Approaches to identify exposure to real-life chemical mixtures in the general population

Ilse Bente Ottenbros

Approaches to identify exposure to real-life chemical mixtures in the general population

Ilse Ottenbros

Ilse Bente Ottenbros
Approaches to identify exposure to real-life chemical mixtures in the general population

Doctoral thesis, Utrecht University

Cover image: Midjourney
Lay-out: RON Graphic Power | www.ron.nu
Printing: ProefschriftMaken | www.proefschriftmaken.nl

ISBN: 978-90-393-7582-2
DOI: https://doi.org/10.33540/1943

Approaches to identify exposure to real-life chemical mixtures in the general population

Methoden voor het identificeren van de werkelijke blootstelling aan mengsels van chemische stoffen in de algemene bevolking

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

woensdag 1 november 2023 des ochtends te 10.15 uur

door

Ilse Bente Ottenbros

geboren op 26 september 1991
te Alkmaar

Promotoren:
Prof. dr. ir. R.C.H. Vermeulen
Prof. dr. ir. E. Lebret

Copromotor:
Dr. J.J. Vlaanderen

Beoordelingscommissie:
Prof. dr. C.R. Berkers
Prof. dr. J. Klánová
Prof. dr. ir. J. Kromhout
Prof. dr. ir. J. Legler
Prof. dr. A.P. van Wezel

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van het *Research and Innovation Program* van de Europese Unie, onder *grant agreement* no. 733032.

TABLE OF CONTENTS

Chapter 1

GENERAL INTRODUCTION

Necessity to expand current risk assessment from single to multiple chemicals exposure

We humans are exposed to a plethora of manufactured chemical compounds every day throughout our entire lifetime. These exposures may come from various routes such as ambient environments, indoor and occupational environments, diet, consumer products, or medication. All these combined exposures potentially lead to an almost infinite number of combinations of compounds (Huhn et al., 2021; Vermeulen et al., 2020; Wild, 2012). Current exposure and risk assessment strategies often focus on single compounds, and when multiple compounds are considered, they are mostly originating from the same chemical family, such as phthalates, PCBs or dioxins (Bopp et al., 2018; Drakvik et al., 2020). With an increasing number of chemicals on the market, a mixture exposure and risk assessment approach is becoming more and more important (Bopp et al., 2019; Kienzler et al., 2016). Besides, new chemicals and new applications of existing chemicals are continuously introduced to the market, resulting in an exponential growth of the number of chemicals and potential combinations. This together makes a strong need to address the risks of chemical mixtures in relation to human health within the current policy and regulatory fields (Bopp et al., 2019; European Commission, 2020; Louro et al., 2019). Concurrently, developments in high-resolution analytical technologies (such as suspect screening, see Box 1) allow us to detect an increasing amount of chemicals.

The most prominent challenges in the risk assessment of chemical mixtures are i) how to measure combined exposures to chemical mixtures, ii) how to accurately assess the health risks from the chemical mixtures someone is exposed to, and iii) how to translate these risks to the policy and regulatory fields. This thesis focusses on the first point, how to measure and describe exposure patterns of chemical mixtures, focusing on co-occurrence of chemicals.

What is a chemical mixture

When discussing the term 'chemical mixtures', researchers from different fields of expertise may interpret it differently. Also, in literature different terms are used when referring to chemical mixtures, some call it 'multiple exposures' (Bopp et al., 2018), others 'cumulative exposures' (US EPA, 2003), and by definition chemical mixtures are covered in the exposome concept (Vineis et al., 2020; Wild, 2005, 2012). Common rationales to group chemicals include their chemical family, exposure route, use category, or supposed working mechanism. This thesis focusses on the combination of manufactured chemicals co-occurring within the same individual, or sample taken at the individual level, reflecting internal/external exposure to multiple chemicals. This includes multiple sources and routes of exposure. A cluster of correlated chemicals (in relation to the remaining chemicals measured) is considered as one chemical mixture. This perspective of a chemical mixture does include protracted and sequential exposures of long-lived chemicals, but might exclude (sequential) exposure to short-lived chemicals depending on the time of sampling (Santos et al., 2020).

Measuring exposure to chemical mixtures in the population

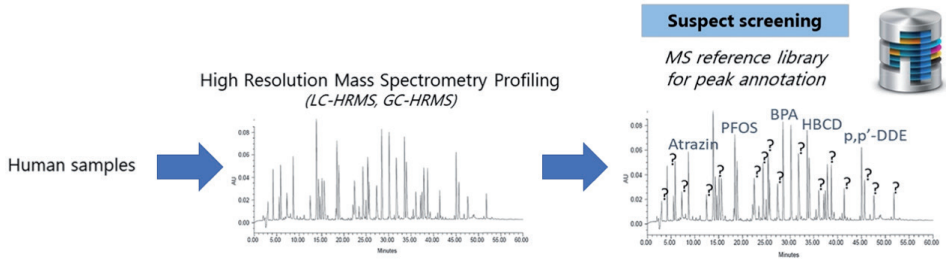
Measuring exposure to real-life chemical mixtures can be challenging, specifically to measure all potential sources and pathways. This would imply measuring or modelling all these aspects per chemical separately; both strategies require a substantial amount of time and effort to

collect all necessary data. Besides, large individual differences occur due to life-style or other personal factors affecting exposure, such as living location, age, or social-economic status. To fully reflect the chemical mixture at individual level, data collection at individual level is key, and preferable a combination of multiple measurements should be included to fully capture the complete chemical burden. These individual measures can be external (such as silicone wristbands), or internal (such as human biomonitoring). To conduct external individual measurements at population level these must be easy to apply and scalable. Passive monitoring would for these reasons be preferred above active monitoring. One method to passively measure chemical mixtures at the individual level are silicone wristbands, which provide the opportunity to measure over a longer period of time with minimal burden to the participant (O'Connell et al., 2014; Samon et al., 2022). In the extract of the wristbands a few hundred chemicals can be measured (Wacławik et al., 2022). However, wristbands only reflect the exposure routes inhalation and dermal. For an aggregated exposure measure including ingestion, internal measures would be preferred. Human biomonitoring (HBM) reflects the internal exposure concentration of chemicals in human tissues (Zare Jeddi et al., 2022), providing an efficient measurement approach to assess co-occurrence of chemical mixtures internally. Within HBM samples a wide range of chemicals can potentially be analyzed, although most are rapidly metabolized and biomarkers of exposure might not always be readily available. Besides, the application of HBM in population-based studies is often limited by the selection of a set of chemicals (targeted measures), the number of samples (due to participant burden, ethical aspects (invasive), or difficulties collecting and storing the samples), and often focus on a specific situation such as occupational settings or certain age groups. Exceptions of European population-based HBM studies covering a wide range of chemicals are for example the German Environmental Survey (GerES, including phthalates, DINCH, PFAS, bisphenols, PAHs, heavy metals, acrylamide, pesticides, aprotic solvents, UV filters and flame retardants) (Schwedler et al., 2020), and the Flemish Environment and Health Studies (FLEHS, including phthalates, PFAS, bisphenols, PAHs, flame retardants, heavy metals, pesticides, PCBs and dioxins) (Govarts et al., 2020; Schoeters et al., 2017). An example of an exception outside Europe is a large HBM dataset from the US, the National Health and Nutrition Examination Survey (NHANES), within which a method was developed to identify the most prevalent mixtures in human by applying frequent itemset mining (finding frequent patterns in the data). Ninety chemical combinations were identified, consisting of relatively few chemicals that occur in at least 30% of the US population (Kapraun et al., 2017).

To overcome the limitation of measuring a selected set of markers within HBM samples, the application of high-resolution analytical advances as suspect screening (SS) (see Box 1) makes it possible to identify a larger set of relevant chemicals (Pourchet et al., 2020; Walker et al., 2019). SS measures co-occurring exposures (parent chemicals plus metabolites) in HBM samples such as urine, also including chemicals for which no analytical reference standard is readily available (Huber et al., 2022). For SS, supporting annotation databases and follow-up work are necessary to assign every detected signal to an identified chemical with a given confidence level (Pourchet et al., 2020; Schymanski et al., 2014).

Box1. Suspect screening approach (Text and figure adapted from Pourchet et al. 2020)

Suspect Screening (SS) is typically based on high-resolution mass spectrometry (HRMS), which resulting peak-profile (defined by the molecular features: accurate mass, retention time and mass spectrum) is matched with a reference library for peak annotation. SS qualitatively assesses detection rates or semi-quantitative data. SS approaches are useful in analyzing a large set of exposure markers, enabling a better description of the exposure pattern; also, SS can be useful for prioritizing future (targeted) developments.



Approaches of this thesis to identify chemical mixtures

This thesis presents three approaches to assess real-life co-occurrence patterns in the general population. These approaches can be distinguished as, i) visualization of chemical mixtures, ii) measuring real-life chemical mixtures at individual level, and iii) application of an analytical screening approach to assess pesticide mixtures. The first part of the thesis concentrates on visualization techniques applied on existing HBM data. By visualization of co-occurring compounds, clusters in exposure markers (chemical mixtures) can be identified. The second part describes the application of the second and third approach on pesticide mixtures, for which samples were collected (silicone wristbands and urine samples) and SS was applied on the collected urine samples.

i) Visualization of chemical mixtures

Due to analytical advancements as SS, higher dimensional exposure data will become available. To describe patterns in these type of data statistical approaches such as variable selection (sparse linear regression) or principal component analysis can be applied (Stafoggia et al., 2017). Another approach is graphical modeling, such as correlation network models combined with a clustering algorithm, which describe and summarize co-occurrence patterns by highly correlated groups or clusters of compounds, where correlations are low between and high within clusters (Huhn et al., 2021; Santos et al., 2020). Correlation network analyses are increasingly being used in bioinformatics, for example to identify clusters in highly correlated genes (Friedman et al., 2012; Langfelder et al., 2008). An advantage of these network approaches is that they are intuitive to interpret, considering the interdependencies between all compounds included in the network model. Also, these networks can be used to assess differences between strata by covariates such as time or smoking status.

ii) Measuring pesticide mixtures at individual level

To describe real-life exposure to pesticide mixtures accurately, exposure measures at the individual level are preferred. To fully characterize the complete chemical burden, it is necessary to use a combination of different measurement technologies. Measurement technologies such as HBM and external personal monitoring could be one of these technologies to be used. HBM is effective to capture chemical mixtures internally (Ganzleben et al., 2017; Louro et al., 2019; Zare Jeddi et al., 2022), and measurements such as those in blood or urine provide an aggregated measure of exposure from different exposure routes (inhalation, ingestion, dermal) and sources. While this aggregated measure could be a very efficient manner to capture the mixture, collecting HBM data from large populations and over longer periods of time remains a challenge. Moreover, compounds with a very short lifetime within the body (due to excretion and or metabolization) can easily be missed through HBM. External personal monitoring such as silicone wristbands are a passive sampling method which is easy to implement and can be worn for several days or weeks, allowing larger timeframes due to their low-impact to the wearer. A large set of chemicals can be measured in wristbands, reflecting exposure from the entire time of wearing. By the different nature of measure, wristbands do however only reflect the inhalation and dermal exposure routes and do not capture dietary intake.

iii) Application of an analytical screening approach on pesticide mixtures

Novel and harmonized SS approaches based on high resolution mass spectrometry are efficient in detecting a broad range of exposure markers, being able to capture the complexity of chemical mixtures (Huber et al., 2022; Pourchet et al., 2020; Sobus et al., 2018). Specifically in HBM samples with often limited volumes, SS enables the detection of more chemicals within the same sample, and are relatively quick compared to targeted measurements.

