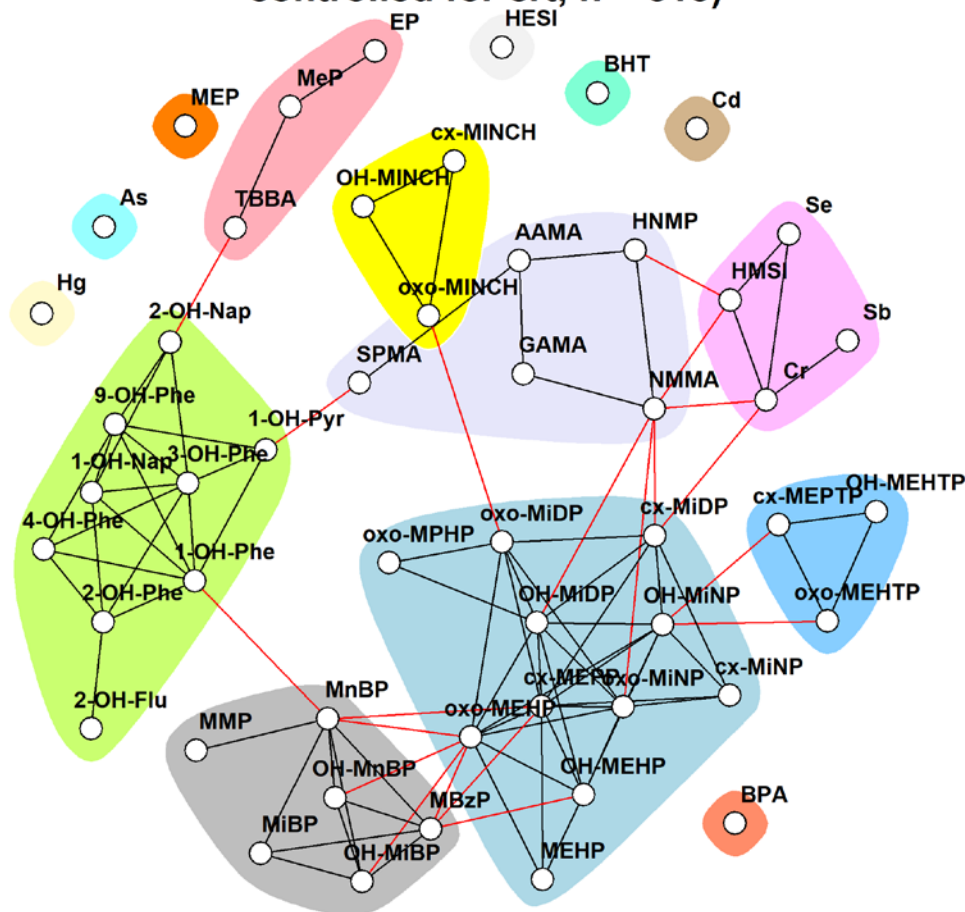
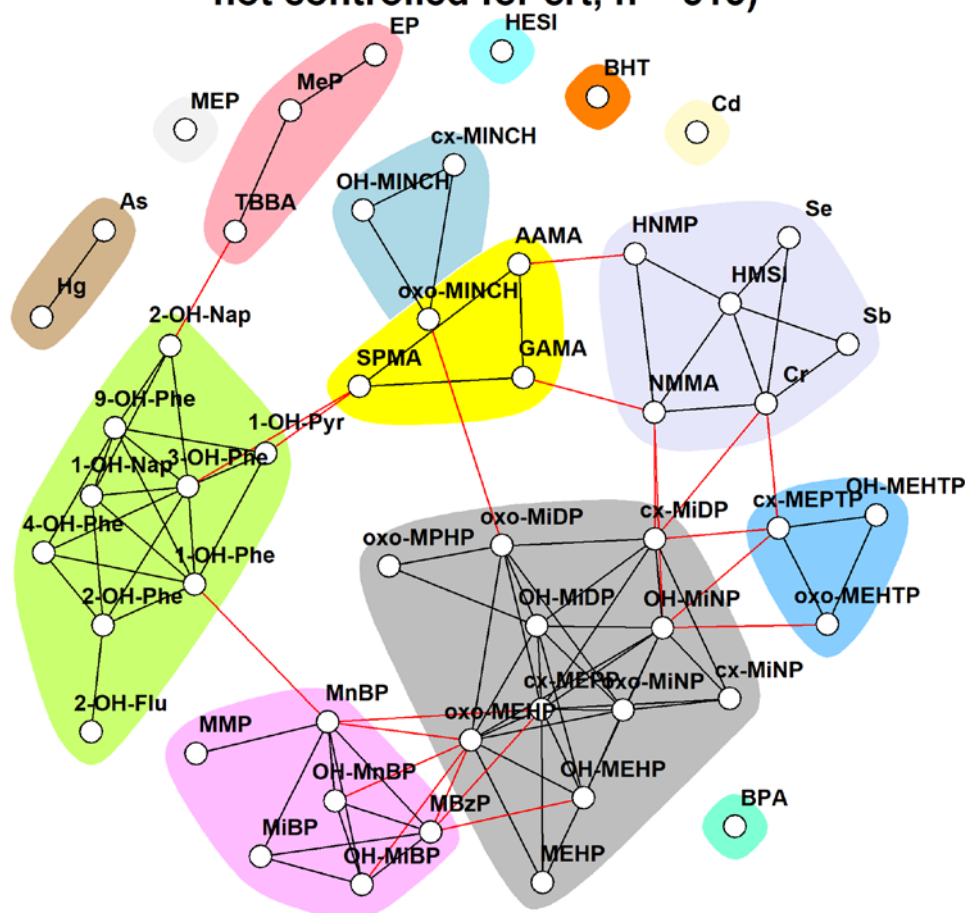


Figure 6.2.2: Network of GerES V with crt adjustment and control; controlled for other covariates



### Network GerES V (subsample, adjusted but not controlled for crt, n = 515)



**Figure 6.2.3: Network of GerES V with crt adjustment but not controlled for crt in multivariate analyses; controlled for other covariates**

The differences between the network on crt-adjusted and crt-controlled (Fig 6.2.2) versus only crt-adjusted (Fig 6.2.3) data are negligible. The major differences are the additional community of mercury and arsenic, and distribution of acrylamide, aprotic solvents, selenium, chromium, antimony, NMMA, and SPMA into two different communities (plum and yellow in the non-crt-controlled network in Figure 6.2.3).

However, not correcting for creatinine in any form results in considerably different communities. As can be seen in Figure 6.2.4 (networks not adjusted or controlled for crt), three heavily interrelated communities (green, pink, grey) mask the detailed communities detected in the network when adjusting and controlling for crt, possibly due to a similar degree of dilution being reflected in stronger correlations.

These findings indicate that it is important to correct for creatinine when aiming at analysing mixtures and at least to adjust the compounds concentrations by it, but yet need to be confirmed by further studies.



### Network GerES V (subsample, not adjusted or controlled for crt, n = 503)

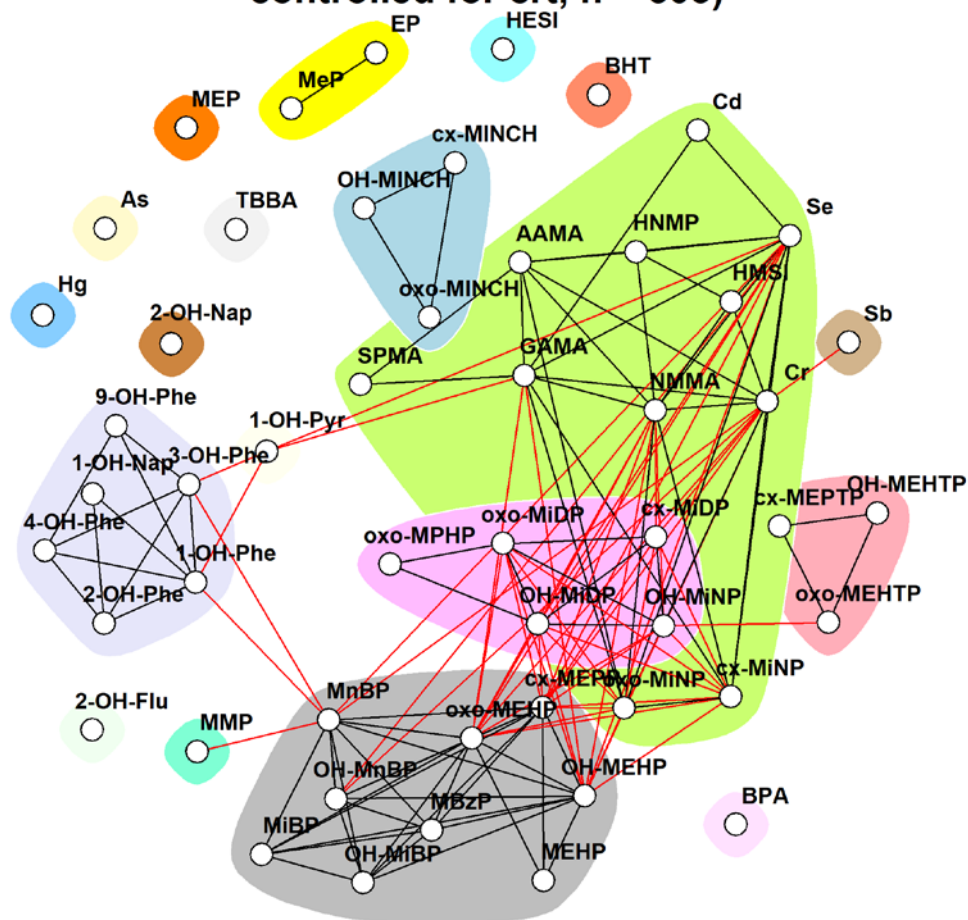
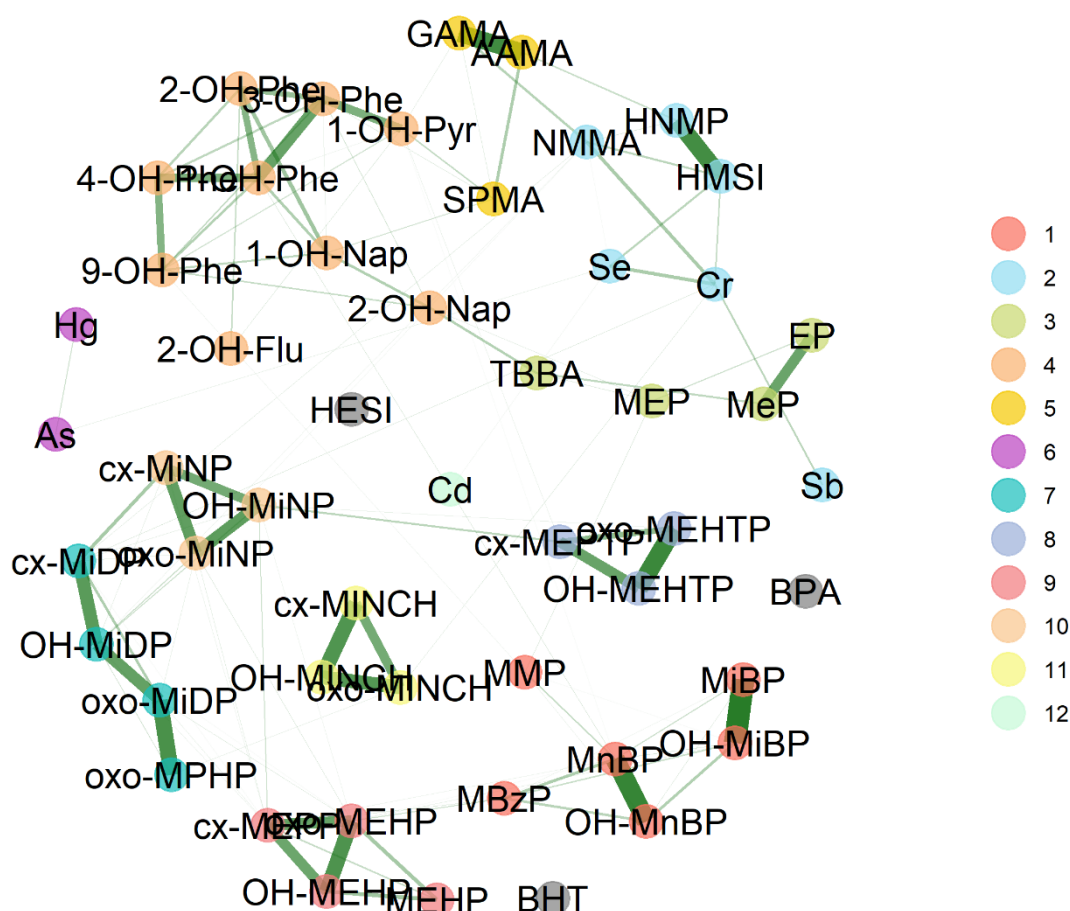


Figure 6.2.4: Network of GerES V without any crt correction; controlled for other covariates

In a second step, the additional method of weighted networks where the strength of links between compounds can be assessed (section 5.4.2). Figure 6.2.5 indicates that links were stronger (i.e., thicker lines) within a parent compound or substance group and that the strongest links were observed within acrylamide, aprotic solvents, parabens EP and MeP, DINCH, DEHTP, and several phthalates.

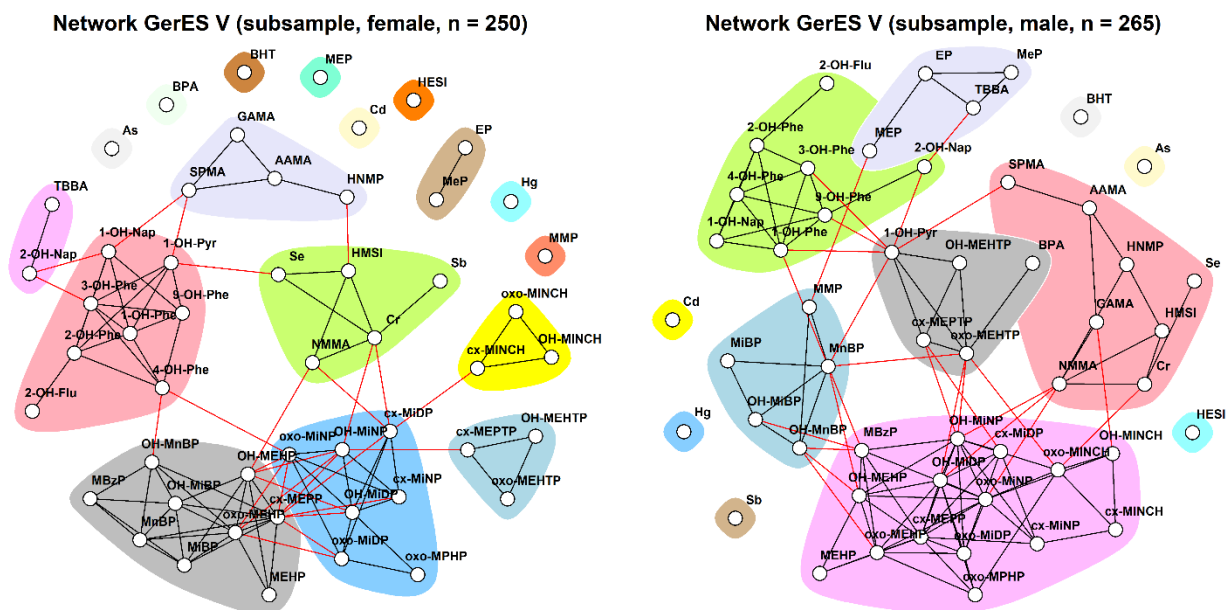


**Figure 6.2.5: Weighted network of GerES V with bootstrap technique applied, with crt adjustment and crt control; controlled for other covariates**

To check the impact of major exposure determinants on mixtures, networks were stratified by sex (female vs. male), smoking status (smoking vs. non-smoking), age (age 10 and younger vs. older than 10 years), ISCED of the household (low and medium vs. high), and BMI (lower than 25 vs. higher than 25). Concentrations were adjusted by and controlled for creatinine.

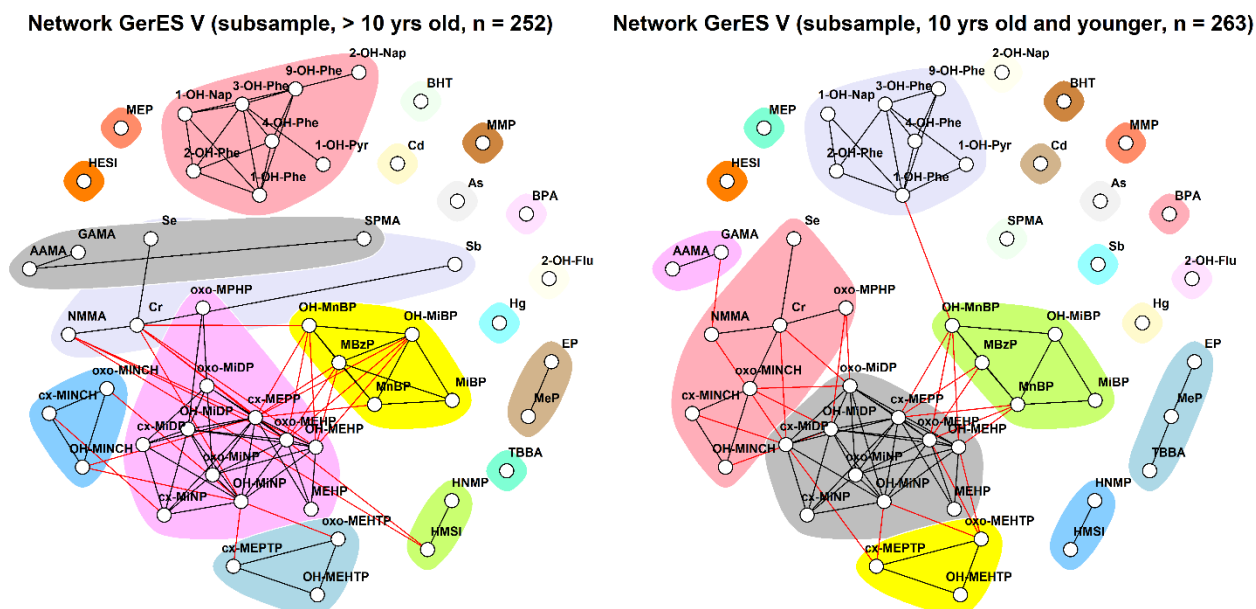
Since only 12 participants were smoking, network analyses might be specific to those participants rather than the group of smoking children and were not further explored. Networks of the non-smoker group ( $n = 503$ ; not shown here) looked very similar to those of the total sample.

A comparison of the boys' and girls' networks (Figure 6.2.6) indicates that there are more differences than similarities. While the girls have 9 communities and 8 standalone compounds, the boys only have 6 larger communities and 6 standalone compounds. The groups share the PAHs and DEHTP communities to some degree. There are, however, considerable differences in the phthalate, acrylamide, and element communities. In boys the phthalates form different communities together with DINCH (BBzP, DEHP, DiNP, DiDP, DINCH (lavender) vs. DiBP, DnBP, and DMP (blue)) than in girls (BBzP, DEHP, DnBP, DiBP (grey) vs. DiNP, DiDP (blue) vs. DINCH in an extra community (yellow)).



**Figure 6.2.6: Stratification of GerES V by sex; controlled for crt, smoking status, age, education, BMI**

Figure 6.2.7 shows stratified networks by the median age which was 10 years. Both children older than 10 and 10 years old and younger show a community each for PAHs, 2 aprotic solvents (HNMP & HMSI), and DEHTP metabolites. Interestingly, DINCH forms a community with NMMA, and elements selenium and chromium (salmon) in younger but not older children in which each element and DINCH belong to three separate communities. In addition, the parabens – a sometimes observed standalone community – form their community with TBBA in the younger group.