Real-life chemical mixture example: pesticides

The second half of this thesis focusses on co-occurring pesticides as a prime example of mixtures. For this, pesticides are measured in silicone wristbands and in urine samples. On the urine samples an SS approach will be applied to detect a large set of biomarkers. Pesticides are a relevant chemical mixture due to their societal attention and worries (Schaub et al., 2020), as well as a point of departure by EFSA for future development of strategies to evaluate mixtures (EFSA, 2013). Also, pesticides have inherently toxic properties, such as neurotoxicity, reproductive issues, and developmental problems (Kim et al., 2017; Kori et al., 2020; Mostafalou et al., 2017), which have resulted in increased regulatory attention (EFSA, 2014, 2016). Pesticides often occur as mixtures, for example by application of farmers with multiple pesticides simultaneously, as well as by dietary intake, due to the large variety present in food items. In real-life, dietary intake can co-occur with potential residential and/or occupational exposures (Carles et al., 2017; Deziel et al., 2015; Rizzati et al., 2016). In the literature there is growing evidence that living in agricultural areas where pesticides are applied contributes to higher pesticide exposure (Dereumeaux et al., 2020; Figueiredo et al., 2021; López-Gálvez et al., 2019). Despite

a large body of literature on dietary exposure to pesticides (Ntzani et al., 2013; Oates et al., 2011), it is still unclear to which combinations of pesticides the general population in Europe is exposed. Moreover, most pesticide exposure measurement campaigns focus on a limited set of targeted pesticides, mainly pyrethroids and non-specific markers of organophosphorus pesticides. Given the long list of registered pesticides, many targeted assays would be required to assess the presence of all urinary pesticides and their metabolites in each sample. Another option to assess pesticide mixtures at individual level would be to collect external personal samples in which a large number of targeted pesticides could be measured, such as silicone wristbands. However, to reflect the internal pesticide mixture (including diet), urine samples are more commonly used.

Layout of this thesis

With the application of the abovementioned three approaches, in my thesis I aim to:

- Describe which chemical mixtures can be detected in European HBM data, including their variation across measurement campaigns or covariates such as smoking status.
- Assess the applicability of correlation networks to visualize patterns and clusters in chemical mixtures based on European HBM data.
- Present two measurement methods to assess pesticides mixtures, HBM and silicone wristbands.
- Address exposure to pesticide mixture exposure patterns in six European countries, with a specific focus on variations in these patterns by living location, season and age.
- Describe and interpret differences in pesticide exposure levels, by various covariates such as pesticide usage, distance to agriculture and diet.

Part one of this thesis addresses the visualization of chemical mixtures graphically by correlation network analysis. Correlation networks provide an intuitive graphical approach to describe which biomarkers in the dataset are (in)dependent of each other, these networks are combined with a hierarchical clustering approach to detect separate groups or communities within these networks. In **chapter 2,** the correlation network approach is applied to the Belgium FLEHS datasets, illustrating the applicability of correlation networks in HBM type of datasets. The network approach is described in detail, combined with the application of a community detection algorithm. Differences between networks are used to assess variability across sampling campaigns, or covariates as smoking status (comparative network analysis). In **chapter 3,** the network approach with the community detection algorithm is applied on a larger scale, on four different cohorts across Europe (Belgium, Czech Republic, Germany and Spain). The applicability of networks in different settings is shown, including comparative network analysis across covariates. This chapter also covers some methodological aspects of the application of the network approach on HBM data: a comparison between weighted (measure of correlation) and unweighted (correlated yes/no) networks, and the effect of creatine correction of urinary biomarkers on the estimated network are shown.

The second part addresses exposure to pesticide mixtures. In **chapter 4**, a measurement method is described to assess long term exposure (two months) to pesticide mixtures at the

individual level by application of silicone wristbands. This matrix allows for a large set of pesticides to be measured in a targeted manner, providing insight into which combinations of pesticides the wearers were exposed to. In **chapters 5 and 6** a pesticide suspect screening approach was applied on urine samples in six European countries, as part of the SPECIMEn study. This study was initiated to generate new pesticide exposure data in a harmonized European setting. **Chapter 5** describes the general study protocol of the SPECIMEn study, detection rates, and co-occurrence of multiple pesticides measured in the urine samples. By design of the study, pesticide exposure patterns by age, location and season were compared. In **chapter 6**, the semi-quantitative pesticide SS data was used to explore exposure trends in the data, and exposure routes and sources underlying the large set of pesticides detected in samples from the Netherlands and Switzerland. Exposure sources as pesticide usage, distance to agriculture and diet were considered as covariates in regression models.

Chapters 2, 3, 5 and 6 were conducted in the context of the European Initiative for Human Biomonitoring (HBM4EU). **Chapter 4** was conducted in the context of the Dutch OBO project (Onderzoek Bestrijdingsmiddelen Omwonenden).

In **chapter 7**, the results of the three approaches to identify chemical mixtures are discussed, as well as their applicability in future work, limitations and alternatives. The lessons learnt from the work performed in this thesis are drawn. The challenges in the mixture exposure characterization are discussed, and a perspective will be given on the next steps in this challenging field.

REFERENCES

Bopp, S. K., Barouki, R., Brack, W., Dalla Costa, S., Dorne, J.-L. C. M., Drakvik, P. E., … Bergman, Å. (2018). Current EU research activities on combined exposure to multiple chemicals. *Environment International*, *120*, 544–562. https://doi.org/10.1016/j.envint.2018.07.037

Bopp, S. K., Kienzler, A., Richarz, A.-N., van der Linden, S. C., Paini, A., Parissis, N., & Worth, A. P. (2019). Regulatory assessment and risk management of chemical mixtures: challenges and ways forward. *Critical Reviews in Toxicology*, *49*(2), 174–189. https://doi.org/10.1080/10408444.2019.1579169

Carles, C., Bouvier, G., Lebailly, P., & Baldi, I. (2017). Use of job-exposure matrices to estimate occupational exposure to pesticides: A review. *Journal of Exposure Science & Environmental Epidemiology*, *27*(2), 125–140. https://doi.org/10.1038/jes.2016.25

Dereumeaux, C., Fillol, C., Quenel, P., & Denys, S. (2020). Pesticide exposures for residents living close to agricultural lands: A review. *Environment International*, *134*, 105210. https://doi.org/10.1016/j.envint.2019.105210

Deziel, N. C., Friesen, M. C., Hoppin, J. A., Hines, C. J., Thomas, K., & Freeman, L. E. B. (2015). A Review of Nonoccupational Pathways for Pesticide Exposure in Women Living in Agricultural Areas. *Environmental Health Perspectives*, *123*(6), 515–524. https://doi.org/10.1289/ehp.1408273

Drakvik, E., Altenburger, R., Aoki, Y., Backhaus, T., Bahadori, T., Barouki, R., … Bergman, Å. (2020). Statement on advancing the assessment of chemical mixtures and their risks for human health and the environment. *Environment International*, *134*, 105267. https://doi.org/10.1016/j.envint.2019.105267

EFSA. (2013). Scientific Opinion on the identification of pesticides to be included in cumulative assessment groups on the basis of their toxicological profile. *EFSA Journal*, *11*(7). https://doi.org/10.2903/j.efsa.2013.3293

EFSA. (2014). Conclusion on the peer review of the pesticide risk assessment of the active substance tebuconazole. *EFSA Journal*, *12*(1). https://doi.org/10.2903/j.efsa.2014.3485

EFSA. (2016). Peer review of the pesticide risk assessment of the active substance acetamiprid. *EFSA Journal*, *14*(11). https://doi.org/10.2903/j.efsa.2016.4610

European Commission. (2020). *Chemicals Strategy for Sustainability Towards a Toxic-Free Environment*.

Figueiredo, D. M., Duyzer, J., Huss, A., Krop, E. J. M., Gerritsen-Ebben, M. G., Gooijer, Y., & Vermeulen, R. C. H. (2021). Spatio-temporal variation of outdoor and indoor pesticide air concentrations in homes near agricultural fields. *Atmospheric Environment*, *262*, 118612. https://doi.org/10.1016/j.atmosenv.2021.118612

Friedman, J., & Alm, E. J. (2012). Inferring Correlation Networks from Genomic Survey Data. *PLoS Computational Biology*, *8*(9), e1002687. https://doi.org/10.1371/journal.pcbi.1002687

Ganzleben, C., Antignac, J.-P., Barouki, R., Castaño, A., Fiddicke, U., Klánová, J., … Kolossa-Gehring, M. (2017). Human biomonitoring as a tool to support chemicals regulation in the European Union. *International Journal of Hygiene and Environmental Health*, *220*(2), 94–97. https://doi.org/10.1016/j.ijheh.2017.01.007

Govarts, E., Portengen, L., Lambrechts, N., Bruckers, L., Den Hond, E., Covaci, A., … Vermeulen, R. (2020). Early-life exposure to multiple persistent organic pollutants and metals and birth weight: Pooled analysis in four Flemish birth cohorts. *Environment International*, *145*, 106149. https://doi.org/10.1016/j.envint.2020.106149

Huber, C., Nijssen, R., Mol, H., Philippe Antignac, J., Krauss, M., Brack, W., … Lommen, A. (2022). A large scale multi-laboratory suspect screening of pesticide metabolites in human biomonitoring: From tentative annotations to verified occurrences. *Environment International*, *168*, 107452. https://doi.org/10.1016/j.envint.2022.107452

Huhn, S., Escher, B. I., Krauss, M., Scholz, S., Hackermüller, J., & Altenburger, R. (2021). Unravelling the chemical exposome in cohort studies: routes explored and steps to become comprehensive. *Environmental Sciences Europe*, *33*(1), 17. https://doi.org/10.1186/s12302-020-00444-0

Kapraun, D. F., Wambaugh, J. F., Ring, C. L., Tornero-Velez, R., & Setzer, R. W. (2017). A Method for Identifying Prevalent Chemical Combinations in the U.S. Population. *Environmental Health Perspectives*, *125*(8), 087017. https://doi.org/10.1289/EHP1265

Kienzler, A., Bopp, S. K., van der Linden, S., Berggren, E., & Worth, A. (2016). Regulatory assessment of chemical mixtures: Requirements, current approaches and future perspectives. *Regulatory Toxicology and Pharmacology*, *80*, 321–334. https://doi.org/10.1016/j.yrtph.2016.05.020

Kim, K.-H., Kabir, E., & Jahan, S. A. (2017). Exposure to pesticides and the associated human health effects. *Science of The Total Environment*, *575*, 525–535. https://doi.org/10.1016/j.scitotenv.2016.09.009

Kori, R. K., Mandrah, K., Hasan, W., Patel, D. K., Roy, S. K., & Yadav, R. S. (2020). Identification of markers of depression and neurotoxicity in pesticide exposed agriculture workers. *Journal of Biochemical and Molecular Toxicology*, *34*(6). https://doi.org/10.1002/jbt.22477

Langfelder, P., & Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*(1), 559. https://doi.org/10.1186/1471-2105-9-559

López-Gálvez, Wagoner, Quirós-Alcalá, Horne, Furlong, Avila, & Beamer. (2019). Systematic Literature Review of the Take-Home Route of Pesticide Exposure via Biomonitoring and Environmental Monitoring. *International Journal of Environmental Research and Public Health*, *16*(12), 2177. https://doi.org/10.3390/ijerph16122177

Louro, H., Heinälä, M., Bessems, J., Buekers, J., Vermeire, T., Woutersen, M., … Santonen, T. (2019). Human biomonitoring in health risk assessment in Europe: Current practices and recommendations for the future. *International Journal of Hygiene and Environmental Health*, *222*(5), 727–737. https://doi.org/10.1016/j.ijheh.2019.05.009

Mostafalou, S., & Abdollahi, M. (2017). Pesticides: an update of human exposure and toxicity. *Archives of Toxicology*, *91*(2), 549–599. https://doi.org/10.1007/s00204-016-1849-x

Ntzani, E. E., Ntritsos G, C. M., Evangelou, E., & Tzoulaki, I. (2013). Literature review on epidemiological studies linking exposure to pesticides and health effects. *EFSA Supporting Publications*, *10*(10). https://doi.org/10.2903/sp.efsa.2013.EN-497

O'Connell, S. G., Kincl, L. D., & Anderson, K. A. (2014). Silicone Wristbands as Personal Passive Samplers. *Environmental Science & Technology*, *48*(6), 3327–3335. https://doi.org/10.1021/es405022f

Oates, L., & Cohen, M. (2011). Assessing Diet as a Modifiable Risk Factor for Pesticide Exposure. *International Journal of Environmental Research and Public Health*, *8*(6), 1792–1804. https://doi.org/10.3390/ijerph8061792

Pourchet, M., Debrauwer, L., Klanova, J., Price, E. J., Covaci, A., Caballero-Casero, N., … Antignac, J.-P. (2020). Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: From promises to challenges and harmonisation issues. *Environment International*, *139*, 105545. https://doi.org/10.1016/j.envint.2020.105545

Rizzati, V., Briand, O., Guillou, H., & Gamet-Payrastre, L. (2016). Effects of pesticide mixtures in human and animal models: An update of the recent literature. *Chemico-Biological Interactions*, *254*, 231–246. https://doi.org/10.1016/j.cbi.2016.06.003

Samon, S. M., Hammel, S. C., Stapleton, H. M., & Anderson, K. A. (2022). Silicone wristbands as personal passive sampling devices: Current knowledge, recommendations for use, and future directions. *Environment International*, *169*, 107339. https://doi.org/10.1016/j.envint.2022.107339

Santos, S., Maitre, L., Warembourg, C., Agier, L., Richiardi, L., Basagaña, X., & Vrijheid, M. (2020). Applying the exposome concept in birth cohort research: a review of statistical approaches. *European Journal of Epidemiology*, *35*(3), 193–204. https://doi.org/10.1007/s10654-020-00625-4

Schaub, S., Huber, R., & Finger, R. (2020). Tracking societal concerns on pesticides – a Google Trends analysis. *Environmental Research Letters*, *15*(8), 84049. https://doi.org/10.1088/1748-9326/ab9af5

Schoeters, G., Govarts, E., Bruckers, L., Den Hond, E., Nelen, V., De Henauw, S., … Baeyens, W. (2017). Three cycles of human biomonitoring in Flanders – Time trends observed in the Flemish Environment and Health Study. *International Journal of Hygiene and Environmental Health*, *220*(2), 36–45. https://doi.org/10.1016/j.ijheh.2016.11.006

Schwedler, G., Rucic, E., Lange, R., Conrad, A., Koch, H. M., Pälmke, C., … Kolossa-Gehring, M. (2020). Phthalate metabolites in urine of children and adolescents in Germany. Human biomonitoring results of the German Environmental Survey GerES V, 2014–2017. *International Journal of Hygiene and Environmental Health*, *225*, 113444. https://doi.org/10.1016/j.ijheh.2019.113444

Schymanski, E. L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H. P., & Hollender, J. (2014). Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environmental Science & Technology*, *48*(4), 2097–2098. https://doi.org/10.1021/es5002105

Sobus, J. R., Wambaugh, J. F., Isaacs, K. K., Williams, A. J., McEachran, A. D., Richard, A. M., … Newton, S. R. (2018). Integrating tools for non-targeted analysis research and chemical safety evaluations at the US EPA. *Journal of Exposure Science & Environmental Epidemiology*, *28*(5), 411–426. https://doi.org/10.1038/s41370-017-0012-y

Stafoggia, M., Breitner, S., Hampel, R., & Basagaña, X. (2017). Statistical Approaches to Address Multi-Pollutant Mixtures and Multiple Exposures: the State of the Science. *Current Environmental Health Reports*, *4*(4), 481–490. https://doi.org/10.1007/s40572-017-0162-z

US EPA. (2003). Framework for Cumulative Risk Assessment, (EPA/600/P-02/001F).

Vermeulen, R., Schymanski, E. L., Barabási, A.-L., & Miller, G. W. (2020). The exposome and health: Where chemistry meets biology. *Science*, *367*(6476), 392–396. https://doi.org/10.1126/science.aay3164

Vineis, P., Robinson, O., Chadeau-Hyam, M., Dehghan, A., Mudway, I., & Dagnino, S. (2020). What is new in the exposome? *Environment International*, *143*, 105887. https://doi.org/10.1016/j.envint.2020.105887

Wacławik, M., Rodzaj, W., & Wielgomas, B. (2022). Silicone Wristbands in Exposure Assessment: Analytical Considerations and Comparison with Other Approaches. *International Journal of Environmental Research and Public Health*, *19*(4), 1935. https://doi.org/10.3390/ijerph19041935

Walker, D. I., Valvi, D., Rothman, N., Lan, Q., Miller, G. W., & Jones, D. P. (2019). The Metabolome: a Key Measure for Exposome Research in Epidemiology. *Current Epidemiology Reports*, *6*(2), 93–103. https://doi.org/10.1007/s40471-019-00187-4

Wild, C. P. (2005). Complementing the Genome with an "Exposome": The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology. *Cancer Epidemiology, Biomarkers & Prevention*, *14*(8), 1847–1850. https://doi.org/10.1158/1055-9965.EPI-05-0456

Wild, C. P. (2012). The exposome: from concept to utility. *International Journal of Epidemiology*, *41*(1), 24–32. https://doi.org/10.1093/ije/dyr236

Zare Jeddi, M., Hopf, N. B., Louro, H., Viegas, S., Galea, K. S., Pasanen-Kase, R., … Bessems, J. (2022). Developing human biomonitoring as a 21st century toolbox within the European exposure science strategy 2020–2030. *Environment International*, *168*, 107476. https://doi.org/10.1016/j.envint.2022.107476

Chapter 2

Network Analysis to Identify Communities Among Multiple Exposure Biomarkers Measured at Birth in Three Flemish General Population Samples

Ilse Ottenbros[1,2]; Eva Govarts[3]; Erik Lebret[1,2]; Roel Vermeulen[2]; Greet Schoeters[3,4]; Jelle Vlaanderen[2]

[1] Centre for Sustainability, Environment and Health, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
[2] Institute for Risk Assessment Sciences (IRAS), Utrecht University, Utrecht, The Netherlands
[3] VITO Health, Flemish Institute for Technological Research (VITO), Mol, Belgium
[4] Dept of Biomedical Sciences, University of Antwerp, Antwerp, Belgium

Published in Frontiers in Public Health, February 2021
DOI: 10.3389/fpubh.2021.590038

ABSTRACT

Introduction: Humans are exposed to multiple environmental chemicals via different sources resulting in complex real-life exposure patterns. Insight into these patterns is important for applications such as linkage to health effects and (mixture) risk assessment. By providing internal exposure levels of (metabolites of) chemicals, biomonitoring studies can provide snapshots of exposure patterns and factors that drive them. Presentation of biomonitoring data in networks facilitates the detection of such exposure patterns and allows for the systematic comparison of observed exposure patterns between datasets and strata within datasets.

Methods: We demonstrate the use of network techniques in human biomonitoring data from cord blood samples collected in three campaigns of the Flemish Environment and Health Studies (FLEHS) (sampling years resp. 2002–2004, 2008–2009, and 2013–2014). Measured biomarkers were multiple organochlorine compounds, PFAS and metals. Comparative network analysis (CNA) was conducted to systematically compare networks between sampling campaigns, smoking status during pregnancy, and maternal pre-pregnancy BMI.