**Figure 6.2.7: Stratification of GerES V by median age (10 years old), controlled for crt, sex, smoking status, education, BMI.**



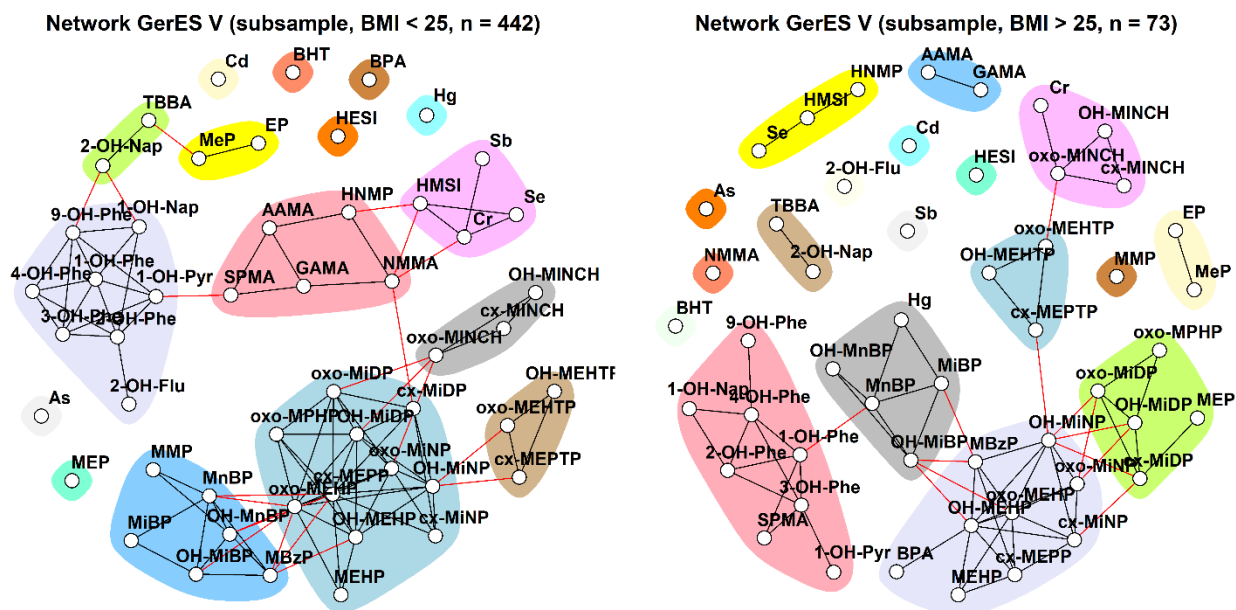
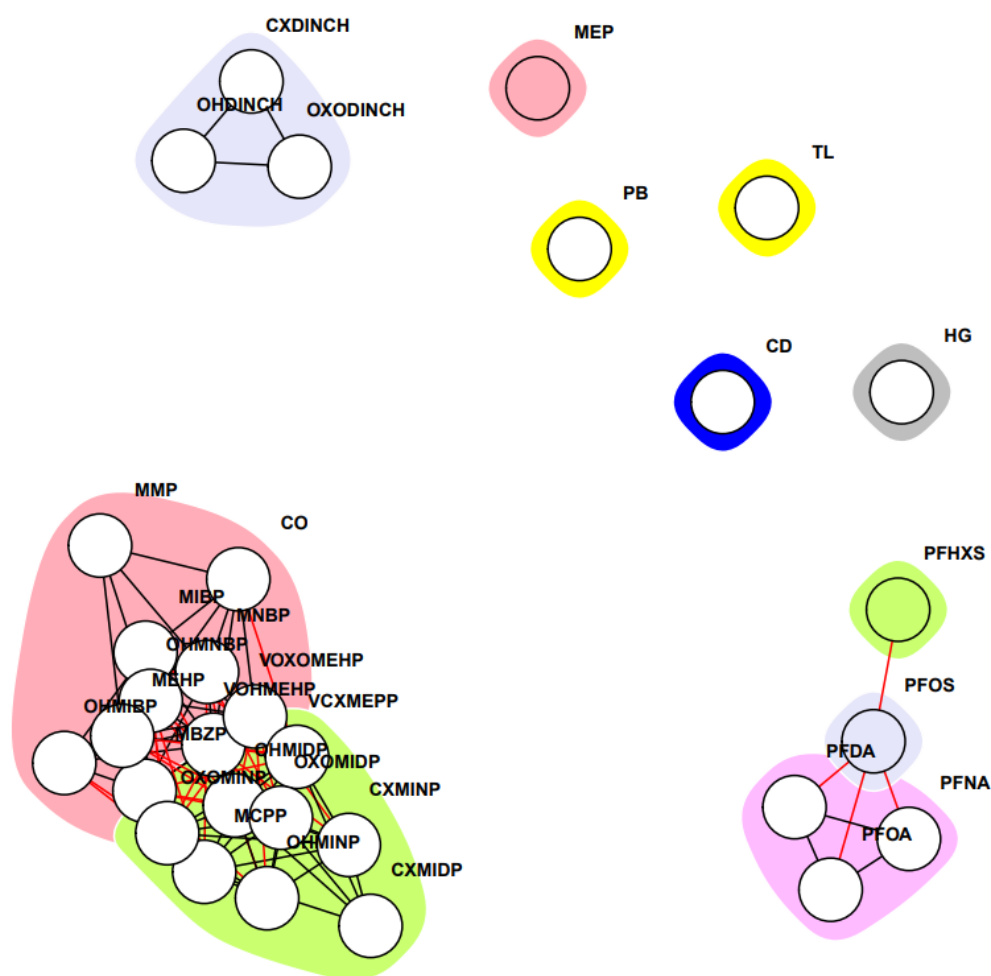


Figure 6.2.9: Stratification of GerES V by BMI, controlled for crt, sex, smoking status, age, and education

### 6.2.3 BIOAMBIENT.ES (Spain)

Figure 6.2.10 shows the results of the unweighted network analysis of 31 biomarkers in the BIOAMBIENT.ES subsample of 163 participants. It shows that several substances (metals mercury, cadmium and lead and MEP, metabolite of the phthalate DEP) were not part of any community. DINCH metabolites appeared grouped in an isolated community. Two communities were found among phthalate metabolites showing dependence among themselves and with cobalt as seen earlier in the Pearson correlations (heatmap and circos plot). For the PFAS, the community formed by PFOA, PFNA and PFDA was connected to PFOS and this one to PFHXS.



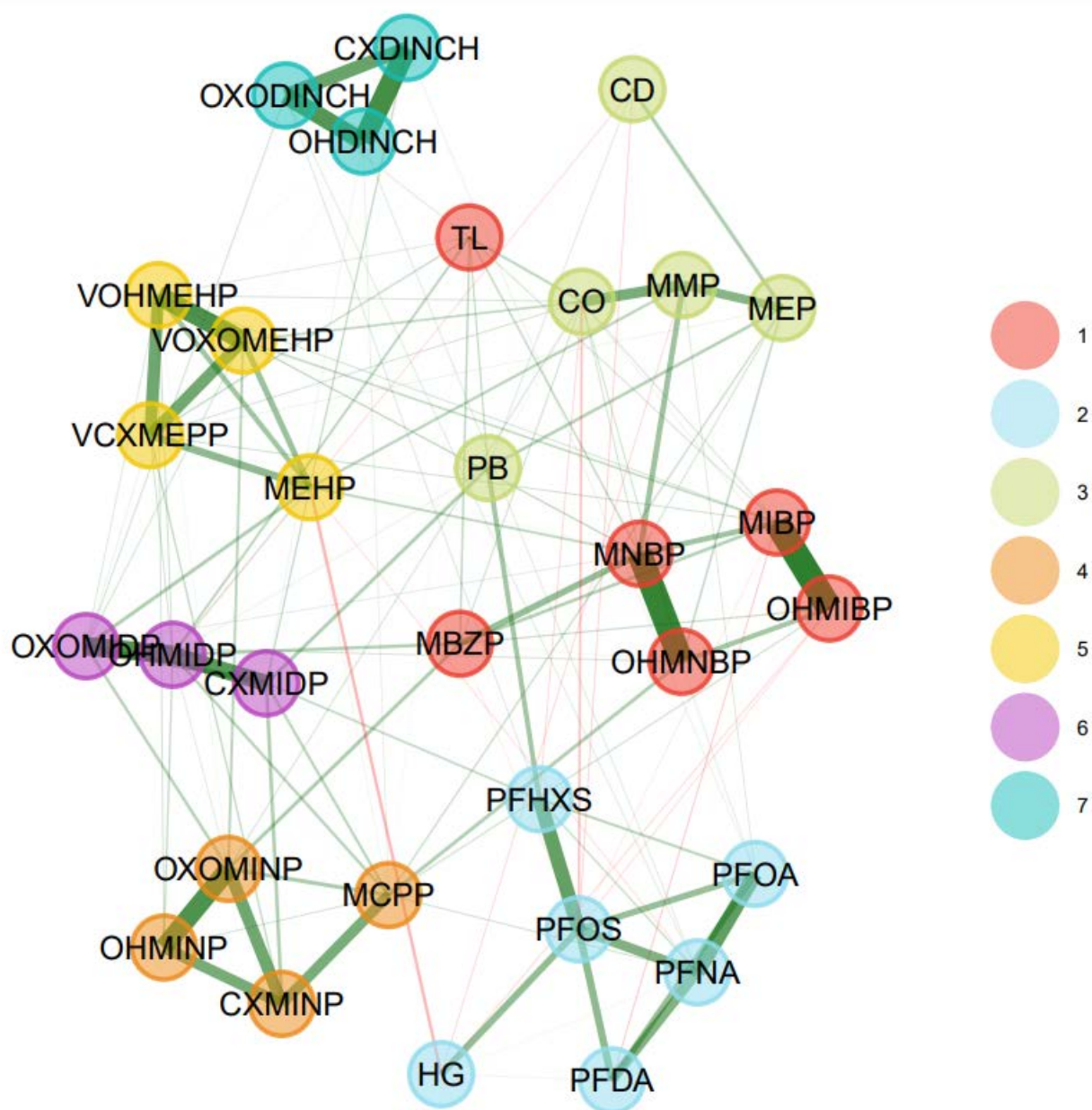


**Figure 6.2.10: Unweighted network of 31 biomarkers in the BIOAMBIENT.ES subsample of 163 participants**

Differences between the unweighted and the weighted network analyses were observed. In the weighted networks, we obtained seven communities showing more dependencies compared to the communities identified with unweighted analysis. For example, the PFAS now form one community, where mercury is also part of that community through a link to PFOS; also, cadmium is linked to MEP, and lead to PFHXS, not observed in the unweighted network analysis. We observed stronger links within the substance groups DINCH and PFAS, and not between metals. Although cobalt is still linked to MMP, this is not the case for thallium, for which links with other biomarkers appeared weak.

We also observed separate grouping within parent compounds in the case of phthalates: DiBP metabolites (MiBP, OH-MiBP), DEHP metabolites (MEHP, OH-MEHP, oxo-MEHP, cx-MEPP), DiNP metabolites (OH-MiNP, oxo-MiNP, cx-MiNP), DiDP metabolites (OH-MiDP, oxo-MIDP, cx-MIDP). However, for DnBP, two metabolites (MnBP and OH-MnBP) grouped together whereas MCPP was grouped together with the DiNP metabolites showing strong links to cx-MiNP.





**Figure 6.2.11: Weighted network analysis of 31 biomarkers in the BIOAMBIENT.ES subsample of 163 participants**

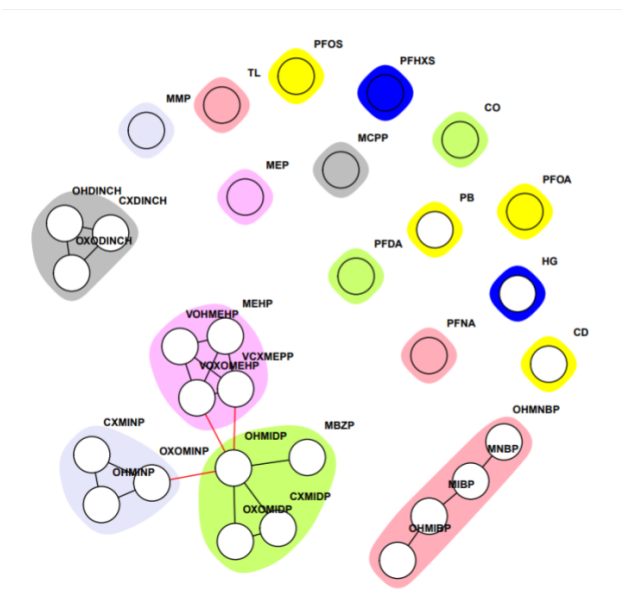
As indicated earlier, we performed stratified analysis for relevant determinants (sex, age, smoking status, ISCED, BMI, and fish consumption), to evaluate their effect on the network and accompanying communities. In these stratified analyses we found that the overall distribution of communities was similar to the one observed in the unweighted network for the general population. Metals tend to appear as standalone compounds, DINCH metabolites form a distinct community, as generally do PFAS metabolites and phthalates metabolites (two main big communities plus MEP). Communities of these two latter substance groups present some differences depending on the analysis. MEP, metabolite of the phthalate substance DEP, always appears separate from other phthalates and substances.

In the stratification by sex (Figure 6.2.12) we observed differences in the PFAS group. Whereas for males all substances form a community, in the females the different metabolites appear as standalone substances. Another difference is related to the metal cobalt, which in males seems to

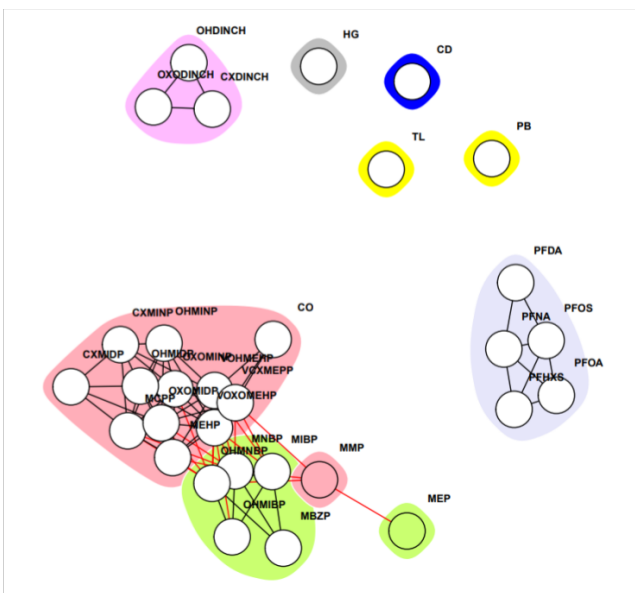
D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 50

form a community with phthalate metabolites of DEHP, DiNP, and DiDP not observed in females. Further differences between females and males relate to the main communities formed by phthalate metabolites and the dependencies across them. In females we find that metabolites of DMP (MMP) and DEP (MEP) appear as standalone substances. MCPP, metabolite of DnBP, also appears alone. We also find three communities showing dependencies between them, a first one includes metabolites of DEHP, a second one includes metabolites of DiNP and the third one composed of metabolites of DiDP and BBzP. A further community, with no dependency to other phthalates, includes metabolites of DiBP and DnBP. In males, the distribution of the phthalate metabolites is in two communities, a bigger community formed by metabolites of DEHP, DiNP and DiDP together with MCPP, metabolite of DnBP, and the second one by metabolites of DnBP, DiBP and BBzP. In this case, MMP shows dependency to both phthalate communities and MEP to the latter.

**Females (N = 86)**



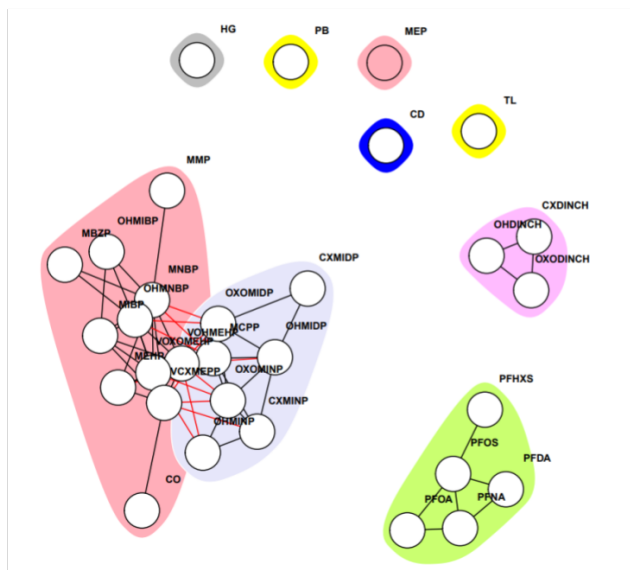
**Males (N = 77)**



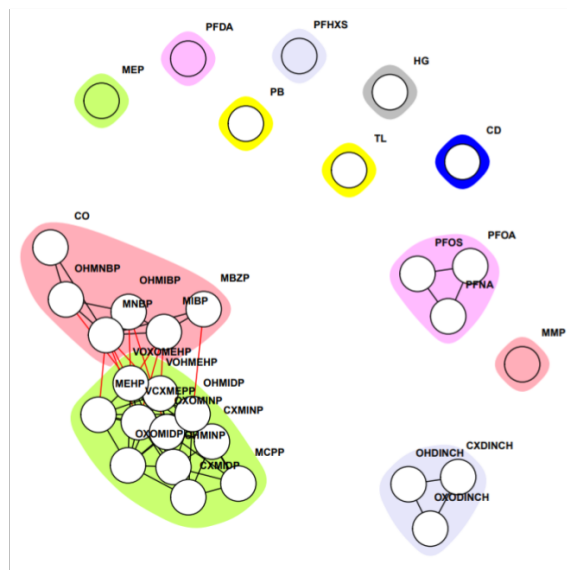
**Figure 6.2.12: Unweighted network of BIOAMBIENT.ES stratified by sex**

In the stratification by age (Figure 6.2.13), the main differences observed relate to the community formed by PFAS, single in the case of those individuals under 40 years of age, and in three (PFOS-PFOA-PFNA, PFDA and PFHxS) not related in those over 40. In this latter group, the phthalate MMP appears as a standalone metabolite.

**<= 39 (N = 92)**



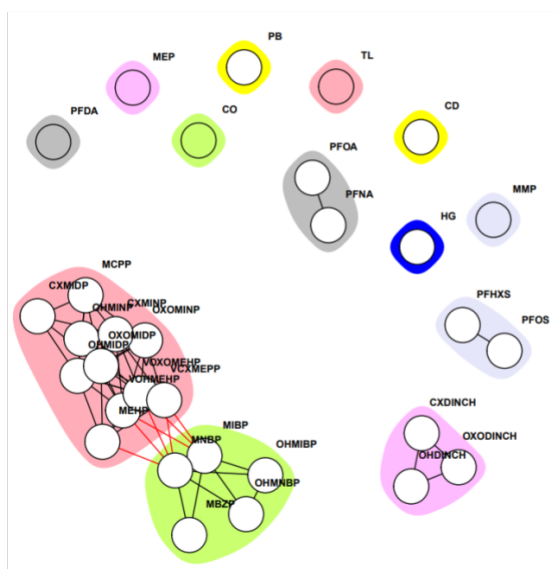
**>= 40 (N = 71)**



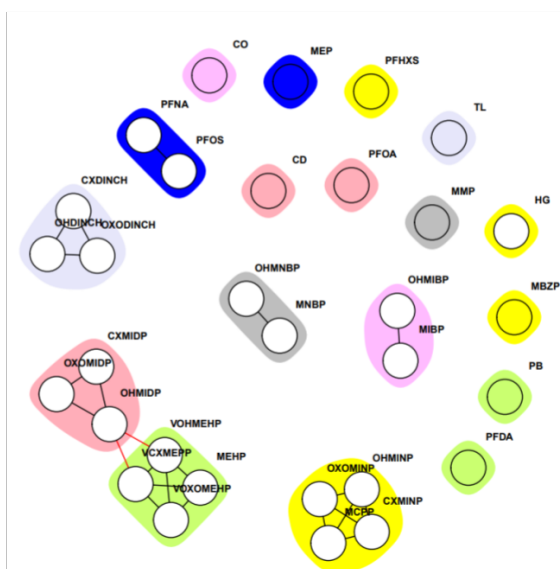
**Figure 6.2.13: Unweighted network of BIOAMBIENT.ES stratified by age**

In the stratification by smoking status (Figure 6.2.14), main differences between the non-smoking and the smoking group are the distribution of communities in PFAS and phthalates. PFAS are in 3 unrelated communities composed of PFDA, PFOA- PFNA and PFHxS- PFOS, respectively, in non-smokers, and in the smoking group PFNA-PFOS appear in a community whereas PFHxS, PFOA and PFDA appear as standalone substances. For phthalates, in the non-smoking group we find the metabolites of DEP (MEP) and DMP (MMP) as standalone substances and two communities composed of metabolites of DEHP, DiNP and DiDP plus MCP (metabolite of DnBP) in a bigger community and metabolites of BBzP, DiBP and DNBP in the smaller one. In the smoking group, each substance appears in a separate community, with the exception of MCP, metabolite of DnBP, which appears in the same community as the metabolites of DiNP.

**Non-smoking (N = 96)**



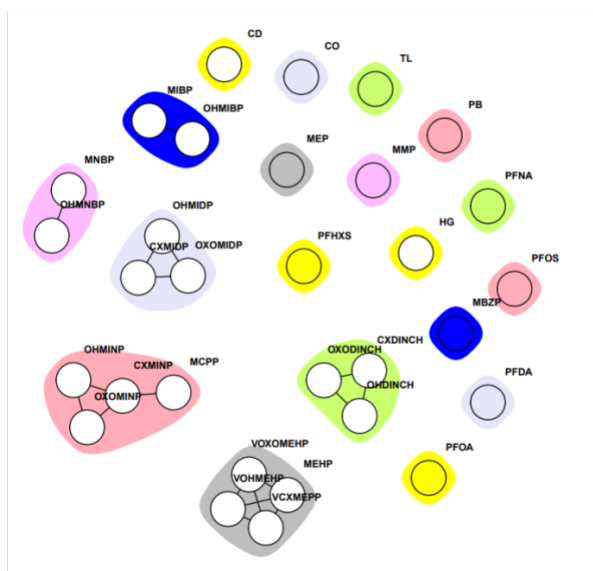
**Smoking (N = 67)**



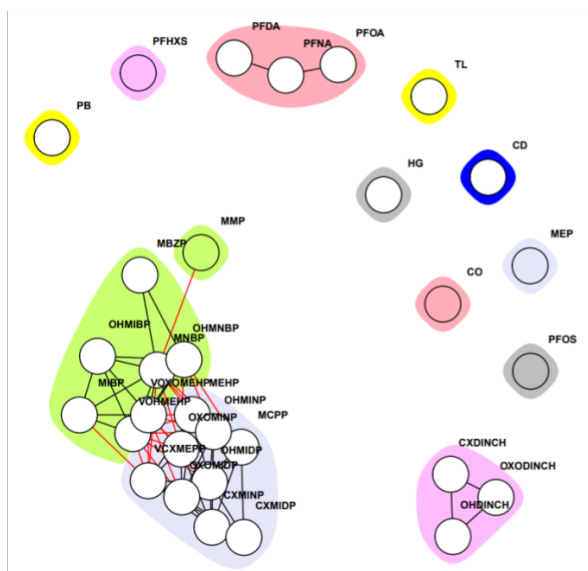
**Figure 6.2.14: Unweighted network of BIOAMBIENT.ES stratified by smoking status**

Figure 6.2.15 shows the stratified networks by education level. We found differences in the dependencies amongst PFAS, which appear as standalone substances in the low-medium ISCED level whereas PFOA-PFNA-PFDA form a community in the high ISCED group. For phthalates, in the low-medium ISCED level each substance appears in a separate community, with the exception of MCPP, metabolite of DnBP, which, as seen earlier, appears in the same community as the metabolites of DiNP, while in the high ISCED group, there are two main communities showing dependencies between them.

#### Low-medium ISCED (N = 103)



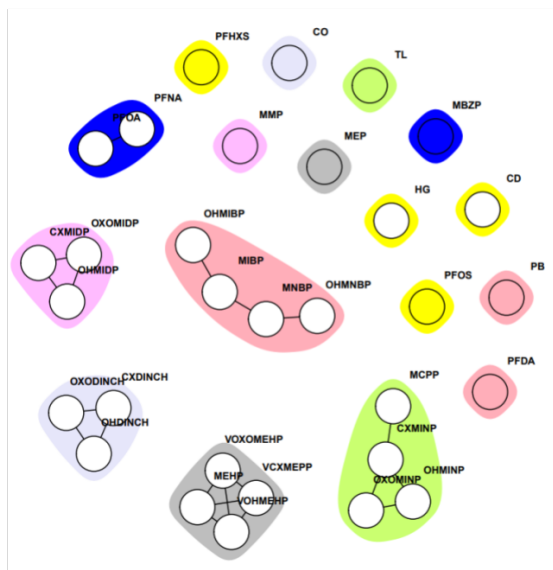
#### High ISCED (N = 49)



**Figure 6.2.15: Unweighted network of BIOAMBIENT.ES stratified by education (ISCED level)**

In the stratification by weight (Figure 6.2.16), we found differences again in the distribution of PFAS. In the overweight group they were grouped in one community whereas in the normal weight group we found PFHxS, PFOS and PFDA as standalone and only PFOA-PFNA formed a community. There are also differences in the distribution of the phthalate metabolites into the communities, they are found as individual substances in the normal weight group, whereas they form communities and show dependencies between them in the overweight group. In this group, cobalt shows dependencies to the phthalate communities not observed in the normal weight group.

### Normal weight (N = 78)



### Overweight (N = 83)

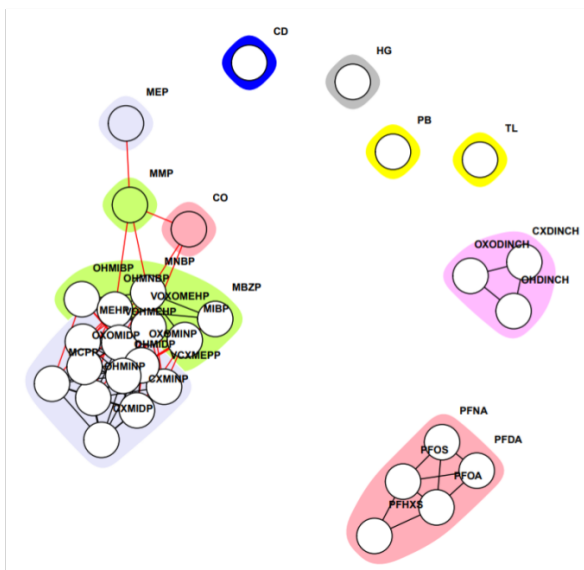
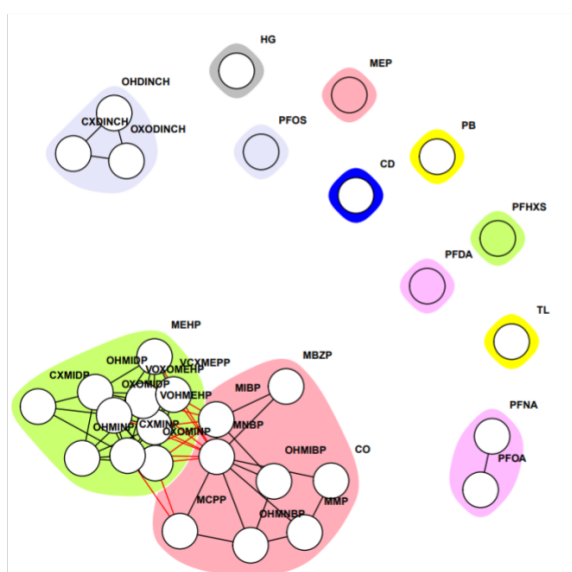


Figure 6.2.16: Unweighted network of BIOAMBIENT.ES stratified by BMI

When evaluating the effect of fish consumption (Figure 6.2.17), we observed less communities, which were mainly composed by substance groups, in the high fish consumption group. Here thallium is included in a phthalate community including metabolites of BBzP, DiBP, DnBP and DEHP, which shows dependencies to the other phthalate communities. Similarly, associations were observed in the initial Pearson correlations and had not been observed in any of the other stratifications analysed before for the BIOAMBIENT.ES subsample. We found again differences in the communities formed by PFAS, which in the high fish consumption group formed a unique community whereas in the low fish group we found four (PFNA-PFOA, PFOS, PFHxS and PFDA). Cobalt formed a community with phthalate metabolites in both groups.

### Low fish (N = 67)



### High fish (N = 94)

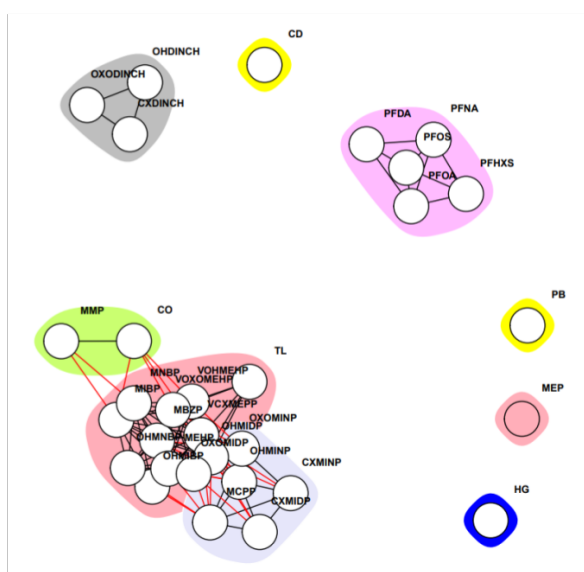


Figure 6.2.17: Unweighted network of BIOAMBIENT.ES stratified by fish consumption

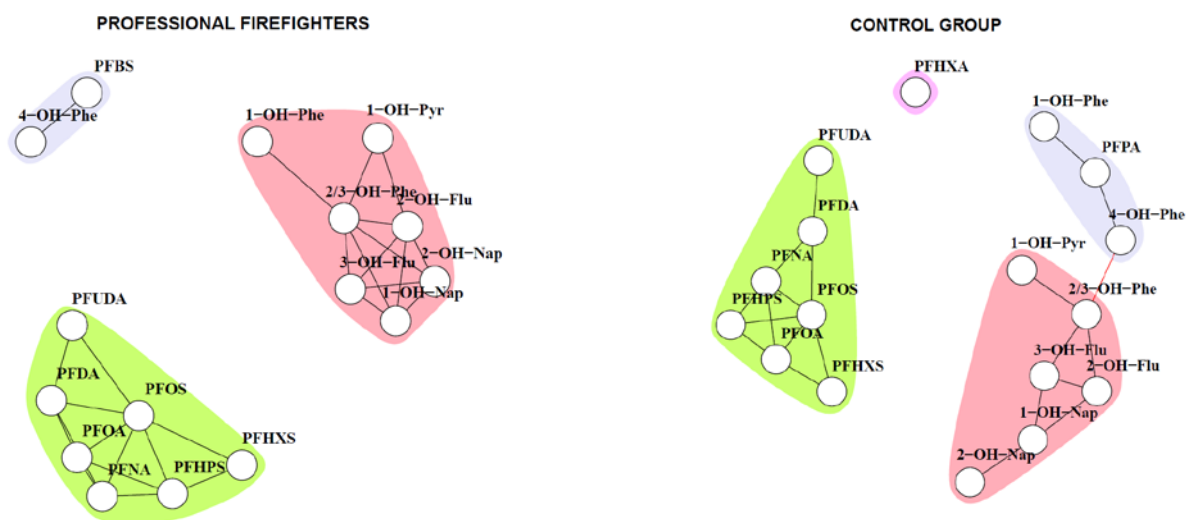


D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 54

#### 6.2.4 CELSPAC - FIREexpo (Czech Republic)

Figure 6.2.18 shows the network analysis for the professional firefighters (n = 52) and the control group (n=55) of the CELSPAC – FIREexpo study. The data were corrected for age and BMI. In this study, all participants were non-smoking males, the information about the education level was not available in this study. Note that the set of substances differ between the two groups due to differences in percentage detects above LOD/LOQ, which with the 60 % cut-off led to small differences in which substances were in- or excluded from the network analysis.

In the professional firefighters, most PFASs and OH-PAHs created a community with the compounds of the same substance group (green and red). Except for the 4-OH-Phe and PFBS which created a stand-alone community (lavender). In the control group, four communities were detected: A community of majority PFAS (green), majority OH-PAHs (red), and then a community of 1-OH-Pyr, PFPA and 4-OH-Phe (lavender). The last two communities were interconnected by the edge between the 4-OH-Phe and the 2/3-OH-Phe. PFHXA was not part of any community.



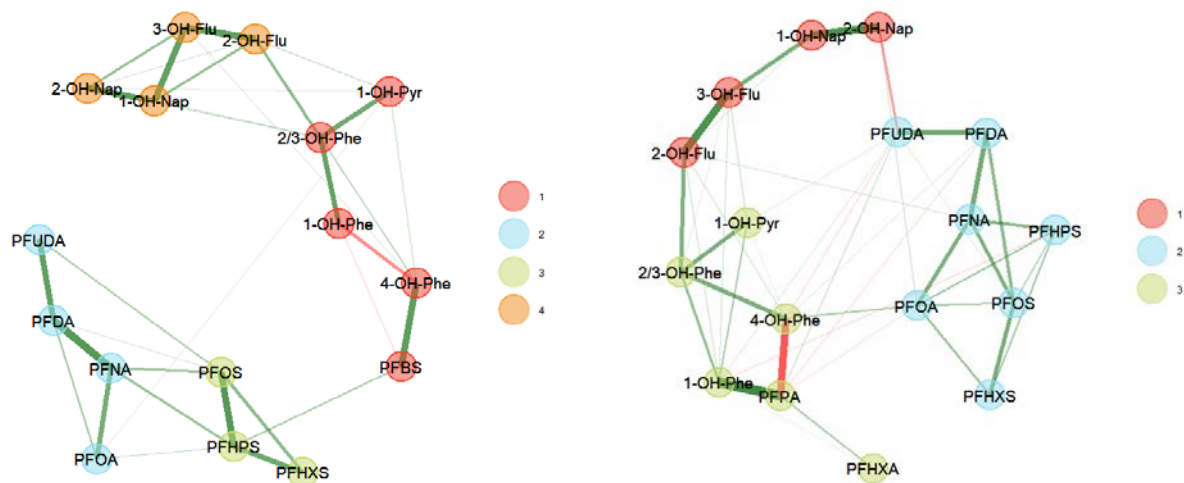
**Figure 6.2.18: Unweighted network for CELSPAC – FIREexpo showing communities of biomarkers for professional firefighters (left) and the control group (right). The data were corrected for age and BMI. Urinary markers (OH-PAHs) were adjusted for creatinine.**

In the weighted networks (Figure 6.2.19), the strength of the links (dependencies) was assessed and visualised. The thicker the line, the stronger the link. In the professional firefighters, the analysed compounds are creating four communities. First containing PFUDA, PFDA, PFNA, and PFOA (blue), and the second (PFOS, PFHPS and PFHXS; lavender), but it can be seen that all PFASs except for the PFBS are connected. Then there are two communities of OH-PAHs, one made of fluorenes and OH-naphthalenes (pink) and one of OH-phenanthrenes, and OH-pyrene, and PFBS, which is more strongly correlated with 4-OH-Phe than with other PFASs. In the control group, there are three communities, one made of OH-naphthalenes and OH-phenanthrenes (lavender), second made of majority PFASs (blue) and a third one made of OH-phenanthrenes, OH-pyrene, PFPA and PFHXA. PFHXA is weakly linked to PFPA and 1-OH-Phe, almost creating a stand-alone community. The links between the compounds of different communities are more present (the communities are not that strictly separated as in the professional firefighters).

The unweighted networks and weighted networks are mostly comparable. Moreover, they are in agreement with observations from correlation heatmaps and circos plots, that in the professional firefighters, the intracommunity links are strong, and there are weak inter-community links,



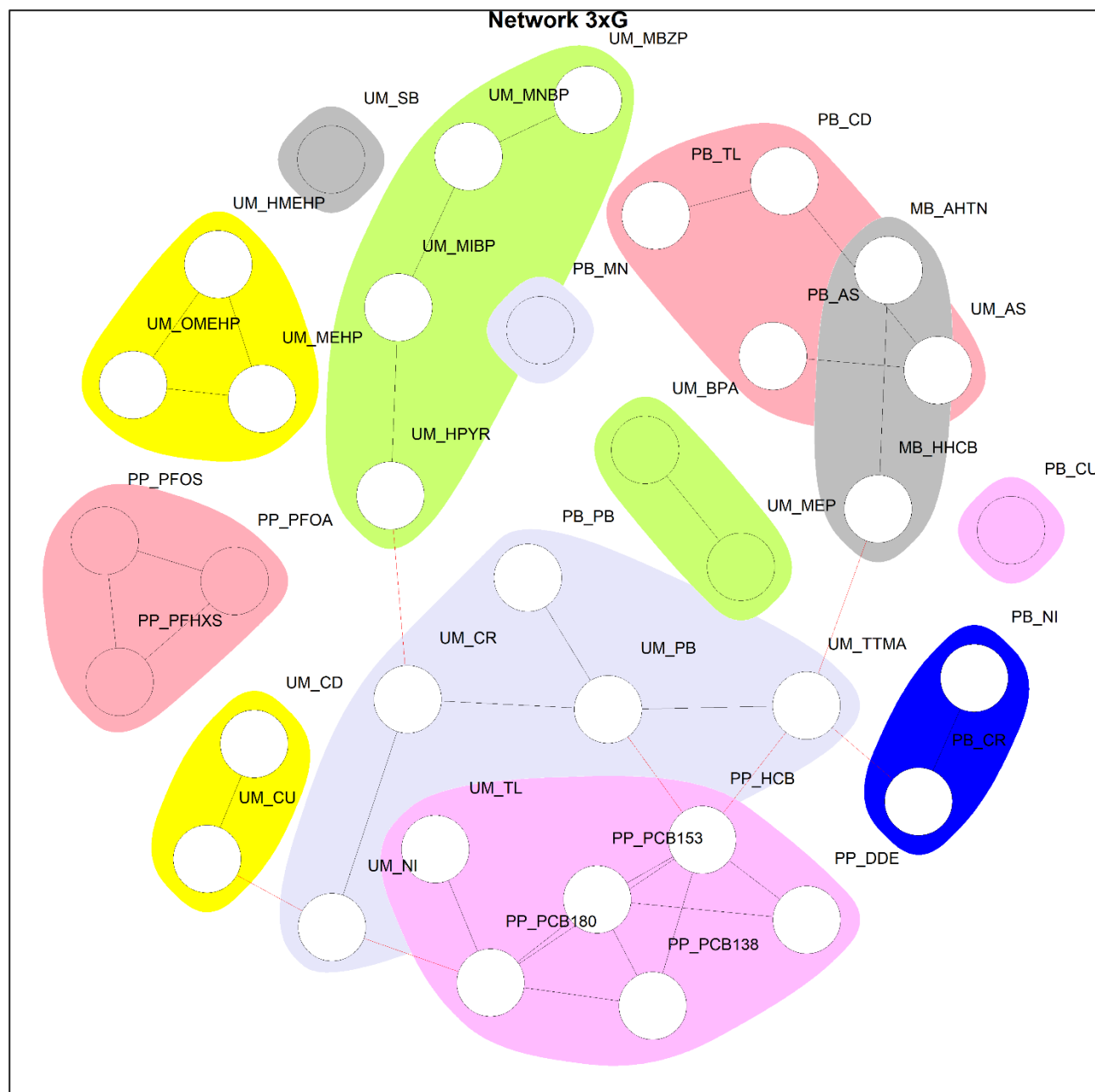
resulting in more strictly separated PFAS and OH-PAHs communities, while in the control group more inter-communities links are present, meaning that sometimes the substances from different groups are more linked together than the substances within the same group.