Results: Network techniques offered an intuitive approach to visualize complex correlation structures within human biomonitoring data. The identification of groups of highly connected biomarkers, "communities," within these networks highlighted which biomarkers should be considered collectively in the analysis and interpretation of epidemiological studies or in the design of toxicological mixture studies. Network analyses demonstrated in our example to which extent biomarker networks and its communities changed across the sampling campaigns, smoking status during pregnancy, and maternal pre-pregnancy BMI.

Conclusion: Network analysis is a data-driven and intuitive screening method when dealing with multiple exposure biomarkers, which can easily be upscaled to high dimensional HBM datasets, and can inform mixture risk assessment approaches.

INTRODUCTION

Throughout their life-time, humans are exposed to a plethora of environmental stressors and chemicals that independently or in interaction may have an impact on health. Whereas chemical risk assessment typically evaluates single compounds, it generally does not appropriately reflect the complexity of concomitant exposure to multiple chemicals in real life. Currently there is yet little insight into commonly occurring exposure mixtures and how these mixtures change between important covariates, e.g., gender, countries, and time. Human biomonitoring (HBM) has the potential to provide a snapshot of exposure to chemicals (Ganzleben et al., 2017), and these data can be used to screen for the presence of clusters of correlated exposures. The identification of these communities is important for the analysis and interpretation within epidemiological studies (which compounds are more related, and should therefore be considered collectively) and for the design of mixture toxicology studies (which combined exposures do occur in the population), thereby informing risk assessors/managers on potential concomitant exposure pathways.

Patterns between multiple biomarkers are not commonly presented (Tamayo-Uria et al., 2019). Increasingly, graphical representation of (partial) correlation patterns such as heatmaps or circular correlation globes (circos plots) are being used. However, here the distinction of groups of correlated compounds is not always straightforward as it depends largely on a-priori ordering by the presenter and on the visual interpretation by the reader. Also, the comparison of multiple circos plots (for example as presented in (Robinson et al., 2015)), is challenging, especially when comparing three or more plots or in high dimensional settings. Networks provide a graphical method to represent groups or communities in the data, which has been used widely in the OMICs world (Gehlenborg et al., 2010; Mitra et al., 2013; Villaveces et al., 2015). Applied to HBM data, networks consist of nodes which represent the biomarkers, and edges that represent the conditional dependence between the biomarkers. Networks give an intuitive interpretation of patterns in the data without prior assumptions (Green et al., 2018). A network may consist of multiple subnetworks (connected nodes). Within a subnetwork, one or more communities of biomarkers can be detected using community detection algorithms (Fortunato, 2010). Communities are groups in which nodes (i.e. biomarkers) are more connected to each other than to the rest of the (sub)network. Communities in exposure biomarker networks might therefore represent common exposure routes (dermal, inhalation or ingestion), external sources (such as lifestyle, social or environmental factors) and/or (bio) chemical properties (e.g. kinetics, distribution).

Further insights can be generated with comparative network analysis (CNA), which is an analytical procedure that allows for the comparison of two or more networks based on (dis) similarities (Emmert-Streib et al., 2016; Ideker et al., 2012; Zhang et al., 2009). CNA can be used to assess the impact of covariates on observed networks. Differences between networks are presented as (dis)similar nodes and edges, which in itself are amendable to community detection as well (Mall et al., 2017).

To pilot and illustrate the use of network techniques in exposure HBM data we applied this methodology to data collected as part of the FLEHS (Flemish Environment and Health

Study) newborn campaigns (Schoeters, Den Hond, et al., 2012). The FLEHS data consists of multiple biomarkers, obtained by targeted analysis of cord blood samples collected directly after birth in three subsequent campaigns over a 12 year period (Flemish Center of Expertise on Environment and Health, 2020; Koppen et al., 2009; Schoeters et al., 2017; Schoeters, Colles, et al., 2012; Schoeters, Den Hond, et al., 2012). Time trends of multiple biomarkers across the subsequent FLEHS newborn campaigns (Persistent Organic Pollutants (POPs) and metals) have been described before, showing varying rates of decline of different biomarker over the three campaigns (Schoeters et al., 2017).

We were particularly interested in the use of network techniques to visualize biomarker correlation patterns within each FLEHS campaign. In addition, we explored the stability of these networks across sampling campaigns, smoking status during pregnancy, and maternal pre-pregnancy BMI using CNA.

MATERIAL AND METHODS

Flemish Environment and Health Study (FLEHS)

In the newborn campaigns of FLEHS, cord blood samples have been collected at three points in time, FLEHS I (N=1196): 2002-2004, FLEHS II (N=255): 2008-2009 and FLEHS III (N=281): 2013-2014. The FLEHS campaigns are conducted in a population sample that is representative for the geographical distribution and the population density of the population in Flanders, Belgium. A summary of the characteristics of each campaign, including the p-value, is presented in Table S1 (Supplementary Material). Details of recruitment, sampling, laboratories, limits of detection and quality control measures have been reported before (Baeyens et al., 2014; Den Hond et al., 2009; Schoeters, Den Hond, et al., 2012). Selection of the chemicals was based on health and exposure related criteria, and technical criteria, extensively discussed by experts as part of the biomonitoring studies (Schoeters et al., 2017). The biomonitoring studies were approved by the Ethical Committee of the University of Antwerp (FLEHS I and II) and of the University hospital of Antwerp (FLEHS III).

Biomarkers

Chemicals measured in cord blood of newborns were included for analysis if more than 60% of the measurements was above the Limit of Detection (LOD). In FLEHS I, seven biomarkers fulfilled this requirement: cadmium, lead, *p,p'*-DDE, HCB, PCB138, PCB153 and PCB180. In FLEHS II, twelve biomarkers: cadmium, lead, *p,p'*-DDE, PCB138, PCB153 and PCB180, arsenic, copper, manganese, thallium, PFOS and PFOA. In FLEHS III, nineteen biomarkers fulfilled this requirement: all from FLEHS II plus the additional biomarkers: HCB, PCB118, PCB146, PCB170, PCB180, PFHXS and PFNA. For the CNA comparisons between the three campaigns six corresponding biomarkers were included, and between FLEHS II and III twelve corresponding biomarkers.

Imputations and data preparation

Concentrations of biomarkers were natural log transformed because distributions were skewed. *p,p'*-DDE, HCB, and PCB concentrations were expressed as concentrations per gram blood lipid and as such corrected for differences in dietary fat intake. Hence it is expected that the correlations are independent of blood fat levels (O'Brien et al., 2017). Biomarker values below LOD 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 and determinants (cholesterol, maternal age, maternal pre-pregnancy BMI, parity, singleton or multiples, and maternal smoking during pregnancy) were imputed by using a single imputation strategy stratified per campaign, using the R package *mice*. Determinants 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. The geometric mean, minimum and maximum (based on imputed data) biomarker concentrations, and percentage of missing samples are presented in Table S2 (Supplementary Material). Pearson correlation structures between the natural logarithm transformed biomarkers per sampling campaign are presented by heatmaps and circos plots in Figure 1 and the Supplementary Material (Figures S1-2).

For comparisons across sampling campaigns, analytical datasets were created in which biomarker concentrations were residualized using a linear model incorporating predictors for maternal age, pre-pregnancy BMI and maternal smoking during pregnancy, following the corrections described by Schoeters et al. (Schoeters et al., 2017). For comparisons across covariate categories of smoking and BMI, analytical datasets of FLEHS III were created. The datasets stratified by smoking were adjusted for maternal age and maternal pre-pregnancy BMI; the datasets of BMI strata for maternal age and smoking during pregnancy.

Network graph estimation and community detection

We used undirected and unweighted network analysis to describe the conditional independence between multiple variables, making use of the packages *huge* and *igraph*, using R (v3.5.0) (Csárdi et al., 2006; Zhao et al., 2012). A node in the network represents a biomarker, and an edge reflects conditional dependency given all other variables (Zhao et al., 2012). For comparison purposes, weighted network analysis was applied as well, making use of the package *EGAnet (v0.9.6)* (Hudson Golino et al., 2020).

The graph estimation was conducted using the graphical lasso, which involves penalized 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 L_1 penalty. The graphical lasso cycles through the variables, fitting a modified lasso regression to each variable in turn. Regularization 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 maximum lambda. Optimal lambda selection was conducted using the stability approach to regularization selection method (StARS) (Liu et al., 2010), 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* 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 et al., 2009; Pons et al., 2005). Nodes were colored according to the community they were assigned to. Sensitivity analysis was performed by comparing the networks with and without inclusion of the *mice* imputed values (samples missing at random, see table S2 for percentages of missings).

The low-dimensional setting of the FLEHS data also allows for the application of correlation networks. (Yu et al., 2019). We compared our approach to an application of weighted correlation networks for the data in FLEHS I, II, and III. Weighted networks were estimated by the *EGA*net package, this Exploratory Graph Analysis technique was based on the Graphical lasso model and an EBIC tuning parameter of 0.5 was used (Hudson Golino et al., 2020; Hudson F. Golino & Demetriou, 2017; Hudson F. Golino & Epskamp, 2017). A parametric bootstrap (1,000 iterations) was used to estimate the median network structure. Communities in the EGA network were estimated using the *walktrap* algorithm. 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).

Networks were constructed for each measurement campaign separately. Secondly, networks were constructed for different strata of the dataset of FLEHS III. Where FLEHS III was either split by maternal smoking status during pregnancy (yes n=33; no n=248), or by maternal pre-pregnancy BMI category (≤ 25 kg/m^2 or low-normal n=195; > 25 kg/m^2.

Comparative Network Analysis (CNA)

Systematically comparing networks, or CNA, is of interest to assess the impact of covariates on networks derived in HBM data. Networks can be compared on their similarities or their dissimilarities. Multiple network comparison methods have been described before, and some can be computationally challenging (Emmert-Streib et al., 2016; Tantardini et al., 2019). In this paper, we focus on exact graph matching, which involves the exact correspondence between two or more graphs with the exact same set of nodes. We call an edge 'conserved' if it is present in all of the input graphs. The complement of conserved edges is represented in a network graph (network of conserved edges). CNA can also assess the presence of edges in network B which are not present in network A. These results can be interpreted as 'additional' or different edges, and are presented in a network graph as well (network of differential edges). The CNA as applied in this paper focusses on differences in network structure, and not on differences in the detected communities. To assess the stability of the independently derived networks across the FLEHS sampling campaigns we conducted CNA to identify the conserved edges between the networks across campaigns. To evaluate the influence of covariates, differences between derived networks were assessed between the strata: high versus low-normal maternal pre-pregnancy BMI, and non-smoking versus smoking during pregnancy. Within the deduced conserved and differential networks multiple subnetworks were distinguished, within which the *walktrap* algorithm was applied for community detection (only if the subnetwork consisted of 6 or more nodes).