**Figure 6.2.19: Weighted network for CELSPAC – FIREexpo professional firefighters (left) and the control group (right)** The data were corrected for age and BMI. Urinary markers (OH-PAHs) were adjusted for creatinine.

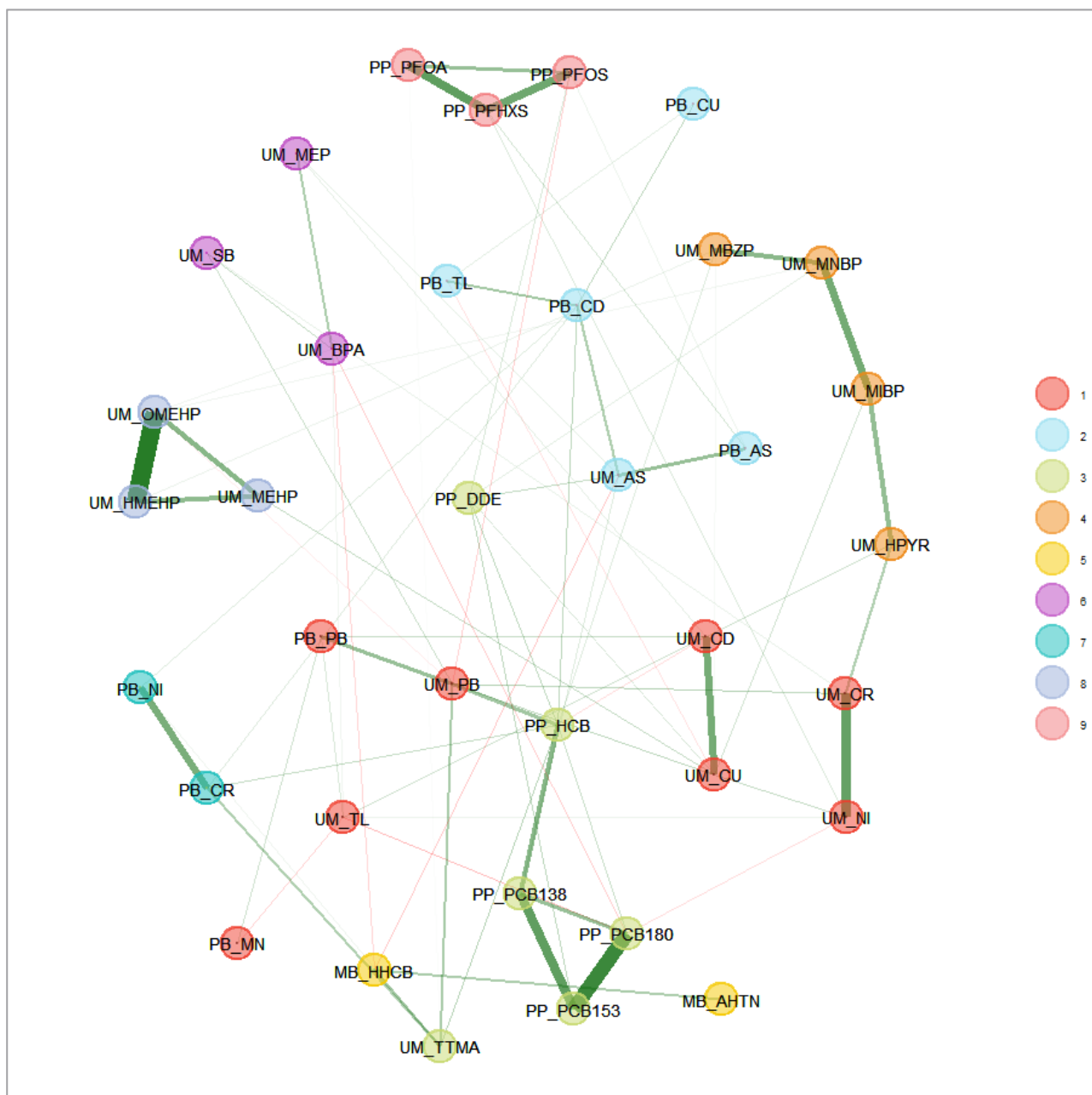
## 6.2.5 3xG (Belgium)

Network analyses were conducted to identify groups of highly connected biomarkers, i.e., so-called "communities". Figures 6.2.20 - 6.2.22 show the communities for the 3xG subset. Three methods to identify communities were used, which include unweighted network analysis as well as weighted network analysis using both the Glasso and the Bootstrap algorithms.



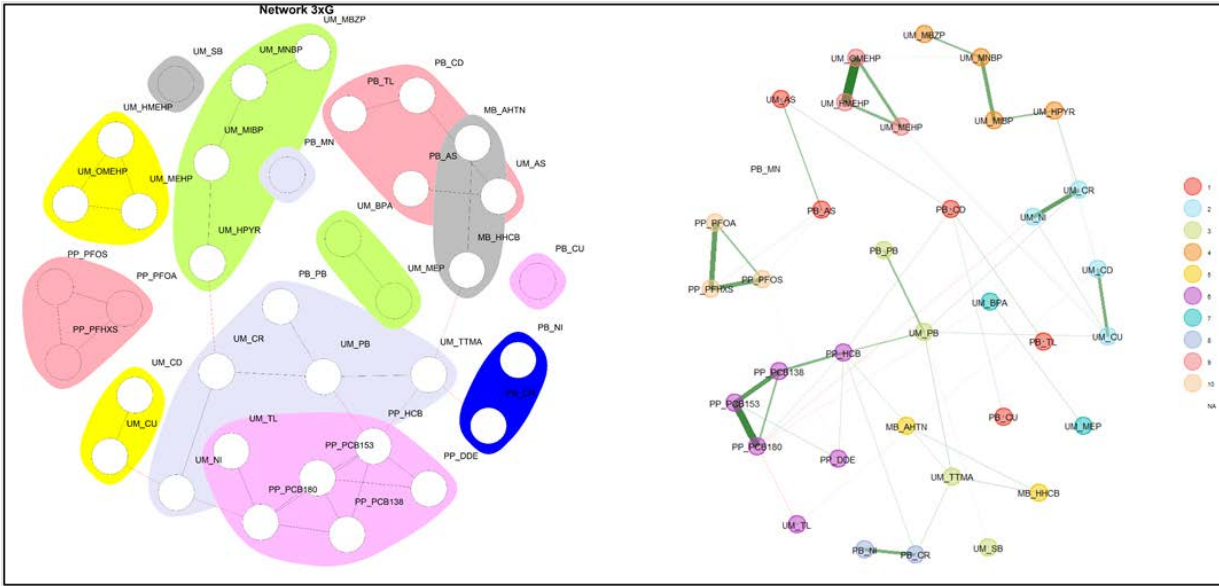
**Figure 6.2.20: Network analyses of 3xG, showing different communities of biomarkers of exposure using the unweighted network approach and Glasso algorithm. Matrices in which biomarkers were measured are morning urine of the mother (UM), whole blood of the mother (MB), cord blood whole blood of the newborn (PB) and cord blood plasma of the newborn (PP).**





**Figure 6.2.22: Network analyses of 3xG, showing different communities of biomarkers of exposure using the weighted network approach and Glasso algorithm. Matrices in which biomarkers were measured are morning urine of the mother (UM), whole blood of the mother (MB), cord blood whole blood of the newborn (PB) and cord blood plasma of the newborn (PP).**

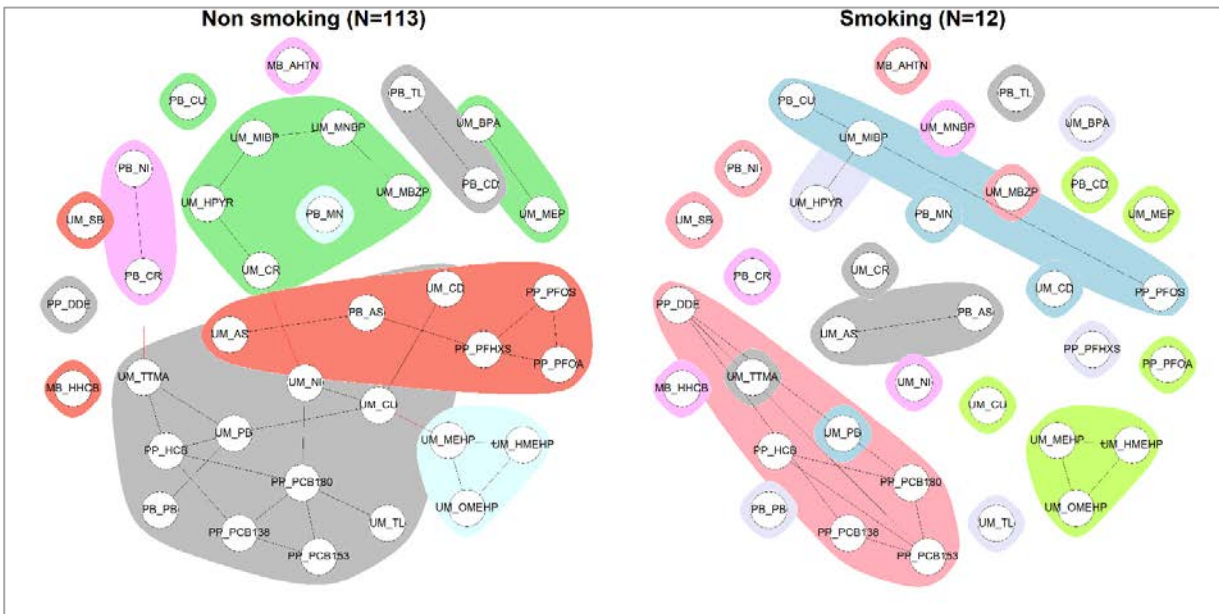
Figure 6.2.23 compares the unweighted and the weighted network, side by side. As can be observed, there are not many differences between both networks: for example, copper and cadmium (morning urine of the mother) form a community in the unweighted network, but are included in community 2 in the weighted network, together with nickel and chromium. Copper measured in blood is included in community 1 with other metals in the weighted network, while it appears separately in the unweighted network. We hypothesise that these changes might arise from the different parameters and algorithms used to construct the two networks.



**Figure 6.2.23: Unweighted versus Weighted (bootstrapped) network**

As a further step, unweighted networks were used to compare how different subpopulations may differ in communities. Determinants (see Table 5.3.4) had been chosen after deliberation within the working group. No interpretation of the results is done in this report, we limit ourselves to a description of the results.

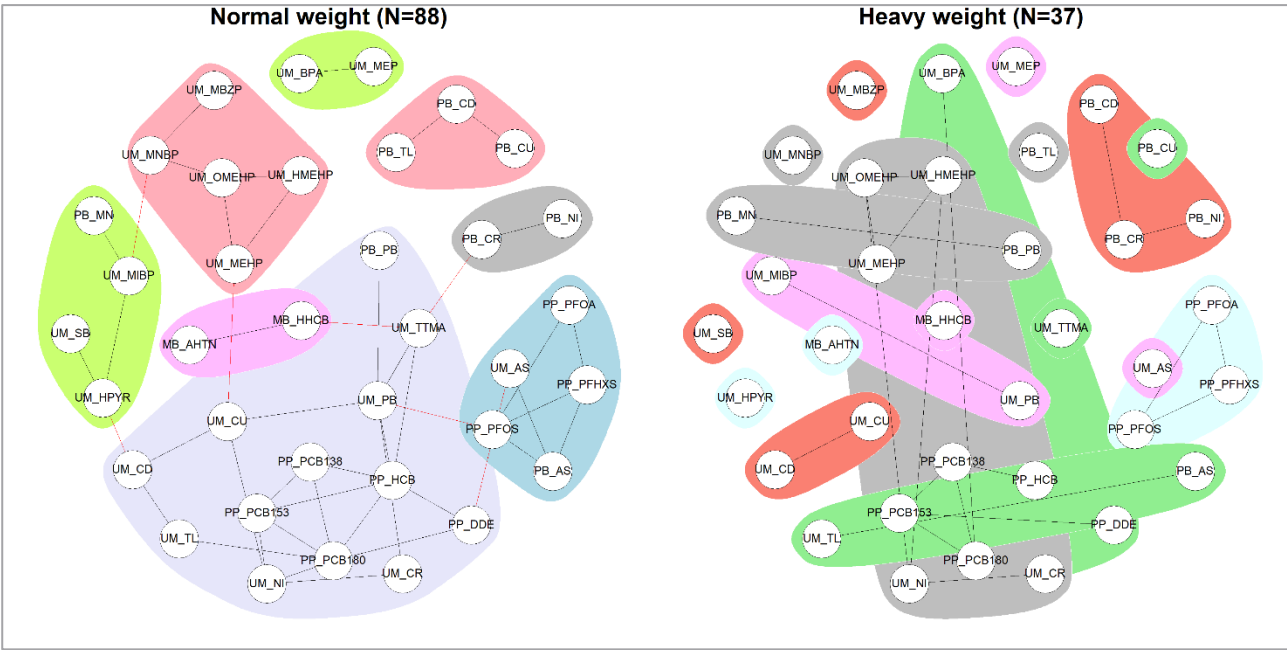
Figure 6.2.24 shows the different communities formed for the subset of participants for which the variable ‘smoking of the mother during pregnancy’ was answered positively or negatively. As shown in the figure, the number of observations for the group of mothers who smoked during pregnancy is quite small (N=12). Therefore, conclusions that might be extracted in the future from this figure should be considered with caution. Nonetheless, it can be noticed that communities observed in the non-smokers group (N=133) and in the whole group (see Figure 6.2.20) are different, suggesting an influence in the overall communities of those smoking participants.



**Figure 6.2.24: Unweighted network stratified for variable smoking of the mother during pregnancy. Matrices in which biomarkers were measured are morning urine of the mother (UM), whole blood of the mother (MB), cord blood whole blood of the newborn (PB) and cord blood plasma of the newborn (PP).**

Figure 6.2.25 represents the different communities formed between the groups for which BMI of the mothers is normal and BMI of the mothers is above 25 kg/m<sup>2</sup>. In this figure, differences can be observed in the PCB community, which includes TI, Cu and BPA for those mothers that belong to the high BMI category. The community that includes DEHP metabolites appear to be slightly different between the two groups as well, including Cr and Ni in the group of high BMI mothers. Furthermore, in the group of mothers within a normal BMI, the community formed by the PFASs included arsenic both measured in the urine of the mother and the cord blood of the newborn, which is not the case in the group of the mothers with a higher BMI.

Finally, Figure 6.2.27 represents the different communities formed between the group where mothers have a low fish consumption and the group where mothers have a high fish consumption, according to the categories described in Section 5. Differences in networks can be observed for communities formed by PCBs, which include copper measured in cord blood and DDE for the subset of high fish consumption. The community formed by CR and NI measured in the urine of the mother also appear to include different biomarkers between the two groups, including CU and CD for low fish consumption mothers, two biomarkers that form an individual community in the high fish consumption mothers. Interestingly, CR and NI measured in the cord blood of the newborn include also different biomarkers depending on the fish consumption group. Finally, TTMA and HHCB form a unique community in high fish consumption mothers, while they are in separate communities in low fish consumption mothers.

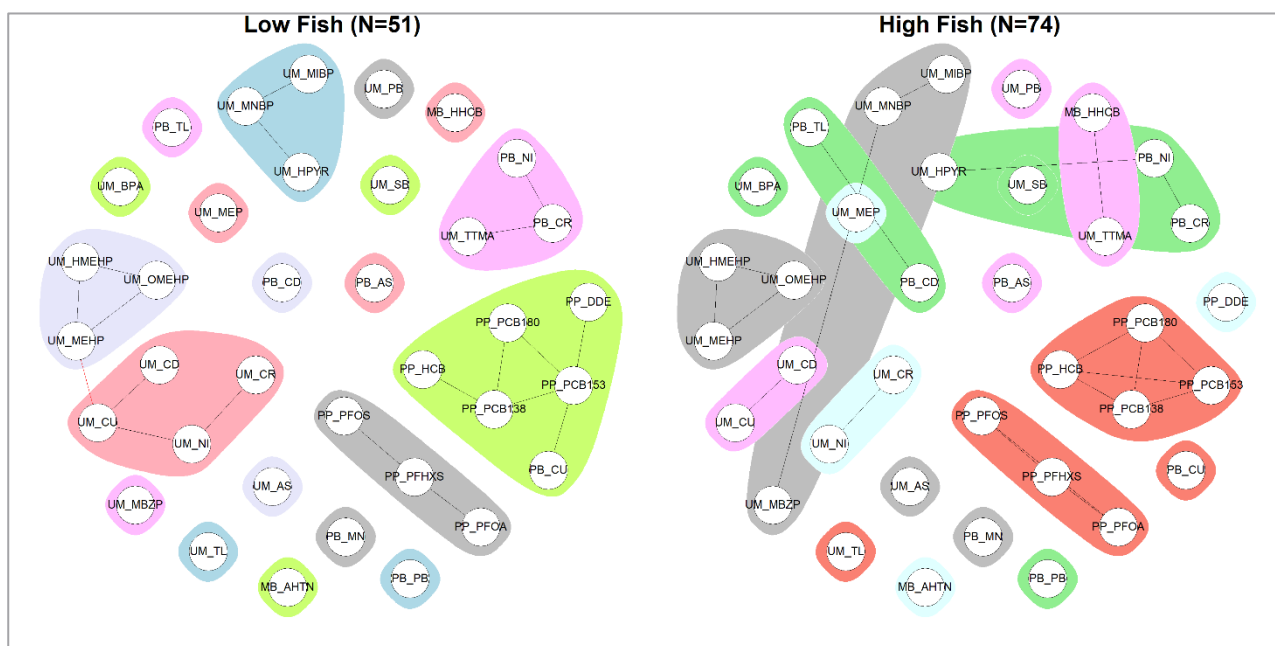


**Figure 6.2.25: Unweighted network stratified for variable BMI of the mother before pregnancy. Matrices in which biomarkers were measured are morning urine of the mother (UM), whole blood of the mother (MB), cord blood whole blood of the newborn (PB) and cord blood plasma of the newborn (PP).**

Figure 6.2.26 represents the different communities formed between the group where mothers have a low ISCED and the group where mothers have a high ISCED. Interestingly, the DEHP community is conserved throughout both groups, while that formed by the DIBP metabolites is not. Furthermore, communities that include heavy metals seem to not be equal between the two groups.







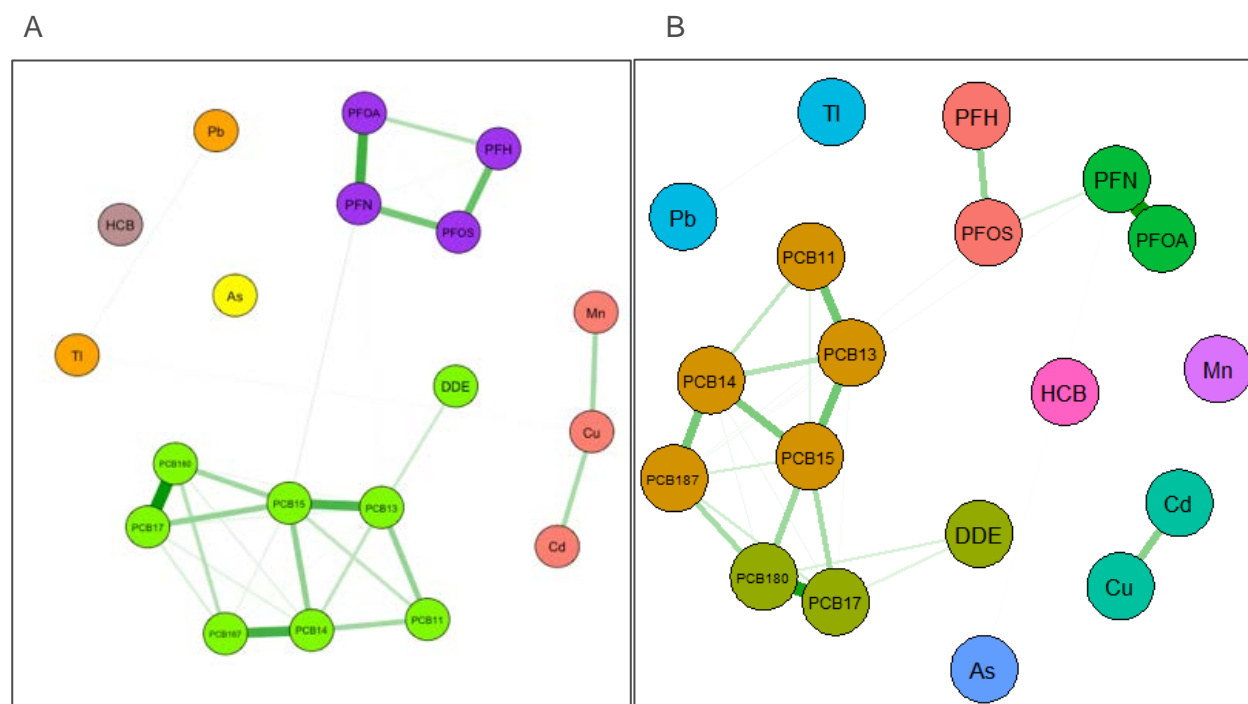
**Figure 6.2.27: Unweighted network stratified for variable fish consumption of the mother during pregnancy. Matrices in which biomarkers were measured are morning urine of the mother (UM), whole blood of the mother (MB), cord blood whole blood of the newborn (PB) and cord blood plasma of the newborn (PP)**

## 6.3 SNMU analysis

### 6.3.1 FLEHS results from SNMU analyses

#### 6.3.1.1 Comparison of datasets

As explained in part 5.2.1.1, in order to have only positive values, the initial dataset on residualised biomarker concentrations used in Ottenbros et al. (2021) underwent an extra exponential transformation. In a first step, the network analysis method is applied to the exponential residualised biomarker concentration dataset in order to check the impact on mixtures of this extra transformation (Figure 6.3.1). For this comparison the bootstrapped weighted network method (EGAnet) was used. This network is compared with results obtained with the initial residualised biomarker concentrations and available in the supplementary material from Ottenbros et al. (2021).

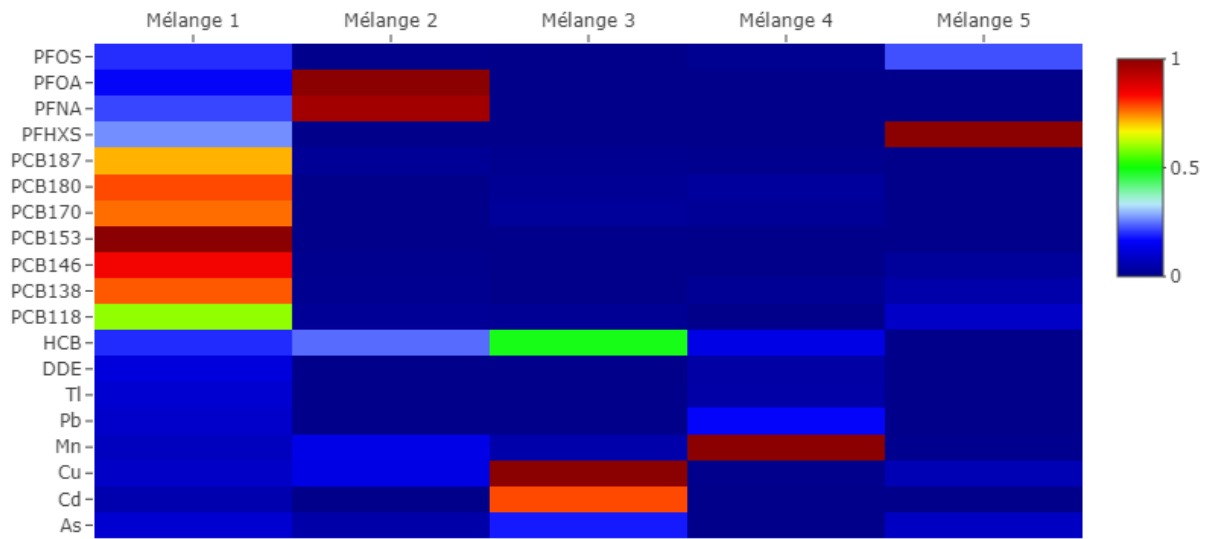


**Figure 6.3.1: Weighted networks from FLEHS III A. Residualised biomarker concentrations B. Exponential residualised biomarker concentrations**

The network obtained from the dataset created for the SNMU method (Figure 6.3.1B) is similar to the network presented in Ottenbros *et al.* 2021 (Figure 6.3.1A). Indeed, there are similar connections between the PCB substances, two metals (Cu and Cd), and the PFAS substances. However, we can notice some differences, such as the absence of a link between Cu and Mn, between PFHXS and PFOA, and between PCB138 and DDE using the exponential dataset. Moreover, p,p'-DDE is well connected to the PCB group, but via PCB170 and PCB180 instead of the PCB138 using the residualised biomarker concentrations. Despite these differences, the two-compared networks can be considered as equivalent because the same mixtures are identified in both when looking at the strongest connections. The impact of the exponential transformation on the dataset is considered as negligible. In order to compare only the methods and not their combined effect with the exponential transformation, results from SNMU method will be compared to the network obtained with the non-negative dataset obtained with the exponential transformation (Figure 6.3.1B).

### 6.3.1.2 Mixture identification with SNMU

Applying the SNMU approach on the exponential residualised biomarker concentrations leads to optimal values of 11, 17, 15, 10, 9, and 5 explaining from 91% to 72% of the total variance. The selection of the number of mixtures is done according to the procedure described in section 5.2.4, and results in  $k=5$ . This value seems to be a good compromise since the percentage of variance explained (72%) is still high, the number of mixtures is not too high compared to the low number of substances analysed and all mixtures are relevant when clustering is applied. Moreover, 5 mixtures are the optimal value closest to the number of mixtures identified on the unweighted network (3 mixtures if we consider the strongest connections). The obtained matrix U of dimensions 19×5 (19 studied substances, 5 identified mixtures) is presented in Figure 6.3.2.

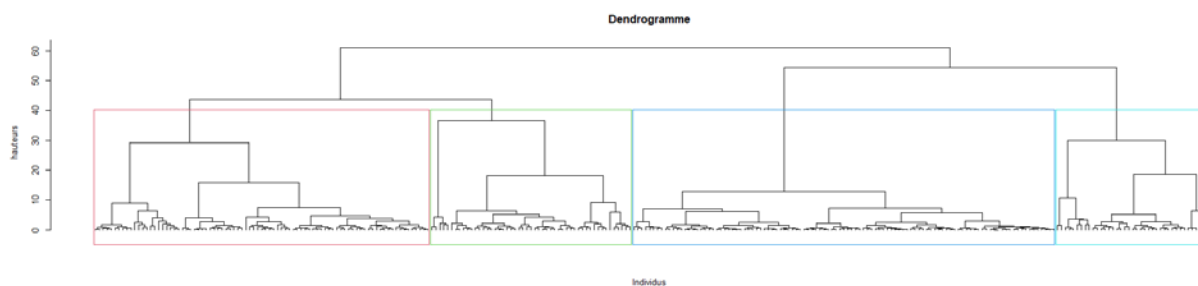


**Figure 6.3.2: Mixture U matrix explaining 72% of the total variance of the exposure of the FLESH III population**

Each column represents the composition of one mixture. A dark blue element of matrix U corresponds to a non-contributory element, whereas a dark red one indicates that the substance is highly contributing to the mixture. Mixture 1 (44% of the variance) is mainly composed of PCBs (PCB153, PCB146, PCB180, PCB138, PCB170, PCB187, PCB118). Mixture 2 (8.5% of the variance) reveals high correlations between PFOA and PFNA. Mixture 3 (8.7% of the variance) is mainly composed of the metals copper and cadmium associated to HCB. Mixture 4 (5.1% of the variance) is mainly composed of one substance, the manganese. Strictly speaking, mixture 5 (5.4% of the variance) consists of PFHXS alone, but it is interesting to note that the second most important substance of this mixture is PFOS. The precise composition is available in Annex 1 presenting the substances weights of substances composing mixtures.

### 6.3.1.3 Co-exposure profiles

Following the mixture identification, the clustering method applied to the matrix V leads to a division of the FLEHS III population into 4 clusters of individuals with similar exposure profiles (Figure 6.3.3).

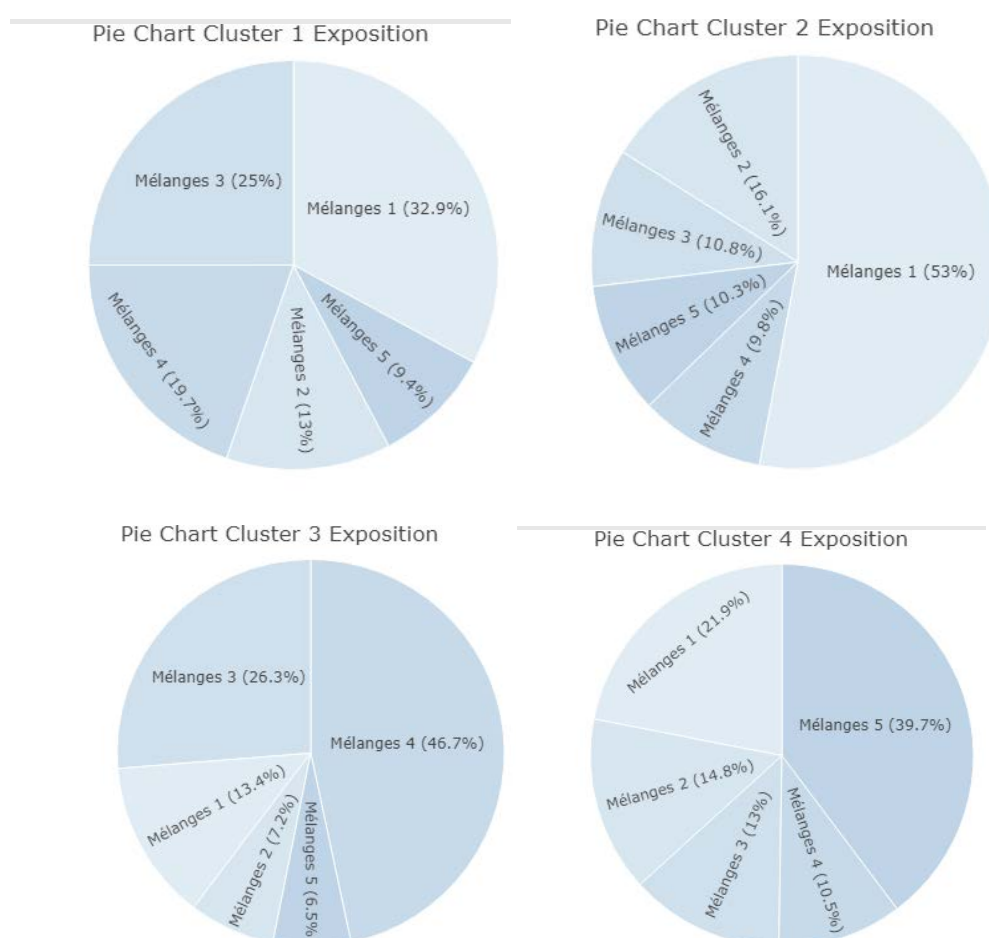


**Figure 6.3.3: Dendrogram of the clustering of the FLEHS III population regarding its exposure to the obtained 5 mixtures**

The contribution of mixtures per cluster, also called the mixture weights, is described in pie charts (Figure 6.3.4). The first cluster included 38% of the study population (107 women). The mixtures that contributed the most to this cluster are mixtures 1 “PCBs” (32.9%), mixture 3 “Cu-Cd-HCB” (25%), and mixture 4 “Mn” (19.7%). The second cluster included 18% of the study population (51 women) and was highly characterised by mixture 1 “PCBs” (53%), and to a lesser extent mixture 2

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 65

“PFOA, PFNA” (16.1%) and mixture 3 “Cu-Cd-HCB” (10.8%). Cluster 3 is the smallest with 14% of the study population (38 women). Mixtures 4 “Mn” (46.7%), mixture 3 “Cu-Cd-HCB” (26.3%) and mixture 1 “PCBs” (13.4%) are descriptive of cluster 3. The last 30% of the study population (85 women) is mainly exposed to mixtures 5 “PFHXS, PFOS” (53%), mixture 1 “PCBs” (21.9%) and mixture 2 “PFOA, PFNA” (14.8%). Women in this cluster tend to be more primiparous mother than the rest of the study population according to the Fisher test (62% of primiparous mothers in this cluster against 45% in the total population).



**Figure 6.3.4: Contribution of mixtures to the 4 selected clusters of the FLEHS III population**

#### 6.3.1.4 Integration of SNMU and classification results

In order to characterise the clusters' exposures, we can compare the exposure to each substance (i.e., the exponential residualised concentrations) between cluster populations and the total studied population. Annex 2 summarises the main mixtures FLEHS III clusters are exposed to. For each substance in a mixture, the substance weights, the means and the 95<sup>th</sup> percentiles of exposure are reported, as well as the result of the exposure comparison tests between the clustered population and the total population. For example, the 51 mothers of cluster 2 are significantly more exposed to the main substances of mixture 1, namely PCBS, than the all studied population. Whereas the 85 women of cluster 4 are statistically more exposed to the main substances of mixture 5, namely PFHXS and PFOS.

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebrecht, Madeline Carsique, Eva Govarts, et al.	Page: 66

The relative exposure level of clusters with respect to the whole population is summarised in Table 6.3.1. Signs + or – imply a significant higher or lower difference in the mean exposure using a t-test  $p$ -value < 0.05 compared to the total population.

**Table 6.3.1: Relative exposure level of clusters: + higher exposure in the cluster than in the total population; - lower exposure in the cluster than in the total population; ~no significant exposure difference between the cluster and the total population**

Cluster	Number of individuals	Mixture (main components)				
		N°1 (PCBs)	N°2 (PFOA, PFNA)	N°3 (Cu, Cd, HCB)	N°4 (Mn)	N°5 (PFHXS, PFOS)
Cluster 1	107	-		-	-	
Cluster 2	51	+	+	~		
Cluster 3	36	-		+	+	
Cluster 4	85	-	~			+

### 6.3.2 Discussion and conclusions on SNMU and SNMU – network analysis comparison

The SNMU approach by combining a reduction dimension method and a clustering method identifies 5 mixtures explaining 72% of the variance of the exposure of the 281 mothers grouped in 4 clusters. The output results could be used to prioritise the mixtures to be assessed by toxicologists and as input in epidemiological analysis.

The 5 mixtures identified from FLEHS III data by analysing matrix U obtained with the SNMU method (Figure 6.3.2) are compared to the ones obtained by network analysis method (Figure 6.3.1A). Mixture 1 (PCBs), mixture 2 (PFOA-PFNA), mixture 4 (Mn), mixture 5 (PFHXS-PFOS) from SNMU method are also identified in the unweighted network as we can see strong edges between these substances. No connection between HCB and any metals is observed in the network, whereas mixture 3 obtained with the SNMU approach is composed of Cu, Cd and HCB. Both methods identify very similar mixtures, confirming the stability of these methods to identify mixtures based on co-exposures. However, each method provides different and complementary outputs. Indeed, network analysis provides an intuitive graphical representation displaying degrees and directions of correlations. With this method, the choice of the number of mixtures ( $k$  value) or the number of clusters is automatised. However, this method does not provide a precise composition of mixture. Although outputs derived from SNMU approach are more complex providing heatmap and tables, this method can provide details by providing substance weights representing the contribution of each substance to the mixtures. This method is also well adapted for non-negative data (such as HBM data) and with many zero entries, therefore data would not have required to be transformed as much as it does for the networks. Indeed, the data could have been used raw without any additional transformations, without implementing missing data and the selection criterion of substances (more than 60% of measured concentrations above LOD per substance) might have been bypassed for some substances. Moreover, this approach combines a classification analysis by clustering individuals according to their exposure to the mixtures and integrating their characteristics. This last point avoids the need to display several graphical networks to compare the characteristics of individuals. Lastly, the matrices generated by SNMU could be reused to guide toxicological and epidemiological studies. For toxicological testing, the weights of the matrix U can be used to define the percentage of each substance in the mixture. For



epidemiological studies, the matrix U can be used as input in regression models to study the link between exposure to the obtained mixtures with biomarkers of effects. Another way to use the outputs is to indicate the cluster to which each individual belongs to in epidemiological models.

However, it should be noted that the dataset used for this analysis is low-dimensional (only 19 substances considered), whereas the methods are developed to work in high dimension. Indeed, biomonitoring data are often few in number of measured substances for a single individual, which limits the co-exposure analysis.

## 6.4 Toxicity weighting

### 6.4.1 Results of UBA-IOM data application

We applied the approach as described in section 5.3.6 and 5.4.4 and used the HBGV database to find corresponding values. As can be seen in Fig. 6.4.1 there are multiple communities and three of them contain multiple compounds with available HBGVs.