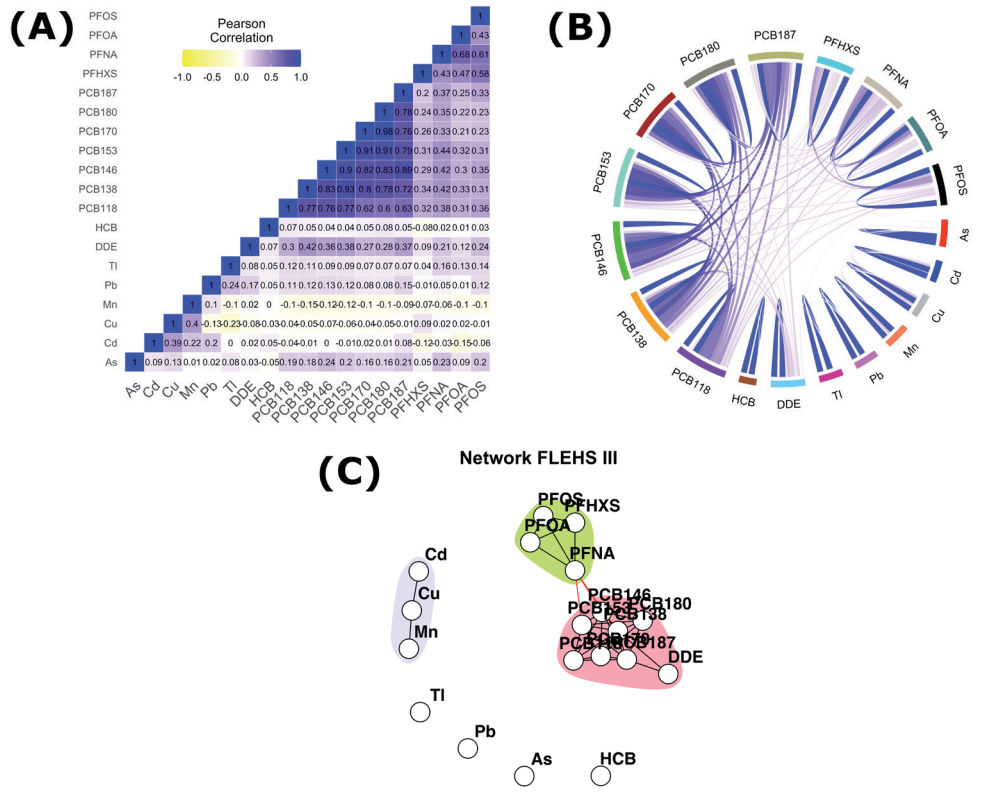


Figure 1. *Heatmap (A), circular correlation globe (B) and network including community detection (C) of FLEHS III, nineteen biomarkers, n=281. Data is corrected for maternal age, smoking during pregnancy and maternal pre-pregnancy BMI. The heatmap is based on Pearson correlation between the biomarkers. 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.*

RESULTS

The study population, summary statistics and time trends of individual biomarkers over the three sampling campaign have been described previously (Schoeters et al., 2017; Schoeters, Den Hond, et al., 2012). An overview of the study characteristics and the concentrations per biomarker are presented in Tables S1 and S2.

The FLEHS III dataset consists of nineteen biomarkers, and has been used to illustrate the network techniques since it is most data rich. Figure 1 presents the heatmap, correlation globe and network for the FLEHS III dataset. For comparison purposes, we present two alternative approaches to represent correlation structures in HBM data. In both the heatmap (Figure 1A) and the circular correlation globe (Figure 1B) correlation structures become apparent. The identification of communities of strongly correlated markers using these visualizations is not straight forward as it depends largely on the subjective interpretation of the reader. The heatmaps and correlation globes for FLEHS I and II are presented in figures S1 and S2 A-B.

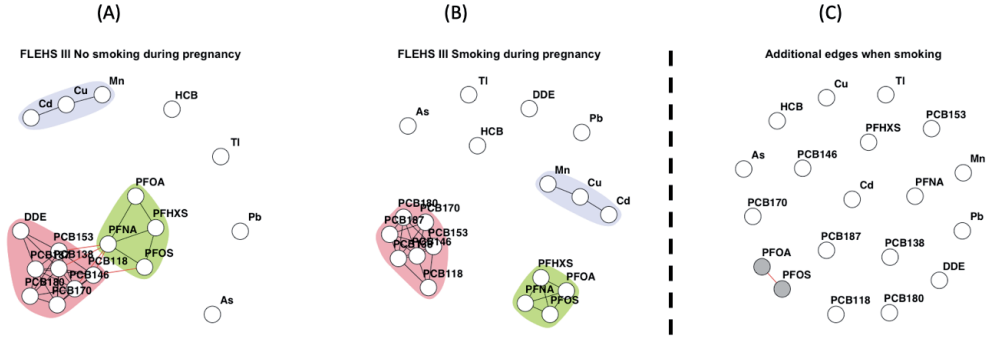


Figure 2. *(A) Network based on individuals FLEHS III where the mother did not smoke during pregnancy (n=248), (B) and mothers who smoked during pregnancy (n=33). Within both (A) and (B) networks, each dot or node represents a biomarker, each edge represents a connection between the biomarkers, the different colors represent a community within a subnetwork. (C) Results of the CNA, dissimilar, or additional, edges when the mother smoked during pregnancy, only nodes part of a subnetwork are colored in gray. Data is corrected for maternal age and maternal pre-pregnancy BMI.*

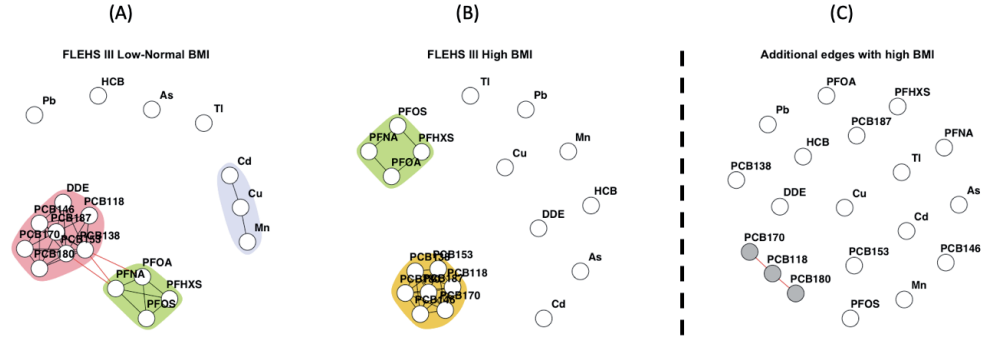


Figure 3. *(A) Network based on individuals from FLEHS III split by low-normal maternal pre-pregnancy BMI (BMI ≤ 25 kg/m², n=195), (B) and high maternal pre-pregnancy BMI (BMI > 25 kg/m², n=86). Within both (A) and (B) networks, each dot or node represents a biomarker, each edge represents a connection between the biomarkers, the different colors represent a community within a subnetwork. (C) Results of the CNA, dissimilar, or additional, edges when the mother had a high BMI, only nodes part of a subnetwork are colored in gray. Data is corrected for maternal age and smoking status during pregnancy.*

Network estimation and community detection

In the obtained network for FLEHS III, three communities were estimated. The markers of HCB, arsenic, thallium and lead were not part of a community. A subnetwork consisted of two connected communities, one with PCBs and *p,p'*-DDE, and one with PFAS (PFOA, PFOS, PFHXS and PFNA). The link between the two communities, marker PFNA within the PFOA community was connected to PCB138 and PCB153. The other community consisted of cadmium, copper and manganese; and was not connected to any other communities. When we compare these networks to weighted networks derived in the same data (Figure S4), we observe the same communities of PCBs, PFAS and the metals cadmium, copper and manganese. Additionally, the metals thallium and lead also form a community. The markers for HCB and arsenic remain not part of any community.

The networks of FLEHS I and II are presented in the Supplementary Material (Figures S1-S2 C). In the network for FLEHS I two subnetworks were estimated, one consisting of cadmium and lead, and the other consisting of PCB138/153/180, HCB and *p,p'*-DDE. In the network for FLEHS II four subnetworks were found, of which two were equal to FLEHS III (PCBs and PFAS). The community of the metals cadmium and lead was equal to FLEHS I. The weighted network for FLEHS I (including *walktrap* community detection algorithm) the community for the metals as the unweighted network (Figure S4). The markers for *p,p'*-DDE and HCB were estimated as a separate community, connected to PCB138/153/180. It can be seen that between the latter two communities the edges were strong. Within the weighted network for FLEHS II the exact same communities as the unweighted network were estimated (Figure S4). Sensitivity analysis was performed by comparing the networks with and without inclusion of the imputed values. No differences between those networks were found.

Comparative Network Analysis

Differential networks (smoking during pregnancy)

Figure 2A and 2B present the networks consisting of biomarkers collected during FLEHS III, stratified by smoking status during pregnancy. 248 mothers did not smoke during pregnancy and 33 mothers did smoke during pregnancy. Equal to the total FLEHS III dataset, two subnetworks were identified for mothers who did not smoke during pregnancy. The graph of non-smoking mothers only differed by the connection of the community PCBs with PFAS, PFOS was also linked with the PCB community (Figure 2A). When the mother did smoke during pregnancy, three subnetworks were distinguished, one consisting of PCBs without *p,p'*-DDE, one of PFASs, and one with cadmium, copper and manganese. (Figure 2B). Compared to figure 2A and the network for the total FLEHS III dataset, the network of mothers who smoked had no connection between PCBs and PFAS. The results from the CNA presented in Figure 2C show one small subnetwork (colored in gray), reflecting the change in connection between PFOS and PFOA, that were not connected when the mother did not smoke during pregnancy, while they were connected when the mother did smoke. The CNA of the edges only present when the mother did not smoke during pregnancy are shown in Figure S5. Here multiple edges between PFNA and PCBs, PFOS with PCB118, and *p,p'*-DDE with multiple PCBs were shown to be only estimated within the network of mothers who did not smoke during pregnancy.