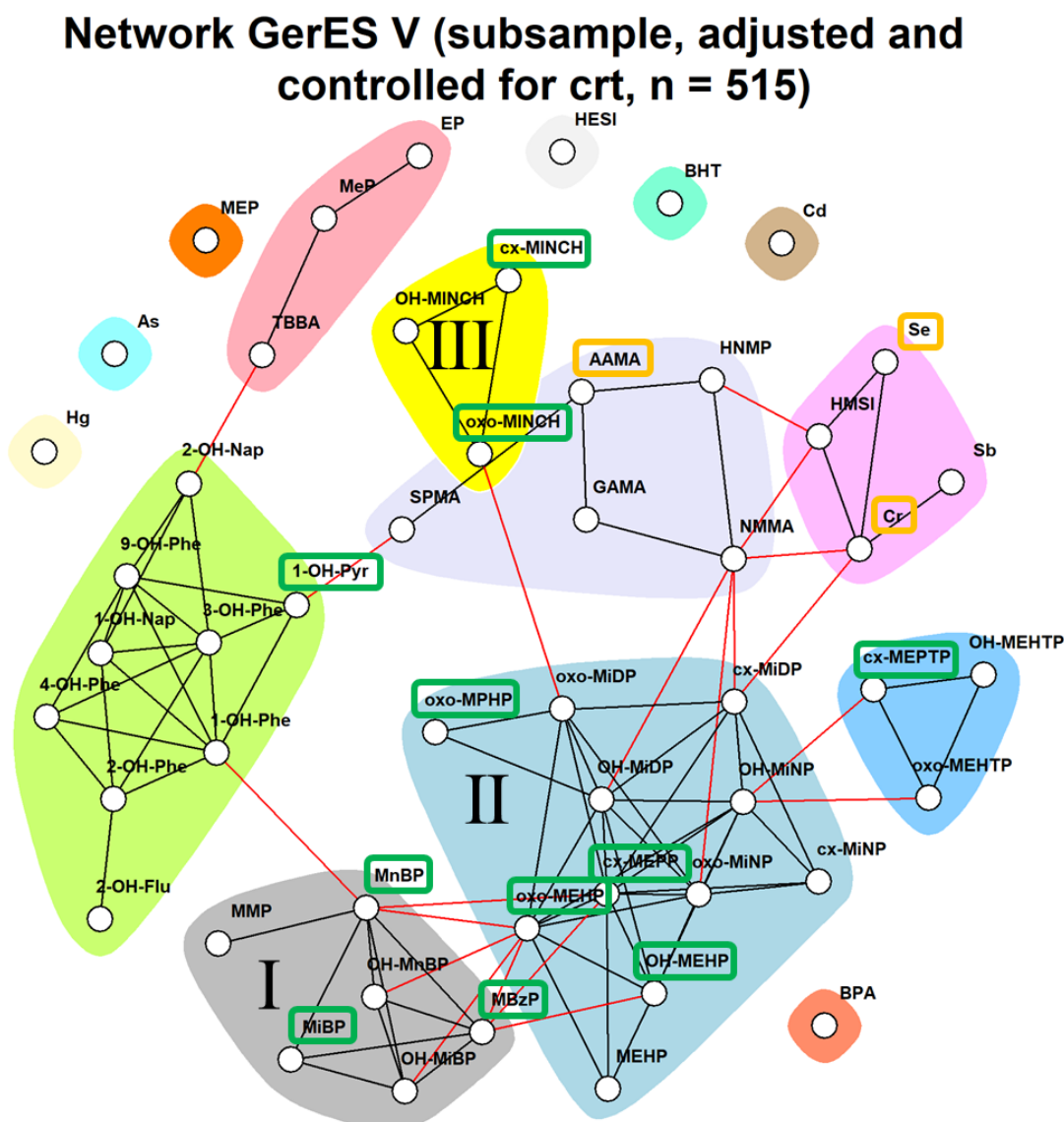
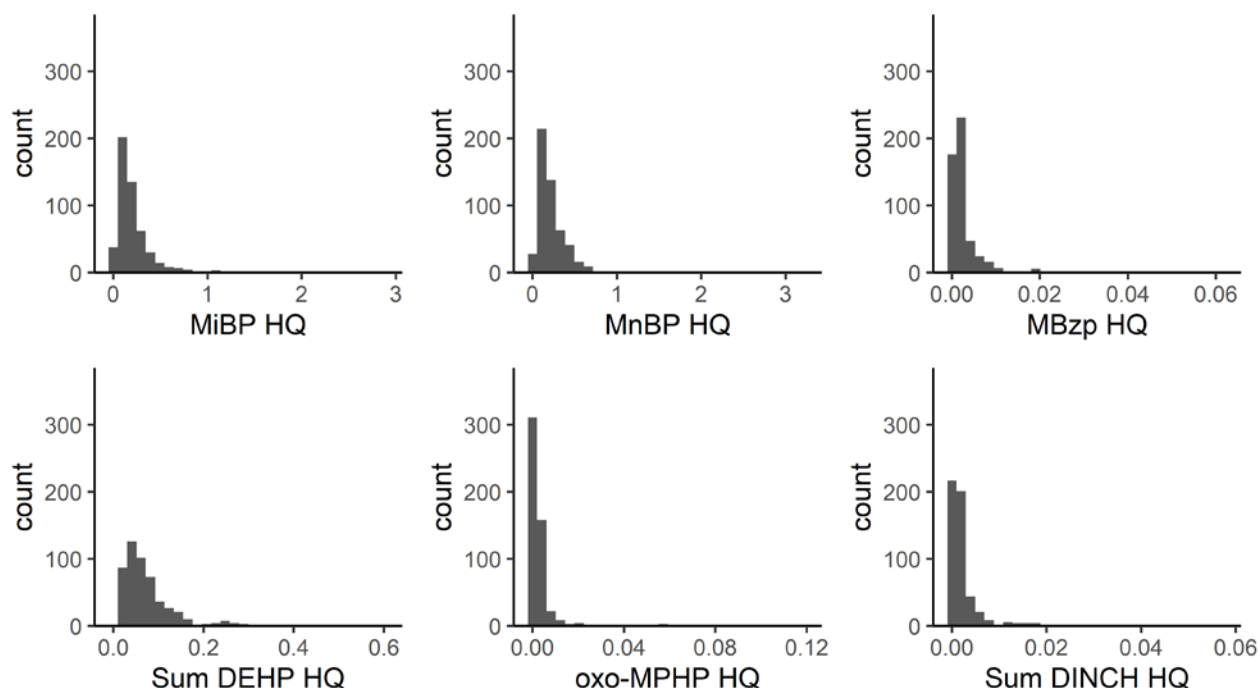
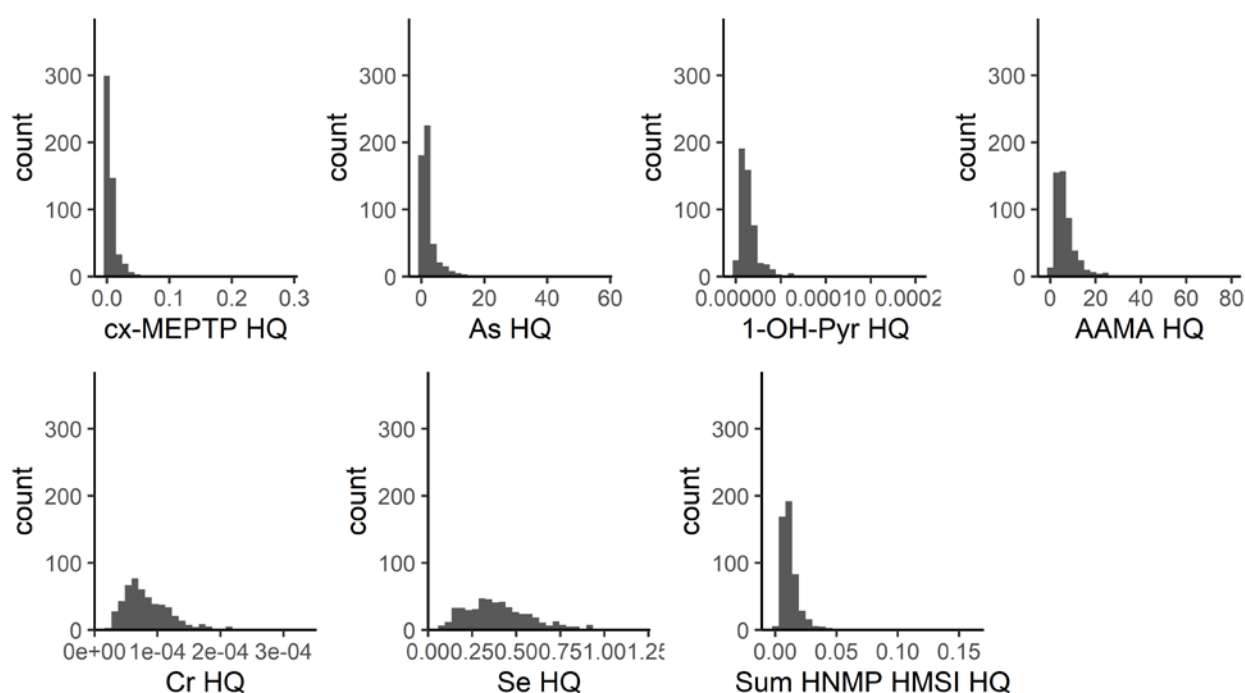


Fig 6.4.1: Network and the available guidance values (HBGVs in green; BEs in yellow). Three communities with two or more HBGVs are present.

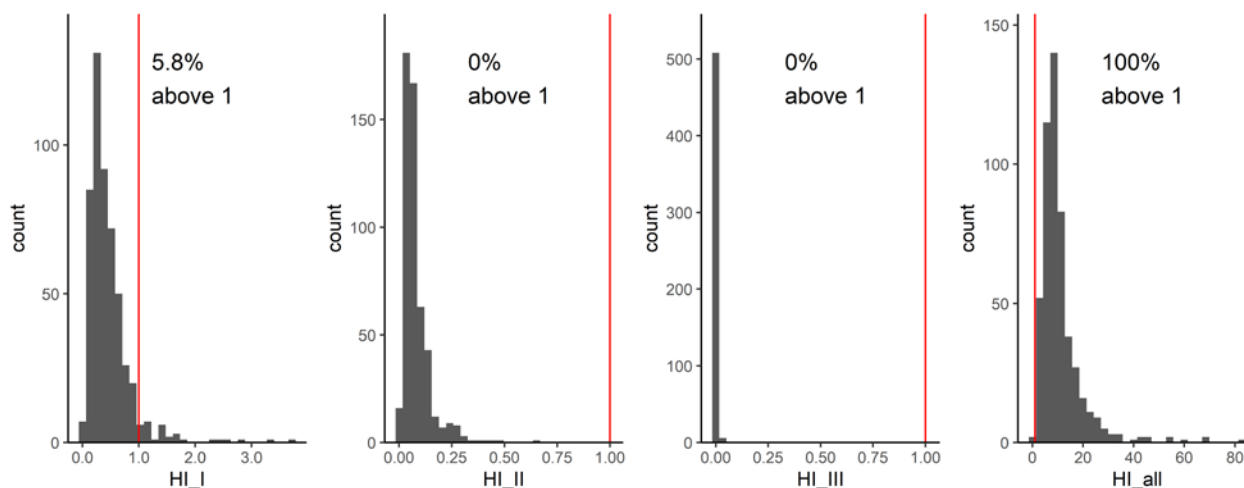
Figure 6.4.2 shows the distribution of individual hazard quotients (HQ) of the six substances involved in communities I to III. The differences between compounds are striking and in cases exceedance of the HBMGVs can be observed. This indicates that the compounds with high hazard quotients will make a large contribution to the hazard index.



**Fig 6.4.2: Distributions of hazard quotients for the six substances that are involved in the three communities. For DEHP and DINCH the available HBGVs for the sums were used (oxo-MEHP, cx-MEPP, OH-MEHP and cx-MINCH, oxo-MINCH, respectively).**



**Fig 6.4.3: Distributions of hazard quotients for the other substances that are single members of communities. The HQs for Arsenic and AAMA are based on BEs and show excessive exceedance.**



**Figure 6.4.4: Distributions for hazard indices of the three communities and combined for all substances for which HQs could be calculated. The red line indicates the value of 1. For the first community a small fraction exceeds 1.**

Hazard indices (HI) were calculated for the three communities I, II and III and are shown in Figure 6.4.4. Community I contains the phthalate metabolites MiBP and MnBP with high hazard quotients, which leads to hazard indexes that exceed the value of 1 in 5,8% of the individuals. The combined hazard indices from all substances with available HBGVs show high values and are dominated by the high hazard quotients of arsenic and acrylamide.

#### 6.4.2 Discussion and conclusions on toxicity weighting

The use of toxicity weighting the communities is a novel approach and presents an opportunity to gauge the hazard potential of a newly discovered mixture. Setting the exposure in relation to guidance values allows an interpretation whether the communities should be further investigated. There are a number of knowledge gaps which need to be addressed to allow this approach to be applied more completely, and which would enhance the utility of toxicity weighting for risk screening. Missing HBGVs for members of the community present a problem and can cause the hazard to be underestimated. Eleven of the 51 substances had a HBGV. Of the eight communities in figure 6.4.1, six had at least one HBGV. Two communities had only one HBGV. Three BE values were available to improve the coverage. For the large community studied in this network about 50% of the biomarkers had an associated HBGV value. While the complete hazard indices of the communities cannot be calculated with missing values, using an incomplete set and still finding exceedance is a reason for concern and requires further investigation. HBGVs may also change as information on chemicals changes – which means that the HI's may differ depending on the HBGVs used. Incorporation of metabolites is also something to be considered, as HBGVs may not be available for all metabolites of a parent compound or may only be available for the sum of compounds.

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebet, Madeline Carsique, Eva Govarts, et al.	Page: 70

## 7 Discussion and conclusions

### 7.1 Methodological aspects

#### 7.1.1 Network analyses

For the further exploration of co-occurrence of substances, four existing databases on HBM data were used. These databases varied considerably due to differences in study design, population sampled, biological samples collected and chemical analytical procedures. To the degree possible, harmonised approaches for data preparation, imputation of missing data, descriptive statistics and statistical procedures were followed, where possible based on HBM4EU procedures developed in other WP's. Also, the network analyses were optimised per dataset, potentially leading to differences in the optimised parameter setting. Nonetheless, comparisons across datasets should be done with great care, given the differences in study designs and laboratories.

Prior to the network analysis, missing data on individual biomarkers were imputed; a (arbitrary) cut-off of 60 % above LOD/LOQ was uniformly selected for inclusion of the data in the network analysis. At this point, no sensitivity analysis was performed with different cut-off points, or alternative imputation strategies to assess the influence on network results.

Biomarker levels were normalised against creatinine and blood lipids, where appropriate. This was considered necessary to avoid 'spurious' correlations based on e.g., the level of dilution or concentration of urine samples. In datasets where only urine samples were used, other means to correct for creatinine were an option, e.g., by including creatinine as a covariate in the analysis. This was explored in the German dataset. Since other databases had biomarkers measured in urine and in blood, this procedure was not the default and individual biomarker levels were normalised against individual creatinine or blood lipid values. Other, more optimal normalisation strategies, e.g., differentiation by age group, or taking into account physical activity in firefighters and using specific gravity approaches were not explored here. The exercise to compare different creatinine normalisation approaches on the German dataset indicated that normalisation is indeed necessary, since the network on 'raw' unnormalised and uncontrolled data showed three heavily interrelated communities, thereby masking communities identified with normalisation and/or control for creatinine level. This is probably the results of differences in the dilution/concentration level of the urine individual samples.

In the four studies, different selections of study (sub)populations with different numbers of measured substances could be made. Generally, the more substances measured, the smaller the study population with a complete data set. For this exercise, we selected the largest possible number of substance available, thus reducing the number of participants to include in the analysis. This reduction of the data may have limited the stability of the obtained networks. It did hamper comparison between strata of covariates in the differential network analysis, particularly for smoking, where only a small fraction of the population were smokers. While the obtained results certainly show the feasibility of analysis on strata of covariates of interest, i.e., potential determinants of HBM levels and patterns of co-occurrence of substances, interpretation of the results in terms of the observed patterns should be done with great care. Our research question 3 "Are such patterns indicative for specific sources or pathways of exposure?" cannot be directly answered yet, based on covariate analysis on the available datasets. It is therefore recommended that future HBM studies should aim to collect data on the full range of chemicals of interest in sufficiently large study populations measured in the same individuals, to assess the actual mixture exposures in the population and co-occurrence in the body. This will allow sufficient observations to also assess the influence of covariates with sufficient power.

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 71

## 7.2 Results discussion

### 7.2.1 Network analysis

Overall, the observed correlations were moderate, with more positive than negative associations. Negative correlations were generally below 0.3. With the exception of parent-metabolite relations, correlations were generally below 0.8. In this study, the focus was on the dependencies in HBM levels, not so much on the absolute levels. In prioritising mixtures of concern, of course also the absolute levels should be considered.

The network analysis in all four studies identified, as expected, several communities of chemical families, e.g., PCBs, PFAS. Also, links between parent and metabolites were observed, e.g., for acrylamide. However, also connections across chemical families were observed, possibly reflecting common exposure or ADME patterns.

In the German and Spanish data, metals were quite often not part of communities. Also, in the German and Spanish unweighted networks, MEP was not part of a community. In the weighted networks, MEP was part of a community in both countries, but due to differences in measured substances, a direct comparison cannot be made. In the in the Belgium 3xG data, patterns were different. This may result from the differences in matrices in which the metals were measured. In the Belgium FLEHS data, Cd, Cu and Mn formed a community, while other metals were not part of a community (Ottenbros et al, 2021). In the German study, only urine samples were used, while in the other Spanish and Belgium studies, also blood samples were included. The observed differences in inclusions of metals in communities may therefore stem in differences in exposure pattern (recent versus more long-term) and/or ADME processes. Also, in the Belgian study, the relation between mother and newborn may give rise to different community patterns, i.e., in metals.

As mentioned earlier, the results of the differential network analysis into the possible effects of covariates/determinants should be interpreted with care, particularly where the number of observations per stratum is small compared to the number of chemicals measured. While it is too early for an in-depth interpretation of underlying mechanisms for the observed differences in differential network analysis, these first findings do illustrate the potential use to explore lifestyle and exposure patterns on HBM mixture patterns. For example, co-occurrence of parabens (MeP and EP) with Lysmeral (TBBA) in a community in the German network results may point at the role of cosmetics.

The weighted and unweighted network analyses yielded generally similar results. To a large degree, the links in the weighted analyses reflect the observed communities in the unweighted networks. While weighted network analysis is more computationally intensive and less fit for high dimensional data in comparison to the unweighted networks, and advantage is the indication of the relative strength of the links and the direction of the association. The latter seems less important in these datasets, since negative correlations were relatively scarce and overall rather weak. Where the links in the unweighted networks all seem as important, the weighed networks focus attention on the stronger associations. Combined with their HBM levels, this may be more indicative for communities of concern than weak or undifferentiated links.

### 7.2.2 SNMU

As SNMU approach and network analysis provide similar results, they have their own strength and limitations, and they can complement each other. A combination of the two approaches would provide a great representation and understanding of the dataset thanks to network analysis, and deeper analysis and interpretation thanks to SNMU approach. This has already been proposed in the paper by Chazelas et al. (2021). In this article food additive mixtures have been derived by the Non-negative matrix factorisation which is a previous version of the SNMU, and clusters have been



D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 72

identified according to additive intake. Then the network of partial correlations between food additive intakes is displayed to get a more visual understanding of the dataset.

### 7.2.3 Toxicity weighting

To address research question 5 “Can we develop aggregate/hybrid indicators that encapsulate toxicity of the mixture in a meaningful way, e.g., by using HI or CAGs?” and 6 “Can we define mixture levels of excess risk, based on toxicity-based aggregation of HBM mixture data?”, we explored toxicity weighting based on available HBGV and BE’s, using the German network analysis results and applying hazard quotients and hazard index approach. This was done for all available HBGV, ignoring differences in working mechanisms and modes of action, assuming dose addition across the board. Of the 51 substances in the network only 11 HBGVs and 3 BEs were available. Six of the eight communities had more than one HBGV. Two communities had only one HBGV.

HQ for individual substances in communities I to III all were below one. The HI for community I showed HI > 1 with individual values > 3. The other two communities had HI below 1.

This exercise shows that in principle toxicity weighting can be applied at the level of identified communities, the limiting factor being the lack of HBGV’s. This factor would be magnified when suspect screening techniques are applied to HBM. Thus, a more generic approach, with wider coverage of toxicity weights is needed. The approach suggested by Zhao and coworkers (Zhao et al. 2021) may prove fruitful for better coverage. However, the data in that database are mainly from animal experiments from a wide range of different experimental designs, tests and assays, which may make comparison to HBM values less appropriate.

## 7.3 Closing remarks

The specific research questions we initially aimed at addressing in this study were:

1. How can we rank order individuals on the basis of body burdens to mixtures?
2. What patterns can we observe amongst body burdens of different substances within individuals? Or, are people with high exposure levels for some substances more likely to have high exposure levels on other substances as well?
3. Are such patterns indicative for specific sources or pathways of exposure?
4. Can we identify hotspots or risk groups with high body burdens of mixtures?
5. Can we develop aggregate/hybrid indicators that encapsulate toxicity of the mixture in a meaningful way, e.g., by using HI or CAGs?
6. Can we define mixture levels of excess risk, based on toxicity-based aggregation of HBM mixture data?

**Question 1:** “How can we rank order individuals on the basis of body burdens to mixtures?” was addressed in two ways. Through the SNMU analysis where individuals were clustered based on mixture weights into four clusters, across the identified mixtures. In the toxicity weighting exercise, we ranked individuals on HQ and HI, overall and per mixture community. This is a composite of the individual body burden weighted against available HBGV’s.

**Question 2:** “What patterns can we observe amongst body burdens of different substances within individuals? Or, are people with high exposure levels for some substances more likely to have high exposure levels on other substances as well?” is in essence answered by both the network analysis and SNMU, showing communities/clusters of associated co-occurring HBM biomarker levels.



D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 73

**Question 3:** “Are such patterns indicative for specific sources or pathways of exposure?”. The patterns identified by communities in the network analysis and SNMU showed combinations of chemicals family together with other substances from other sources, regulated through different regulatory silos. The differential network analysis to assess effects of determinants/covariates at the source level was only cursorily explored, by looking at smoking status, and fish diet in this phase. In both cases, relatively few observations in the high exposure stratum were available, potentially limiting the ability assess stable network differences. Nonetheless, the observed community of parabens with Lismeral in the German network analysis does suggest cosmetics as a source, showing the possible use of further exploration and interpretation of the observed communities in network analysis. The SNMU approach includes an analysis of determinants characterising clusters of individuals and thus specific sources or pathways can be included.

**Question 4:** “Can we identify hotspots or risk groups with high body burdens of mixtures?” With the SNMU approach a first inroad was made to address this, by identifying clusters of individuals with more or less homogeneous mixture exposure patterns. Also, the individual HI distributions allow exploration when combined with covariate/determinant information, e.g., georeferenced data, proximity to nearby sources, occupational or lifestyle information. This has not yet been explored in this report.

**Question 5:** “Can we develop aggregate/hybrid indicators that encapsulate toxicity of the mixture in a meaningful way, e.g. by using HI or CAGs?” This question was specifically addressed in the Toxicity Weighting sections of this report, with a focus on the HI approach. Using data from the German network analysis, enriched with information on HBGVs and BE’s hazard quotients and hazard indices were calculated for individuals for three identified mixture communities.

**Question 6:** “Can we define mixture levels of excess risk, based on toxicity-based aggregation of HBM mixture data?” With the hazard index approach, in principle the distribution of HI’s at the level mixtures communities can be explored as exemplified in the Toxicity Weighting sections in this report. In this first exploration, the appropriateness to pool substances with different modes of action was ignored, and dose addition of all substances with HBGV or BE was considered suitable as a proof-of-principle.

## 7.4 Conclusions

- The co-occurrences of mixtures of substances using HBM data is feasible through network analysis and SNMU, each with its own strength and limitation. Combined application to explore and quantify co-occurrence is recommended.
- The SNMU analysis of the FLEHS data indicate that a substantial part of the variation in the HBM data (> 70 %) can be captured with a limited number of clusters. While this needs to be replicated in other data, there is no reason to believe that this will be very different in other HBM datasets.
- Existing databases of early HBM studies are useful for the first exploration of co-occurrence of substances in the human body.
- Existing studies typically have limited number of individuals in which the full range of chemical substances have been measured. This limits the ability to identify patterns of co-occurrence and even more so to study the role of determinant/covariates, with fewer observations per stratum. For future studies it is recommended to expand the number of observations with full range of chemicals substances, to improve the ability to study mixtures, e.g., by suspect screening and/or non-targeted analysis.

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebet, Madeline Carsique, Eva Govarts, et al.	Page: 74

- The stability and consistency of identified networks and mixture communities deserves further study, particularly for high dimensional data when strata of covariates are being studied.
- Toxicity weighting of mixture communities/clusters is feasible, but severely limited by the shortage of available HBGV's or other indicators of toxic potency of the substances involved. More generic inroads need to be explored, the more so when suspect screening and untargeted screening is wider applied in HBM studies.
- So far, four existing datasets were subjected to network analysis techniques, in one dataset network analysis and SNMU were compared, and toxicity weighting was explored in one dataset. The datasets varied widely in study population, study design, matrixes under study and chemicals analysed. It is recommended to expand current analysis to a wider set of existing data.

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 75

## 8 References

- Agier L. et al. 2016. A Systematic Comparison of Linear Regression-Based Statistical Methods to Assess Exposome-Health Associations. *Environ. Health Perspect.* doi:10.1289/EHP172
- Apel P, Rousselle C, Lange R, Sissoko F, Kolossa-Gehring M, Ougier E. Human biomonitoring initiative (HBM4EU) - Strategy to derive human biomonitoring guidance values (HBM-GVs) for health risk assessment. *Int J Hyg Environ Health.* 2020 Sep;230:113622. doi: 10.1016/j.ijheh.2020.113622. Epub 2020 Oct 9. PMID: 33045523.
- ANSES 2013. Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail. Expert appraisal on recommending occupational exposure limits for chemical agents Reference Document for the derivation and the measurement of exposure limit values for chemical agents in the workplace (OELs). <https://www.anses.fr/en/system/files/VLEP2009sa0339RaEN.pdf>
- ANSES 2015. Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail. Request no. 2007-SA-0431 – OEL permanent mission BLV cobalt and its compounds <https://www.anses.fr/fr/system/files/VLEP2007SA0431RaEN.pdf>
- Béchaux, C., Zetlaoui, M., J. Tressou, Leblanc, J.Ch., Héraud, F., Crépet, A. (2013) Identification of pesticide mixtures and connection between combined exposure and diet, *Food Chemical and Toxicology*, 59, 191–198.
- Barrera-Gómez, J. et al. 2017. A systematic comparison of statistical methods to detect interactions in exposome-health associations. *Environ. Heal. A Glob. Access Sci. Source* 16.
- Chazelas E, Druésne-Pecollo N, Esseddik Y, de Edelenyi FS, Agaesse C, De Sa A, Lutchia R, Rebouillat P, Srour B, Debras C, Wendeu-Foyet G, Huybrechts I, Pierre F, Coumoul X, Julia C, Kesse-Guyot E, Allès B, Galan P, Hercberg S, Deschasaux-Tanguy M, Touvier M. 2021. Exposure to food additive mixtures in 106,000 French adults from the NutriNet-Santé cohort. *Sci Rep.* 11(1):19680. doi: 10.1038/s41598-021-98496-6.
- Crépet, A., Vasseur, P., Jean, J., Badot, P.M., Nessler, F., Vernoux, J.V., Feidt, C., Mhaouty-Kodja, S. 2022. Integrating selection and risk assessment of chemical mixtures: a novel approach applied to a breastmilk survey. *Environmental Health Perspective.* 130(3):35001.
- Csárdi G, Nepusz T. The igraph software package for complex network research. *Inter J Comp Syst.* (2006) 1695:1–9. Available online at: <https://igraph.org/> (accessed July 27, 2020).
- Drakvik E, Altenburger R, Aoki Y, Backhaus T, Bahadori T, Barouki R, Brack W, Cronin MTD, Demeneix B, Hougaard Bennekou S, van Klaveren J, Kneuer C, Kolossa-Gehring M, Lebret E, Posthuma L, Reiber L, Rider C, Rüegg J, Testa G, van der Burg B, van der Voet H, Warhurst AM, van de Water B, Yamazaki K, Öberg M, Bergman Å. 2020. Statement on advancing the assessment of chemical mixtures and their risks for human health and the environment. *Environ Int.* 134:105267.
- DFG 2020. List of MAK and BAT Values 2020. Permanent Senate Commission for the investigation of Halht Hazards of Chemical Compounds in the Work Area. [https://series.publisso.de/sites/default/files/documents/series/mak/lmbv/Vol2020/Iss2/Doc002/mbwl\\_2020\\_eng.pdf](https://series.publisso.de/sites/default/files/documents/series/mak/lmbv/Vol2020/Iss2/Doc002/mbwl_2020_eng.pdf)
- European Commission. 2012. Communication from the Commission to the Council: The combination effects of chemicals – Chemical mixtures, COM, Brussels, 31.5.2012. Available at: [http://ec.europa.eu/environment/chemicals/effects/effects\\_en.htm](http://ec.europa.eu/environment/chemicals/effects/effects_en.htm). (Accessed: 23rd February 2022)

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 76

European Food Safety Authority. 2016. Public consultation on the terms of reference of the Scientific Committee Working Group on "Harmonisation of risk assessment methodologies for human health and ecological risk assessment of combined exposure to multiple chemicals". Technical report. EFSA Supporting publication.

<https://www.efsa.europa.eu/sites/default/files/consultation/161024a.pdf>.

Friedman J, Hastie T, Tibshirani R. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*. (2008) 9:432–41. doi: 10.1093/biostatistics/kxm045

Greet Schoeters, Rosa Lange, Marike Kolossa, et al. (2020). HBM4EU. Deliverable Report AD5.4 - Reporting for first and second set of substances.

German Human Biomonitoring Commission. Reference and HBM Values.

<https://www.umweltbundesamt.de/en/topics/health/commissions-working-groups/human-biomonitoring-commission/reference-hbm-values>

Gillis, Nicolas, et François Glineur. 2010. « Using underapproximations for sparse nonnegative matrix factorization ». *Pattern Recognition* 43: 1676-87.

<https://doi.org/10.1016/j.patcog.2009.11.013>.

Gillis, Nicolas, et Robert J. Plemmons. 2013. « Sparse Nonnegative Matrix Underapproximation and Its Application to Hyperspectral Image Analysis ». *Linear Algebra and Its Applications* 438 (10): 3991-4007. <https://doi.org/10.1016/j.laa.2012.04.033>.

Golino H, Christensen AP. EGAnet: Exploratory Graph Analysis: A Framework for Estimating the Number of Dimensions in Multivariate Data Using Network Psychometrics. (2020). Available online at: <https://cran.r-project.org/package=EGAnet> (accessed July 27, 2020).

Govarts E, Portengen L, Lambrechts N, Bruckers L, Den Hond E, Covaci A, Nelen V, Nawrot TS, Loots I, Sioen I, Baeyens W, Morrens B, Schoeters G, Vermeulen R. 2020. Early-life exposure to multiple persistent organic pollutants and metals and birth weight: Pooled analysis in four Flemish birth cohorts. *Environ Int*. 145:106149

Hays SM, Aylward LL. Interpreting human biomonitoring data in a public health risk context using Biomonitoring Equivalents. *Int J Hyg Environ Health*. 2012 Feb;215(2):145-8. doi: 10.1016/j.ijheh.2011.09.011. Epub 2011 Dec 20. PMID: 22192580.

HSE UK 2020. EH40/2005 Workplace exposure limits: Containing the list of workplace exposure limits for use with the Control of Substances Hazardous to Health Regulations 2002 (as amended) <https://www.hse.gov.uk/pubns/priced/eh40.pdf>

Kienzler A, Bopp SK, van der Linden S, Berggren E, and Worth A. 2016. Regulatory assessment of chemical mixtures: Requirements, current approaches and future perspectives. *Regulatory Toxicology and Pharmacology* 80, 321–334.

Kirman CR, Aylward LL, Hays SM, Krishnan K, Nong A. Biomonitoring equivalents for DDT/DDE. *Regul Toxicol Pharmacol*. 2011 Jul;60(2):172-80. doi: 10.1016/j.yrtph.2011.03.012. Epub 2011 Apr 3. PMID: 21466830.

Liu H, Roeder K, Wasserman L. Stability Approach to Regularization Selection (StARS) for High Dimensional Graphical Models. (2010). Available online at: <http://arxiv.org/abs/1006.3316> (accessed July 27, 2020).

Lubin JH, Colt JS, Camann D, Davis S, Cerhan JR, Severson RK, et al. Epidemiologic evaluation of measurement data in the presence of detection limits. *Environ Health Perspect*. (2004) 112:1691–6. doi: 10.1289/ehp.7199

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 77

Meek, M.E.(Bette), Boobis, A.R., Crofton, K.M., Heinemeyer, G., Raaij, M.V., Vickers, C., Risk assessment of combined exposure to multiple chemicals: A WHO/IPCS framework, Regulatory Toxicology and Pharmacology (2011), doi: 10.1016/j.yrtph.2011.03.010

New Zealand Government 2020. Worksafe. Workplace exposure standards and biological indices. Ed. 12-1. <https://www.worksafe.govt.nz/topic-and-industry/monitoring/exposure-standards-and-biological-exposure-indices>

Orman GK, Labatut V. “A comparison of community detection algorithms on artificial networks,” in International Conference on Discovery Science. Portugal: Springer (2009). p. 242–56. doi: 10.1007/978-3-642-04747-3\_20

Ottenbros I, Govarts E, Lebret E, Vermeulen R, Schoeters G, Vlaanderen J. 2021. Network Analysis to Identify Communities Among Multiple Exposure Biomarkers Measured at Birth in Three Flemish General Population Samples. Front Public Health. 9:590038.

Pérez-Gómez B, Pastor-Barriuso R, Cervantes-Amat M, Esteban M, Ruiz-Moraga M, Aragonés N, Pollán M, Navarro C, Calvo E, Román J, López-Abente G, Castaño A; BIOAMBIENT.ES. 2013. BIOAMBIENT.ES study protocol: rationale and design of a cross-sectional human biomonitoring survey in Spain. Environ Sci Pollut Res Int. 1193-202. doi: 10.1007/s11356-012-1320-3.

Pons P, Latapy M. Computing Communities in Large Networks Using Random Walks. (2005). p. 284–93.

Rotter S, Beronius A, Boobis AR, Hanberg A, van Klaveren J, Luijten M, Machera K, Nikolopoulou D, van der Voet H, Zilliacus J, Solecki R. 2018. Overview on legislation and scientific approaches for risk assessment of combined exposure to multiple chemicals: the potential EuroMix contribution. Crit Rev Toxicol. 2018 Oct;48(9):796-814. doi: 10.1080/10408444.2018.1541964.

Schwedler G, Conrad A, Rucic E, Koch HM, Leng G, Schulz C, Schmied-Tobies MIH, Kolossa-Gehring M. 2020. Hexamoll® DINCH and DPHP metabolites in urine of children and adolescents in Germany. Human biomonitoring results of the German Environmental Survey GerES V, 2014-2017. Int J Hyg Environ Health. 229:113397.

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 MA, 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(Supplement C):310-328.

US DoHHS 2004. US Department of Health and Human Services. Agency for Toxic Substances and Disease Registry. Division of Toxicology. Guidance Manual for the Assessment of Joint Toxic Action of Chemical Mixtures.

Zhao, F, Li Li, Yue Chen, Yichao Huang, Tharushi Prabha Keerthisinghe, Agnes Chow, Ting Dong, Shenglan Jia, Shiwei Xing, Benedikt Warth, Tao Huan, and Mingliang Fang. Risk-Based Chemical Ranking and Generating a Prioritized Human Exposome Database. EHP, Vol. 129, No. 4. <https://doi.org/10.1289/EHP7722>

Zhao T, Liu H, Roeder K. The huge package for high-dimensional undirected graph estimation in R. J Mach Learn Res. (2012) 13:1059–62.

## 9 Annexes

### 9.1 Annex 1: Matrix U from SNMU method on FLEHS III

Substances	V1	V2	V3	V4	V5
As	0.014	0.016	0.071	0.000	0.053
Cd	0.007	0.000	0.308	0.000	0.000
Cu	0.012	0.050	0.391	0.001	0.039
Mn	0.011	0.053	0.017	0.695	0.003
Pb	0.013	0.000	0.000	0.112	0.000
Tl	0.014	0.000	0.000	0.027	0.000
DDE	0.017	0.000	0.000	0.026	0.000
HCB	0.027	0.095	0.191	0.090	0.000
PCB118	0.083	0.007	0.005	0.000	0.058
PCB138	0.109	0.003	0.000	0.009	0.031
PCB146	0.119	0.002	0.000	0.000	0.014
PCB153	0.140	0.000	0.000	0.000	0.000
PCB170	0.107	0.000	0.008	0.012	0.000
PCB170	0.110	0.000	0.005	0.017	0.000
PCB180	0.100	0.007	0.003	0.003	0.000
PCB187	0.037	0.000	0.000	0.000	0.656
PFHXS	0.030	0.377	0.000	0.000	0.000
PFNA	0.022	0.390	0.000	0.000	0.000
PFOA	0.028	0.001	0.000	0.006	0.146
PFOS	0.014	0.016	0.071	0.000	0.053



D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebrecht, Madeline Carsique, Eva Govarts, et al.	Page: 79

## 9.2 Annex 2: SNMU summary tables

Mixtures	Substances	Substance weight	Cluster 1 - 107 mothers			
			Mixture weight	Exposure		
				Mean	P95	Compared to all
Mixture 1	PCB153	14.00%	32.9%	0.814	1.64	*** (<)
	PCB146	11.90%		0.805	1.69	*** (<)
	PCB180	11.00%		0.842	1.63	*** (<)
	PCB138	10.90%		0.871	2.08	*** (<)
	PCB170	10.70%		0.861	1.65	*** (<)
	PCB187	10.00%		0.88	2.14	*** (<)
	PCB118	8.33%		1.05	2.49	** (<)
	PFHXS	3.73%		0.629	1.24	*** (<)
	PFNA	3.01%		0.889	2.29	*** (<)
	PFOS	2.76%		1.23	2.85	
	HCB	2.73%		1.39	3.05	
	PFOA	2.22%		0.966	2.27	*** (<)
	DDE	1.66%		1.22	5.01	* (<)
	TI	1.41%		1.85	5.67	
	As	1.39%		1.48	5.07	
	Pb	1.29%		1.36	3.79	
	Cu	1.22%		1.14	3.87	** (<)
	Mn	1.05%		0.91	2.15	*** (<)
	Cd	0.73%		1.41	4.31	
Mixture 3	Cu	39.1 %	25%	1.14	3.87	** (<)
	Cd	30.8 %		1.41	4.31	
	HCB	19.1 %		1.39	3.05	
	As	7.12 %		1.48	5.07	
	Mn	1.73 %		0.91	2.15	*** (<)
	PCB170	0.834 %		0.861	1.65	*** (<)
	PCB180	0.532 %		0.842	1.63	*** (<)
	PCB118	0.494 %		1.05	2.49	** (<)
	PCB187	0.27 %		0.88	2.14	*** (<)
Mixture 4	Mn	69.5 %	19.7%	0.91	2.15	*** (<)
	Pb	11.2 %		1.36	3.79	
	HCB	8.99%		1.39	3.05	
	TI	2.74 %		1.85	5.67	
	DDE	2.65 %		1.22	5.01	* (<)
	PCB180	1.74 %		0.842	1.63	*** (<)
	PCB170	1.16 %		0.861	1.65	*** (<)
	PCB138	0.939 %		0.871	2.08	*** (<)
	PFOS	0.625 %		1.23	2.85	
	PCB187	0.271 %		0.88	2.14	*** (<)
	Cu	0.115 %		1.14	3.87	** (<)

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebret, Madeline Carsique, Eva Govarts, et al.	Page: 80