Differential networks (maternal pre-pregnancy BMI)

Figure 3 presents networks consisting of biomarkers collected during FLEHS III, stratified by maternal pre-pregnancy BMI. 195 mothers had a low-normal pre-pregnancy BMI, and 86 mothers a high pre-pregnancy BMI. Within the network of the stratum of mothers with a low-normal pre-pregnancy BMI ($\leq$ 25 kg/m^2), two subnetworks were identified. The detected subnetworks and communities were the same as in the total FLEHS III dataset. The PCB community was connected to PFAS, and the community of cadmium/copper/manganese was not connected to any other (Figure 3A).Within the stratum of mothers with a high pre-pregnancy BMI ($>$ 25 kg/m^2) only communities for PCBs and PFAS were estimated, which were not con-

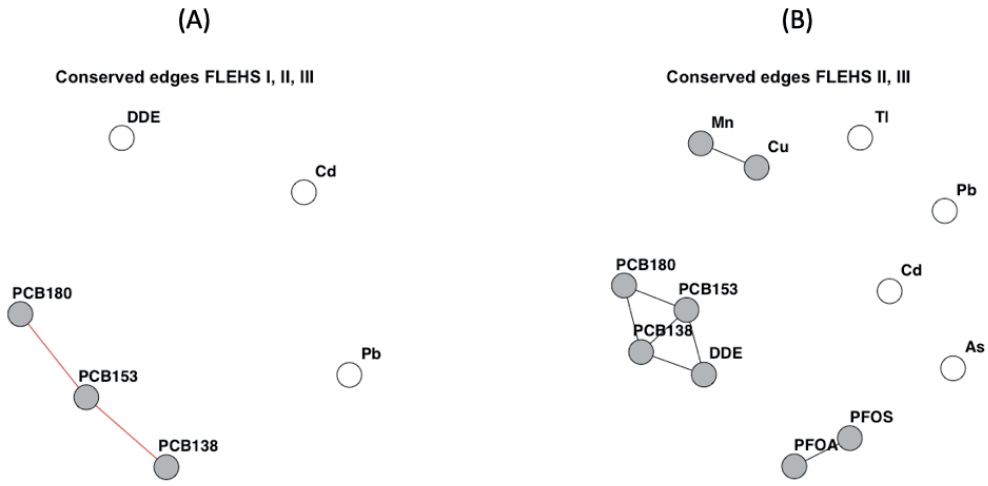


Figure 4. *Results of the CNA across three campaigns (A), or between two campaigns (B). Resulting networks are the similar edges, present in either all three, or both, of the networks per FLEHS campaign. (A) Conserved or similar edges over all three networks of FLEHS I, II and III, based on six biomarkers. (B) Conserved or similar edges between the two networks of FLEHS II and III, based on twelve biomarkers.*

nected (Figure 3B). Also, *p,p'*-DDE was not part of the PCB community. CNA of the networks, presented in Figure 3C, shows the dissimilar edges between the strata. The edges additional to the network for mothers with high pre-pregnancy BMI were identified and colored in gray: PCB118, PCB170 and PCB180. The CNA results showing edges only present for mother with low-normal BMI are shown in Figure S5. Multiple edges between DDE and PCBs were estimated, as well as the edges between manganese, copper and cadmium.

Conserved networks across campaigns

Figure 4A presents the conserved edges across the three networks that were independently derived in the FLEHS I, II, and III datasets (containing the 6 biomarkers measured in all three campaigns). The individual networks derived on the six biomarkers measured in FLEHS I, II, and III are presented in the Supplementary Material (Figure S3). Edges between PCB138, PCB153 and PCB180 were seen in all three campaigns. *p,p'*-DDE, lead and cadmium were not included as a subnetwork of this CNA, as these were not consistently correlated across the three campaigns. Figure 4B presents the conserved edges based on FLEHS II, and III datasets (containing the 12 biomarkers measured in both campaigns). Here, three subnetworks were identified: PFOA and PFOS; *p,p*-'DDE and PCB138/153/180; manganese and copper. These subnetworks identified are the biomarkers that were consistently connected in both sampling campaigns. Arsenic, cadmium, thallium and lead were not included in any of the subnetworks, and therefore not connected to the same biomarkers in both FLEHS II and FLEHS III networks.

DISCUSSION

We provide an application of network analysis in HBM data. The primary utility of this work is to demonstrate that network methodologies can be used to identify prevalent mixtures of chemicals in HBM data. Conditional independence networks provide a data-driven and intuitive approach to highlight the presence of highly connected biomarker measurements without prior assumptions or groupings, about for example sources, chemical properties, pathways or mode of actions. The primary benefit of a network over the heatmap or circos plots is the ease of identification, formalization of the procedure to identify communities and providing a structural approach for comparison of exposure patterns between datasets or across strata within the dataset.

At the same time, some information is potentially lost when describing an HBM dataset using conditional independence networks. Heatmaps and circos plots provide information on the degree of correlation. As such the applied network methodology is an addition to other graphical presentations, not a replacement. The networks as described in the results section are based on unweighted edges, which become of more value in high dimensional HBM data such as untargeted screening data. Weighted partial correlation networks that include information on the degree and direction of association between biomarkers, can provide additional information especially when the number of nodes is not too large and a visual interpretation can be made (Hudson F. Golino & Epskamp, 2017). In addition to graphical tools, approaches such as principal components or cluster analysis (Govarts et al., 2016) can provide insight into complex correlation structures in the data, but are often more difficult to digest visually, especially in high dimensional settings.

Network techniques can be used as a first screening technique to assess patterns in mixtures exposure biomarker data and comparisons across strata of covariates, to assist exposure scientists (pathway, source identification), to assist epidemiologists in taking the communities into account during data analysis and interpretation, and to guide toxicological mixture experiments in identifying real-life mixtures.

Worked example: FLEHS datasets

The application in FLEHS provided some examples of insights that can be acquired by applying networks in HBM data. The community structures we detected in the FLEHS data are in line with earlier findings that groups with similar chemical structures such as PCBs group together (Den Hond et al., 2015). As expected based on previous analyses and literature, due to their often observed high correlation structure, we observed a PCB community in all derived networks, which could be explained by shared sources and similar kinetics (Fisher et al., 2016; Govarts et al., 2016; Lee et al., 2017). We also note, however, that sometimes biomarker *p,p'*-DDE was included in the 'PCB community' highlighting that, when assessing the impact of PCBs, one potentially needs to take into account concurrent exposure of *p,p'*-DDE. This was observed in a previous analyses of the FLEHS data (Govarts et al., 2020), where an association between *p,p'*-DDE and birth weight was observed while correcting for PCBs, which was not observed in a single pollutant model between *p,p'*-DDE and birth weight. Such findings

underline that assessing health risks of combinations of exposure biomarkers reflects better real-world situations and thereby allow more effective risk assessment. Another group of typically highly correlated compounds, the PFAS, were consistently identified as a community in our networks. For the metals the size and composition of the communities varied across the FLEHS campaigns, likely reflecting rather dispersed sources of metal exposure. Within some of the networks, some biomarkers were not included in a subnetwork (such as HCB in FLEHS III), which could be expected since the partial correlation with other biomarkers was very low (no links to other markers in the circos plot), indicating different exposure sources and/or kinetics.

The results of the CNA between the three datasets (Figure 4A), show that the association of the PCBs with *p,p'*-DDE is not always based on the same PCB, and therefore doesn't show as a conserved link across all three campaigns. Multiple explanations can be hypothesized, such as a change in correlation between source and usage over time, causing a change in correlation. Also, the concentration of DDT/DDE/DDD changes over time (e.g. by regulation), as well as the composition of the PCB mixture. The smaller number of samples analyzed in FLEHS II and FLEHS III might also mean that there is a larger impact of random variation or error in the estimated networks, which would explain the observed variation as well.

As an example, the FLEHS data was stratified by smoking status and pre-pregnancy BMI, other strata such as diet (e.g. fish consumption) are also possible. The FLEHS III networks stratified by smoking, both had an equal composition of communities. With the difference that when the mother smoked during pregnancy, an additional edge between PFOA/PFOS was estimated. We could not identify a straightforward explanation for this observation, yet potential explanations would include metabolic changes due to smoking behavior, or a co-exposure that occurs only with smoking women (Rovira et al., 2019). Moreover, only 33 mothers indicated they smoked during pregnancy, which could indicate reduced statistical power to detect true correlations. Also, since this variable indicates if they have ever smoked during pregnancy it could be that the actual smoking frequency was rather low as mothers would be aware of the bad influence of smoking on their unborn child. In the network derived in mothers with high pre-pregnancy BMI we see that the biomarkers form two communities, one with all PCBs and one with PFAS. While both communities were connected when the mother had a low-normal BMI, which could be explained in differences in diet or other lifestyle factors. The results of the CNA between smoking and BMI such as in Figures 2C and 3C give direction in thinking about common exposure sources or common exposures due to lifestyle factors (e.g. dietary habits, low SES, smoking) that contribute to the correlations patterns in HBM exposure biomarkers and will help to prioritize concurrent exposures that could be considered together when assessing exposure-effect associations.

In a biomonitoring study with relatively limited number of markers measured, such as the FLEHS campaigns, weighted networks can be applied as well. In our application a weighted network provided similar insights to our conditional independence method: communities overlap between both methods. The only difference in community was thallium and lead in the FLEHS III dataset, which had the weakest within-community edge. Most likely the detection of this community is just above the threshold; it is dropped in the unweighted network as a result of the slightly different network estimation. In the weighted networks the edges that

connect the communities are clearly less present (thinner) or not present at all. As such there was no significant loss of information by choosing for an unweighted network method. In high dimensional settings the application of weighted networks might become unwieldy and therefore we suggest our method in such settings.

Limitations

There were several limitations to the application of the network analysis in the FLEHS data. First of all, this work is based on a limited set of biomarkers, which reduced the added information of the network estimation, but on the other hand presented easily interpretable networks. Due to the limited number of biomarkers in FLEHS I and II, it was decided to focus the stratification by covariates only on the FLEHS III dataset with nineteen biomarkers. Secondly, the amount of observations was limited. For the comparisons of BMI category or smoking status the amount of observations in one of the strata was limited (minimum of n=33). Thirdly, an underlying assumption of the temporal comparisons between the FLEHS campaigns, is the comparability between the campaigns. Analysis of the biomarkers was done by the same lab in the subsequent campaigns, and control samples were analyzed to assess the comparability of the results. However, different individuals were measured in the different campaigns and slight variations in demographics between participants by campaign could result in different networks.