Mixtures	Substances	Substance weight	Cluster 2 - 51 mothers			
			Mixture weight	Exposure		
				Mean	P95	Compared to all
Mixture 1	PCB153	14.00%	53%	3.2	5.68	*** (>)
	PCB146	11.90%		3.51	6.75	*** (>)
	PCB180	11.00%		3.13	4.98	*** (>)
	PCB138	10.90%		3.13	6.78	*** (>)
	PCB170	10.70%		3.15	6.16	*** (>)
	PCB187	10.00%		3.55	7.41	*** (>)
	PCB118	8.33%		2.79	7.03	*** (>)
	PFHXS	3.73%		1.59	3.87	
	PFNA	3.01%		3.12	10.2	** (>)
	PFOS	2.76%		2.02	5.05	
	HCB	2.73%		1.78	4.58	
	PFOA	2.22%		2.82	13.2	
	DDE	1.66%		3.23	10.2	
	TI	1.41%		1.93	9.31	
	As	1.39%		2.5	11.7	
	Pb	1.29%		1.71	5.02	
	Cu	1.22%		1.3	2.92	
	Mn	1.05%		1.35	3.17	
	Cd	0.73%		1.45	4.11	
Mixture 2	PFOA	39 %	16.1%	2.82	13.2	
	PFNA	37.7 %		3.12	10.2	** (>)
	HCB	9.46 %		1.78	4.58	
	Mn	5.26 %		1.35	3.17	
	Cu	5.02 %		1.3	2.92	
	As	1.6 %		2.5	11.7	
	PCB187	0.677 %		3.55	7.41	*** (>)
	PCB118	0.66 %		2.79	7.03	*** (>)
	PCB138	0.289 %		3.13	6.78	*** (>)
	PCB146	0.166 %		3.51	6.75	*** (>)
	PFOS	0.118 %		2.02	5.05	
Mixture 3	Cu	39.1 %	10.8%	1.3	2.92	
	Cd	30.8 %		1.45	4.11	
	HCB	19.1 %		1.78	4.58	
	As	7.12 %		2.5	11.7	
	Mn	1.73 %		1.35	3.17	
	PCB170	0.834 %		3.15	6.16	*** (>)
	PCB180	0.532 %		3.13	4.98	*** (>)
	PCB118	0.494 %		2.79	7.03	*** (>)
	PCB187	0.27 %		3.55	7.41	*** (>)

Mixtures	Substances	Substance weight	Cluster 3 - 38 mothers			
			Mixture weight	Exposure		
				Mean	P95	Compared to all
Mixture 4	Mn	69.5 %	46.7%	4.67	4.67	*** (>)
	Pb	11.2 %		3.03	3.03	
	HCB	8.99%		1.91	1.91	
	Tl	2.74 %		1.79	1.79	
	DDE	2.65 %		1.93	1.93	
	PCB180	1.74 %		0.887	0.887	*** (<)
	PCB170	1.16 %		0.889	0.889	*** (<)
	PCB138	0.939 %		0.855	0.855	*** (<)
	PFOS	0.625 %		1.01	1.01	** (<)
	PCB187	0.271 %		0.922	0.922	*** (<)
	Cu	0.115 %		2.82	2.82	* (>)
Mixture 3	Cu	39.1 %	26.3%	2.82	7.47	* (>)
	Cd	30.8 %		4.56	17.5	* (>)
	HCB	19.1 %		1.91	3.59	
	As	7.12 %		1.31	2.87	
	Mn	1.73 %		4.67	10.7	*** (>)
	PCB170	0.834 %		0.889	1.8	*** (<)
	PCB180	0.532 %		0.887	1.81	*** (<)
	PCB118	0.494 %		0.919	2.29	** (<)
	PCB187	0.27 %		0.922	2.13	*** (<)
Mixture 1	PCB153	14.00%	13.4%	0.834	1.81	*** (<)
	PCB146	11.90%		0.922	2.31	*** (<)
	PCB180	11.00%		0.887	1.81	*** (<)
	PCB138	10.90%		0.855	2.08	*** (<)
	PCB170	10.70%		0.889	1.8	*** (<)
	PCB187	10.00%		0.922	2.13	*** (<)
	PCB118	8.33%		0.919	2.29	** (<)
	PFHXS	3.73%		0.829	1.85	*** (<)
	PFNA	3.01%		1.07	3.28	** (<)
	PFOS	2.76%		1.01	2.51	** (<)
	HCB	2.73%		1.91	3.59	
	PFOA	2.22%		0.989	2.57	** (<)
	DDE	1.66%		1.93	7.72	
	Tl	1.41%		1.79	5.1	
	As	1.39%		1.31	2.87	
	Pb	1.29%		3.03	6.76	
	Cu	1.22%		2.82	7.47	* (>)
	Mn	1.05%		4.67	10.7	*** (>)
	Cd	0.73%		4.56	17.5	* (>)

Mixtures	Substances	Substance weight	Cluster 4 - 85 mothers			
			Mixture weight	Exposure		
				Mean	P95	Compared to all
Mixture 5	PFHXS	65.6 %	39.7%	3.11	7.32	*** (>)
	PFOS	14.6 %		2.28	5.89	* (>)
	PCB118	5.77 %		1.43	3.36	
	As	5.35 %		1.72	5.73	
	Cu	3.87 %		1.62	3.44	
	PCB138	3.06 %		1.28	2.88	
	PCB146	1.44 %		1.19	2.37	* (<)
	Mn	0.346 %		1.29	3.24	
Mixture 1	PCB153	14.00 %	21.9%	1.18	2.26	
	PCB146	11.90 %		1.19	2.37	* (<)
	PCB180	11.00 %		1.07	1.97	* (<)
	PCB138	10.90 %		1.28	2.88	
	PCB170	10.70 %		1.09	2.01	* (<)
	PCB187	10.00 %		1.15	2.67	* (<)
	PCB118	8.33 %		1.43	3.36	
	PFHXS	3.73 %		3.11	7.32	*** (>)
	PFNA	3.01 %		1.86	4.87	
	PFOS	2.76 %		2.28	5.89	* (>)
	HCB	2.73 %		1.26	3.33	
	PFOA	2.22 %		1.99	4.32	
	DDE	1.66 %		2.22	7.49	
	TI	1.41 %		1.59	4.6	
	As	1.39 %		1.72	5.73	
	Pb	1.29 %		1.57	3.86	
	Cu	1.22 %		1.62	3.44	
	Mn	1.05 %		1.29	3.24	
	Cd	0.73 %		1.2	3.52	* (<)
Mixture 2	PFOA	39 %	14.8%	1.99	4.32	
	PFNA	37.7 %		1.86	4.87	
	HCB	9.46 %		1.26	3.33	
	Mn	5.26 %		1.29	3.24	
	Cu	5.02 %		1.62	3.44	
	As	1.6 %		1.72	5.73	
	PCB187	0.677 %		1.15	2.67	* (<)
	PCB118	0.66 %		1.43	3.36	
	PCB138	0.289 %		1.28	2.88	
	PCB146	0.166 %		1.19	2.37	* (<)
	PFOS	0.118 %		2.28	5.89	* (>)

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebrecht, Madeline Carsique, Eva Govarts, et al.	Page: 83

### 9.3 Annex 3: List of substances used for toxicity weighting in network analyses

Substance (group)	CAS-Biomarker	Biomarker Acronym	Biomarker - Full Name	HB-HBGV found
Acrylamide	81690-92-8	AAMA	N-acetyl-S-(2-carbamoyl)ethyl)cysteine	yes
Acrylamide	GAMA	GAMA	N-Acetyl-S-(2-carbamoyl-2-hydroxyethyl)cysteine	no
Antioxidants	128-37-0	BHT	butylated hydroxytoluene	no
Aprotic solvents	104612-35-3	2-HMSI	2-hydroxy-N-methylsuccinimide	yes
Aprotic solvents	41194-00-7	5-HNMP	5-Hydroxy-N-methyl-2-pyrrolidone	yes
Aprotic solvents	63467-80-1	2-HESI	2-hydroxy-N-ethylsuccinimide	yes
Benzol	4775-80-8	SPMA	S-Phenylmercapturic acid	no
Bisphenols	80-05-7	BPA	bisphenol A	yes
Fragrances	1320-16-7	TBBA	tert-butylbenzoic acid	no
Heavy metals	7439-97-6	Hg	mercury	yes
Heavy metals	7440-43-9	Cd	cadmium	yes
Heavy metals	7440-47-3	Cr (total)	chromium (total)	yes
Metalloids	64436-13-1	AsB	arsenobetaine	no
Metalloids	7440-36-0	Antimony	antimony	no
Metalloids	7440-38-2	As(total)	arsenic (total) from (iAs + MMA + DMA)	yes
Metalloids	75-60-5	DMA	dimethylarsenic	yes
Metalloids	7784-46-5	As(III)	arsenous acid	no
Non-metals	7782-49-2	Se	selenium	yes
PAHs	135-19-3	2-naphthol	2-hydroxynaphthalene	no
PAHs	2433-56-9	1-PHEN	1-hydroxyphenanthrene	no
PAHs	2443-58-5	2-FLUO	2-hydroxyfluorene	no
PAHs	484-17-3	9-PHEN	9- hydroxyphenanthrene	no
PAHs	5315-79-7	1-PYR	1-hydroxypyrene	yes
PAHs	605-55-0	2-PHEN	2- hydroxyphenanthrene	no
PAHs	605-87-8	3-PHEN	3- hydroxyphenanthrene	no

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebrecht, Madeline Carsique, Eva Govarts, et al.	Page: 84

Substance (group)	CAS-Biomarker	Biomarker Acronym	Biomarker - Full Name	HB-HBGV found
PAHs	7651-86-7	4-PHEN	4- hydroxyphenanthrene	no
PAHs	90-15-3	1-naphthol	1-hydroxynaphthalene	no
Parabens	120-47-8	EtP	ethylparaben	no
Parabens	99-76-3	MetP	methylparaben	no
Phthalates	131-70-4	MnBP	mono-n-butyl phthalate	yes
Phthalates	2306-33-4	MEP	monoethyl phthalate	yes
Phthalates	2528-16-7	MBzP	monobenzyl phthalate	yes
Phthalates	30833-53-5	MiBP	monoisobutyl phthalate	yes
Phthalates	40321-98-0	5oxo-MEHP (MEOHP)	Mono(2-ethyl-5-oxo-hexyl) phthalate	yes
Phthalates	40321-99-1	5OH-MEHP (MEHHP)	mono(2-ethyl-5-hydroxyhexyl) phthalate	yes
Phthalates	40809-41-4	5cx-MEHP (MECPP)	Mono(2-ethyl-5-carboxy- pentyl) phthalate	no
Phthalates	4376-18-5	MMP	mono-methyl phthalate	no
Phthalates	4376-20-9	MEHP	mono(2-ethylhexyl) phthalate	yes
Phthalates	5cx-MEPTP	5cx-MEPTP	mono(2-ethyl-5-carboxypentyl) terephthalate	yes
Phthalates	5OH-MEHTP	5OH-MEHTP	mono(2-ethyl-5-hydroxyhexyl) terephthalate	no
Phthalates	5oxo-MEHTP	5oxo-MEHTP	mono(2-ethyl-5-oxo-hexyl) terephthalate	no
Phthalates	936022-00-3	oxo-MiNP (MONP) (MOiNP)	monooxoisononyl phthalate	yes
Phthalates	936022-02-5	cx-MiNP (MCOP) (MCiOP)	mono(carboxyisooctyl) phthalate	yes
Phthalates	cx-MiDP	cx-MiDP (MCNP)	mono(2,7-methyl-7- carboxy-heptyl) phthalate	no
Phthalates	cx-MINCH	cx-MINCH (MCOCH)	cyclohexane-1,2- dicarboxylate-mono-(7- carboxylate-4- methyl)heptyl ester	yes
Phthalates	OH-MiBP	OH-MiBP	2-OH-mono-iso- butylphthalate	no
Phthalates	OH-MiDP	OH-MiDP (MHiDP)	mono-(hydroxyisodecyl) phthalate	no
Phthalates	OH-MINCH	OH-MINCH (MHNCH)	cyclohexane-1,2- dicarboxylate-mono-(7- hydroxy-4-methyl)octyl ester	yes
Phthalates	OH-MiNP	OH-MiNP (MHNP) (MHiNP)	monohydroxyisononyl phthalate	yes



D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebrecht, Madeline Carsique, Eva Govarts, et al.	Page: 85

Substance (group)	CAS-Biomarker	Biomarker Acronym	Biomarker - Full Name	HB-HBGV found
Phthalates	OH-MnBP	OH-MnBP (MHBP)	3-OH-mono-n-butyl phthalate, mono-(3-hydroxybutyl) phthalate	no
Phthalates	oxo-MiDP	oxo-MiDP (MOiDP)	mono-oxo-iso-decyl phthalate	no
Phthalates	oxo-MINCH	oxo- MINCH (MONCH)	cyclohexane-1,2- dicarboxylate-mono-(7-oxo- 4-methyl)octyl ester	no
Phthalates	oxo-MPHP	oxo-MPHP	oxo-mono(propylheptyl) phthalate	yes
Preservatives	NMMA	NMMA	N-methylmalonamic acid	no

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebet, Madeline Carsique, Eva Govarts, et al.	Page: 86

## 9.4 Annex 4: Health-Based Human Biomonitoring Guidance Values

### Description

**Table A4.1: Environmental biological guidance values**

Environmental HBHBGV (Acronym)	Country	Description
<b>Human Biomonitoring guidance value for the general population</b> (HBM-GV <sub>GenPop</sub> )	Germany	The HBM-GV(GenPop) represents the concentration of a substance or its specific metabolite(s) in human biological media (e.g. urine, blood, hair) at and below which, according to current knowledge, there is no risk of health impairment anticipated, and consequently no need for action. They are equivalent to the HBM-I values from the German Human Biomonitoring Commission. They are derived within HBM4EU for priority substances.  (Apel et al. 2020)
<b>Human Biomonitoring I</b> (HBM-I)	Germany	The HBM-I value, derived on the basis of toxicological and epidemiological studies, represents the concentration of a substance in human biological material below which – according to the knowledge and judgement of the HBM Commission – there is no risk for adverse health effects and, consequently, no need for action. (German Human Biomonitoring Commission)
<b>Human Biomonitoring II</b> (HBM-II)	Germany	The HBM-II-value derived on the basis of toxicological and epidemiological studies, represents the concentration of a substance in a human biological material above which – according to the knowledge and judgement of the HBM Commission – there is an increased risk for adverse health effects and, consequently, an acute need for exposure reduction measures and the provision of biomedical advice. The HBM-II-value should thus be regarded as an intervention or action level.  (German Human Biomonitoring Commission)
<b>Biomonitoring equivalents (BE)</b>	-	The concentration or range of concentrations of a chemical or its metabolite in a biological medium (blood, urine, or other medium) that is consistent with an existing health-based exposure guideline. This could be a non-cancer health-based guidance value such as a reference dose (RfD) or tolerable or acceptable daily intake (TDI or ADI) or a cancer-based exposure guidance value such as a risk-specific dose (e.g. the dose associated with a $1 \times 10^{-4}$ cancer risk).  (Hays and Aylward, 2012).

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebrecht, Madeline Carsique, Eva Govarts, et al.	Page: 87

**Table A4.2: Occupational biological guidance values**

Occupational HBHBGV (Acronym)	Country	Description
<b>Human Biomonitoring guidance value workers (HBM-GV<sub>Worker</sub>)</b>	-	<p>The HBM-GV(Worker) represents the concentration of a substance or its relevant metabolite(s) in human biological media aiming to protect workers exposed to the respective substance regularly (each work day), and over the course of a working life from the adverse effects related to medium- and long-term exposure. They are derived within HBM4EU for priority substances.</p> <p>(Apel et al. 2020).</p>
<b>Biologische Arbeitsstoff-Toleranzwerte(Biological Tolerance Values) (BAT)</b>	Germany	<p>The German Commission establishes BAT values ("Biologische Arbeitsstoff-Toleranzwerte": biological tolerance values) and BLW ("Biologische Leit-Werte") to enable the evaluation of the risk to an individual's health which results from exposure to a substance at the workplace.</p> <p>The BAT value describes the occupational-medical and toxicological derived concentration for a substance, its metabolites or an effect parameter in the corresponding biological material at which the health of an employee generally is not adversely affected even when the person is repeatedly exposed during long periods. BAT values are based on a relationship between external and systemic exposure or between the systemic exposure and the resulting effect of the substance. The derivation of the BAT value is based on the average of systemic exposures. (DFG 2020)</p>
<b>Biological Exposure Indices (BEI)</b>	New Zealand	<p>The BEI is a limit value that represents the level of exposure the typical worker can experience without adverse health effects. The values are established by the American Conference of Governmental Industrial Hygienists (ACGIH). (New Zealand Government 2020)</p>
<b>Biological Limit Values (BLV)</b>	France	<p>The biological limit values are recommended by ANSES as biological exposure markers which are considered to be relevant in the workplace. They are intended to protect workers from harmful effects related to exposure to the chemical in question, over the medium- or long-term. They take into account repeated exposure throughout a worker's working life.</p> <p>Depending on the available data, the recommended biological limit values do not all have the same meaning:</p> <ul style="list-style-type: none"> <li>▪ if the body of scientific evidence is sufficient to quantify a dose/response relationship with certainty, the biological limit values (BLVs) will be established on the basis of health data (no effect for threshold substances or risk levels for non-threshold carcinogens);</li> <li>▪ in the absence of such data for substances with threshold effects, BLVs are calculated on the basis of the expected</li> </ul>

D15.7: Report on re-analysis of existing HBM data for priority substances to identify mixtures	Security: Public
WP15: Mixtures, HBM and human health risks	Version: 1.1
Authors: Mirjam Luijten, Erik Lebet, Madeline Carsique, Eva Govarts, et al.	Page: 88

		<p>concentration of the biomarker of exposure (BME) when the worker is exposed to the 8-hour OEL.</p> <ul style="list-style-type: none"> <li>For carcinogens, in the absence of sufficient quantitative data, the biological limit value is calculated on the basis of another effect (pragmatic BLV).</li> </ul> <p>These last values do not guarantee the absence of health effects, but aim to limit exposure to these substances in the workplace. (ANSES 2013, ANSES 2015)</p>
<b>Biologische Leit-Werte (Biological Guidance Value) (BLW)</b>	Germany	<p>The German Commission establishes BAT values ("Biologische Arbeitsstoff-Toleranzwerte": biological tolerance values) and BLW ("Biologische Leit-Werte") to enable the evaluation of the risk to an individual's health which results from exposure to a substance at the workplace.</p> <p>The BLW ("Biological guidance value") is the amount of a chemical substance or its metabolites or the deviation from the norm of biological parameters induced by the substance in exposed humans which serves as an indicator for necessary protective measures. BLWs are assigned only for hazardous materials for which the available toxicological or occupational-medical data are insufficient for the establishment of BAT values (i. e. for carcinogenic substances and suspected carcinogens in the categories 1 to 3 and for non- carcinogens for which the toxicological data are inadequate).</p> <p>BLW values are generally established on the assumption that persons are exposed at work for at most 8 hours daily and 40 hours weekly during their working lives).</p> <p>The BLW is based on occupational-medical information as to the effects of handling the hazardous material together with toxicological data. (DFG, 2020)</p>
<b>Biological Monitoring Guidance Value (BMGV)</b>	United Kingdom	<p>These have been established by the Health and Safety Executive UK. BMGVs are either based on a relationship between biological concentrations and health effects, between biological concentrations and exposure at the level of the Workplace Exposure Limits, or on data collected from a representative sample of workplaces correctly applying the principles of good occupational hygiene practice.</p> <p>The types of data that are available will vary between substances and therefore the route taken to derive the BMGV will vary between substances. (HSE UK, 2020)</p>