Future extensions

While not opportune in our current dataset, further extensions to the currently described methods can be foreseen. For example, rather than focusing on differences in networks across covariates, one could focus on differences in communities: Differential Community Detection (Mall et al., 2017). Since the amount of different communities per network was limited for the FLEHS data, this would not have added much information in the FLEHS datasets, but would in high dimensional HBM datasets. Also, the focus on the community differences would be important for applications in epidemiology, mixture toxicology, and mixture risk assessment. The communities in a network can be considered as starting points for further assessment of mixture health effects or in the design of mixture toxicology studies, providing information on combined exposures that occur at population level. Mixture risk assessment might indirectly use the community information, focusing on a common health effect for all substances in the community. Depending on the risk assessment purpose, it might be of use to apply overlapping community detection, where one biomarker could be part of multiple communities (fuzzy clustering) (Xie et al., 2013).

The application of weighted correlation networks to the FLEHS data did not yield substantially differing insights as compared to the results obtained with the application of the conditional independence methods. This is likely explained by the strong communities that exist in this data and that the correlation matrix is largely positive. However, in other datasets of similar dimensions, weighted network approaches can be a useful addition by providing more information (degree and direction) on the associations between biomarkers, underlying the observed communities. When the number of biomarkers in the dataset increase, the inter-

pretation of weighted networks is likely to become more challenging, although community detection will facilitate interpretation to a great deal. Also, CNA of weighted networks will become more challenging, for example inexact graph matching where networks are assessed as equal within certain criteria (Tantardini et al., 2019), or where the most important nodes and/or edges are extracted (Koutra et al., 2016).

The network approaches presented here will be a worthwhile tool when applied in high dimensional HBM datasets. Technological developments are making such datasets increasingly possible by application of methods such as untargeted high resolution mass spectrometry (Andra et al., 2017; Pourchet et al., 2020; Vermeulen et al., 2020). The application of network analysis could help identifying clusters in the data, including parent compounds and related metabolites. Network analysis on high dimensional data has great potential for mixture risk assessment to describe the complex exposure patterns, their composition and variability. CNA on strata of covariates may identify specific risk groups with particular communities of biomarkers of concern. While initial steps have been made towards the risk assessment of mixtures, these approaches are often either based on the assessment of chemically related compounds (e.g. PCB congeners), or based on toxicology (Boberg et al., 2019; Howdeshell et al., 2017; Kienzler et al., 2016), and not on common occurrence and exposure patterns. Insights into complex correlation networks in HBM data, and the presence of communities within these networks, provide useful information on the presence of mixtures at population level.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

IO participated in the design of this research, performed the statistical analysis, and writing the manuscript. **EG** contributed to the data collection, data analysis and writing the manuscript. **IO** and **EG** had an equal contribution to this manuscript. **JV** conceived and designed the research, participated in data analysis and writing the manuscript. **GS, EL** and **RV** contributed to the design of this research, provided feedback on the statistical analyses and assisted in writing the manuscript. All authors reviewed and approved the final version of the manuscript.

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733032 HBM4EU and from National Institute for Public Health and the Environment's Strategic Programme RIVM (SPR) in which expertise and innovative projects prepare RIVM to respond to future issues in health and sustainability. The FLEHS studies were carried out by the Flemish Center of Expertise on Environment and Health. The studies of the Center were commissioned, financed and steered by the Ministry of the Flemish Community.

Abbreviations

CNA	Comparative Network Analysis
EGA	Exploratory Graph Analysis
FLEHS	Flemish Environment and Health Study
HBM	Human Biomonitoring
HCB	Hexachlorobenzene
LOD	Limit of Detection
PCBs	Polychlorinated biphenyls
PFAS	Per- and polyfluoroalkyl substances
PFHXS	Perfluorohexane sulfonate
PFNA	Perfluorononanoic acid
PFOA	Perfluorooctanic acid
PFOS	Perfluorooctane sulfonate
POPs	Persistent Organic Pollutants
p,p'-DDE	p,p'-Dichlorodiphenyldichloroethylene

REFERENCES

Andra, S. S., Austin, C., Patel, D., Dolios, G., Awawda, M., & Arora, M. (2017). Trends in the application of high-resolution mass spectrometry for human biomonitoring: An analytical primer to studying the environmental chemical space of the human exposome. *Environment International*, *100*, 32–61. https://doi.org/10.1016/j.envint.2016.11.026

Baeyens, W., Vrijens, J., Gao, Y., Croes, K., Schoeters, G., Den Hond, E., … Leermakers, M. (2014). Trace metals in blood and urine of newborn/mother pairs, adolescents and adults of the Flemish population (2007–2011). *International Journal of Hygiene and Environmental Health*, *217*(8), 878–890. https://doi.org/10.1016/j.ijheh.2014.06.007

Boberg, J., Dybdahl, M., Petersen, A., Hass, U., Svingen, T., & Vinggaard, A. M. (2019). A pragmatic approach for human risk assessment of chemical mixtures. *Current Opinion in Toxicology*, *15*, 1–7. https://doi.org/10.1016/j.cotox.2018.11.004

Csárdi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *Inter. J. Comp. Syst.*, *1695*, 1–9.

Den Hond, E., Govarts, E., Bruckers, L., & Schoeters, G. (2009). Determinants of polychlorinated aromatic hydrocarbons in serum in three age classes—Methodological implications for human biomonitoring. *Environmental Research*, *109*(4), 495–502. https://doi.org/10.1016/j.envres.2009.02.007

Den Hond, E., Govarts, E., Willems, H., Smolders, R., Casteleyn, L., Kolossa-Gehring, M., … Schoeters, G. (2015). First Steps toward Harmonized Human Biomonitoring in Europe: Demonstration Project to Perform Human Biomonitoring on a European Scale. *Environmental Health Perspectives*, *123*(3), 255–263. https://doi.org/10.1289/ehp.1408616

Emmert-Streib, F., Dehmer, M., & Shi, Y. (2016). Fifty years of graph matching, network alignment and network comparison. *Information Sciences*, *346–347*, 180–197. https://doi.org/10.1016/j.ins.2016.01.074

Fisher, M., Arbuckle, T. E., Liang, C. L., LeBlanc, A., Gaudreau, E., Foster, W. G., … Fraser, W. D. (2016). Concentrations of persistent organic pollutants in maternal and cord blood from the maternal-infant research on environmental chemicals (MIREC) cohort study. *Environmental Health*, *15*(1), 59. https://doi.org/10.1186/s12940-016-0143-y

Flemish Center of Expertise on Environment and Health. (2020). No Title. Retrieved from http://www.milieu-en-gezondheid.be/en/homepage-eng

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, *486*(3–5), 75–174. https://doi.org/10.1016/j.physrep.2009.11.002

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432–441. https://doi.org/10.1093/biostatistics/kxm045

Ganzleben, C., Antignac, J.-P., Barouki, R., Castaño, A., Fiddicke, U., Klánová, J., … Kolossa-Gehring, M. (2017). Human biomonitoring as a tool to support chemicals regulation in the European Union. *International Journal of Hygiene and Environmental Health*, *220*(2), 94–97. https://doi.org/10.1016/j.ijheh.2017.01.007

Gehlenborg, N., O'Donoghue, S. I., Baliga, N. S., Goesmann, A., Hibbs, M. A., Kitano, H., … Gavin, A.-C. (2010). Visualization of omics data for systems biology. *Nature Methods*, *7*(S3), S56–S68. https://doi.org/10.1038/nmeth.1436

Golino, H., & Christensen, A. P. (2020). EGAnet: Exploratory Graph Analysis: A framework for estimating the number of dimensions in multivariate data using network psychometrics. Retrieved from https://cran.r-project.org/package=EGAnet

Golino, H. F., & Demetriou, A. (2017). Estimating the dimensionality of intelligence like data using Exploratory Graph Analysis. *Intelligence*, *62*, 54–70. https://doi.org/10.1016/j.intell.2017.02.007

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLOS ONE*, *12*(6), e0174035. https://doi.org/10.1371/journal.pone.0174035

Govarts, E., Portengen, L., Lambrechts, N., Bruckers, L., Den Hond, E., Covaci, A., … Vermeulen, R. (2020). Early-life exposure to multiple persistent organic pollutants and metals and birth weight: Pooled analysis in four Flemish birth cohorts. *Environment International*, *145*, 106149. https://doi.org/10.1016/j.envint.2020.106149

Govarts, E., Remy, S., Bruckers, L., Den Hond, E., Sioen, I., Nelen, V., … Schoeters, G. (2016). Combined Effects of Prenatal Exposures to Environmental Chemicals on Birth Weight. *International Journal of Environmental Research and Public Health*, *13*(5), 495. https://doi.org/10.3390/ijerph13050495

Green, S., Şerban, M., Scholl, R., Jones, N., Brigandt, I., & Bechtel, W. (2018). Network analyses in systems biology: new strategies for dealing with biological complexity. *Synthese*, *195*(4), 1751–1777. https://doi.org/10.1007/s11229-016-1307-6

Howdeshell, K. L., Hotchkiss, A. K., & Gray, L. E. (2017). Cumulative effects of antiandrogenic chemical mixtures and their relevance to human health risk assessment. *International Journal of Hygiene and Environmental Health*, *220*(2), 179–188. https://doi.org/10.1016/j.ijheh.2016.11.007

Ideker, T., & Krogan, N. J. (2012). Differential network biology. *Molecular Systems Biology*, *8*(1), 565. https://doi.org/10.1038/msb.2011.99

Kienzler, A., Bopp, S. K., van der Linden, S., Berggren, E., & Worth, A. (2016). Regulatory assessment of chemical mixtures: Requirements, current approaches and future perspectives. *Regulatory Toxicology and Pharmacology*, *80*, 321–334. https://doi.org/10.1016/j.yrtph.2016.05.020

Koppen, G., Den Hond, E., Nelen, V., Van De Mieroop, E., Bruckers, L., Bilau, M., … Schoeters, G. (2009). Organochlorine and heavy metals in newborns: Results from the Flemish Environment and Health Survey (FLEHS 2002–2006). *Environment International*, *35*(7), 1015–1022. https://doi.org/10.1016/j.envint.2009.05.002

Koutra, D., Shah, N., Vogelstein, J. T., Gallagher, B., & Faloutsos, C. (2016). DeltaCon: Principled Massive-Graph Similarity Function with Attribution. *ACM Transactions on Knowledge Discovery from Data*, *10*(3), 1–43. https://doi.org/10.1145/2824443

Lee, W.-C., Fisher, M., Davis, K., Arbuckle, T. E., & Sinha, S. K. (2017). Identification of chemical mixtures to which Canadian pregnant women are exposed: The MIREC Study. *Environment International*, *99*, 321–330. https://doi.org/10.1016/j.envint.2016.12.015

Liu, H., Roeder, K., & Wasserman, L. (2010). Stability Approach to Regularization Selection (StARS) for High Dimensional Graphical Models. Retrieved from http://arxiv.org/abs/1006.3316

Lubin, J. H., Colt, J. S., Camann, D., Davis, S., Cerhan, J. R., Severson, R. K., … Hartge, P. (2004). Epidemiologic Evaluation of Measurement Data in the Presence of Detection Limits. *Environmental Health Perspectives*, *112*(17), 1691–1696. https://doi.org/10.1289/ehp.7199

Mall, R., Ullah, E., Kunjia, K., & Bensmail, H. (2017). Differential Community Detection in Paired Biological Networks. *BioRxiv Preprint*, (Preprint). https://doi.org/https://doi.org/10.1101/147538

Mitra, K., Carvunis, A.-R., Ramesh, S. K., & Ideker, T. (2013). Integrative approaches for finding modular structure in biological networks. *Nature Reviews Genetics*, *14*(10), 719–732. https://doi.org/10.1038/nrg3552

O'Brien, K. M., Upson, K., & Buckley, J. P. (2017). Lipid and Creatinine Adjustment to Evaluate Health Effects of Environmental Exposures. *Current Environmental Health Reports*, *4*(1), 44–50. https://doi.org/10.1007/s40572-017-0122-7

Orman, G. K., & Labatut, V. (2009). A Comparison of Community Detection Algorithms on Artificial Networks (pp. 242–256). https://doi.org/10.1007/978-3-642-04747-3_20

Pons, P., & Latapy, M. (2005). Computing Communities in Large Networks Using Random Walks (pp. 284–293). https://doi.org/10.1007/11569596_31

Pourchet, M., Debrauwer, L., Klanova, J., Price, E. J., Covaci, A., Caballero-Casero, N., … Antignac, J.-P. (2020). Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: From promises to challenges and harmonisation issues. *Environment International*, *139*, 105545. https://doi.org/10.1016/j.envint.2020.105545

Robinson, O., Basagaña, X., Agier, L., de Castro, M., Hernandez-Ferrer, C., Gonzalez, J. R., … Vrijheid, M. (2015). The Pregnancy Exposome: Multiple Environmental Exposures in the INMA-Sabadell Birth Cohort. *Environmental Science & Technology*, *49*(17), 10632–10641. https://doi.org/10.1021/acs.est.5b01782

Rovira, J., Martínez, M. Á., Sharma, R. P., Espuis, T., Nadal, M., Kumar, V., … Schuhmacher, M. (2019). Prenatal exposure to PFOS and PFOA in a pregnant women cohort of Catalonia, Spain. *Environmental Research*, *175*, 384–392. https://doi.org/10.1016/j.envres.2019.05.040

Schoeters, G., Colles, A., Hond, E. Den, Croes, K., Vrijens, J., Baeyens, W., … Loots, I. (2012). The Flemish Environment and Health Study (FLEHS) – Second Survey (2007–2011): Establishing Reference Values for Biomarkers of Exposure in the Flemish Population. In *Issues in Toxicology* (pp. 135–165). https://doi.org/10.1039/9781849733373-00135

Schoeters, G., Govarts, E., Bruckers, L., Den Hond, E., Nelen, V., De Henauw, S., … Baeyens, W. (2017). Three cycles of human biomonitoring in Flanders – Time trends observed in the Flemish Environment and Health Study. *International Journal of Hygiene and Environmental Health*, *220*(2), 36–45. https://doi.org/10.1016/j.ijheh.2016.11.006

Schoeters, G., Hond, E. Den, Colles, A., Loots, I., Morrens, B., Keune, H., … Baeyens, W. (2012). Concept of the Flemish human biomonitoring programme. *International Journal of Hygiene and Environmental Health*, *215*(2), 102–108. https://doi.org/10.1016/j.ijheh.2011.11.006

Tamayo-Uria, I., Maitre, L., Thomsen, C., Nieuwenhuijsen, M. J., Chatzi, L., Siroux, V., … Basagaña, X. (2019). The early-life exposome: Description and patterns in six European countries. *Environment International*, *123*, 189–200. https://doi.org/10.1016/j.envint.2018.11.067

Tantardini, M., Ieva, F., Tajoli, L., & Piccardi, C. (2019). Comparing methods for comparing networks. *Scientific Reports*, *9*(1), 17557. https://doi.org/10.1038/s41598-019-53708-y

Vermeulen, R., Schymanski, E. L., Barabási, A.-L., & Miller, G. W. (2020). The exposome and health: Where chemistry meets biology. *Science*, *367*(6476), 392–396. https://doi.org/10.1126/science.aay3164

Villaveces, J., Koti, P., & Habermann, B. (2015). Tools for visualization and analysis of molecular networks, pathways, and -omics data. *Advances and Applications in Bioinformatics and Chemistry*, 11. https://doi.org/10.2147/AABC.S63534

Xie, J., Kelley, S., & Szymanski, B. K. (2013). Overlapping community detection in networks. *ACM Computing Surveys*, *45*(4), 1–35. https://doi.org/10.1145/2501654.2501657

Yu, D., Zhang, Z., Glass, K., Su, J., DeMeo, D. L., Tantisira, K., … Qiu, W. (2019). New Statistical Methods for Constructing Robust Differential Correlation Networks to characterize the interactions among microRNAs. *Scientific Reports*, *9*(1), 3499. https://doi.org/10.1038/s41598-019-40167-8

Zhang, B., Li, H., Riggins, R. B., Zhan, M., Xuan, J., Zhang, Z., … Wang, Y. (2009). Differential dependency network analysis to identify condition-specific topological changes in biological networks. *Bioinformatics*, *25*(4), 526–532. https://doi.org/10.1093/bioinformatics/btn660

Zhao, T., Liu, H., & Roeder, K. (2012). The huge Package for High-dimensional Undirected Graph Estimation in R. *The Journal of Machine Learning Research*, *13*, 1059–1062.

SUPPLEMENTARY MATERIAL

Table S1. *Characteristics of the FLEHS newborn campaigns: mean (min-max). P-value was calculated by ANOVA for continuous variables, and by Chi-square for categorical outcomes.*

Mothers	FLEHS I	FLEHS II	FLEHS III	p-value
Number of participants	1196	255	281	
Age (years)	29.6 (18.1 – 44.0)	30.3 (18.2 – 42.4)	30.2 (18.9 – 44.8)	0.02
Mean pre-pregnancy BMI (kg/m^2)	23.3 (14.0 – 44.6)	23.5 (16.0 – 47.4)	23.9 (15.2 – 45.2)	0.13
Mean duration of pregnancy (weeks)	39 (31-42)	39 (34-42)	39 (35-42)	0.52
Primipari (%)	60.8	39.8	44.8	<0.0001
Smoking during pregnancy (%)	16.2	11.6	11.7	0.06
Never drinking alcohol during pregnancy (%)	91.7	55.5	74.7	<0.0001

Table S2. *Geometric mean exposure concentrations of the biomarkers, based on the data after imputation and fat-correction, split by measurement campaign. As well as the percentage of missing values per biomarker, the % missing are missing at random, e.g. due to laboratory sample loss or insufficient blood volume.*

Variable (unit) / Campaign	**FLEHS I**		**FLEHS II**		**FLEHS III**	
	Geomean (Min - Max)	**% missing**	**Geomean (Min-Max)**	**% missing**	**Geomean (Min-Max)**	**% missing**
Arsenic (As) (µg/L)			0.55 (0.03;14.40)	2.35 %	0.71 (0.09;19.00)	0 %
Cadmium (Cd) (µg/L)	0.19 (0.00;13.87)	7.02 %	0.08 (0.01;5.31)	2.35 %	0.02 (0.01;0.10)	0 %
Copper (Cu) (µg/L)			598.46 (299.0;994.0)	2.35 %	556.79 (310.0;995.0)	0 %
Manganese (Mn) (µg/L)			31.46 (7.29;80.12)	2.35 %	30.10 (12.46;81.05)	0 %
Lead (Pb) (µg/L)	13.29 (0.63;177.61)	7.36 %	8.60 (2.40;67.40)	2.35 %	6.44 (1.93;43.92)	0 %
Thallium (Tl) (ng/L)			16.89 (8.00;41.00)	2.35 %	18.59 (8.73;43.86)	0 %
p.p'-DDE (ng/g lipid)	108.16 (6.64;1815.53)	6.86 %	77.04 (9.76;641.08)	0.78 %	59.92 (8.85;903.18)	1.78 %
HCB (ng/g lipid)	17.63 (0.94;402.07)	12.46 %			11.61 (1.26;72.07)	1.78 %
PCB118 (ng/g lipid)					3.36 (0.33;15.26)	1.78 %
PCB138 (ng/g lipid)	14.68 (0.56;156.86)	11.62 %	17.20 (1.68;69.78)	0.78 %	10.31 (1.83;39.68)	1.78 %
PCB146 (ng/g lipid)					1.49 (0.24;6.88)	1.78 %
PCB153 (ng/g lipid)	25.69 (0.93;230.21)	10.70 %	26.77 (4.78;108.89)	0.78 %	16.50 (2.91;53.97)	1.78 %
PCB170 (ng/g lipid)					4.34 (0.57;20.18)	1.78 %
PCB180 (ng/g lipid)	20.41 (1.51;153.13)	10.20 %	15.64 (2.68;70.18)	0.78 %	8.57 (0.99;56.42)	1.78 %
PCB187 (ng/g lipid)					2.39 (0.40;21.10)	1.78 %
PFHXS (µg/L)					0.36 (0.06;1.33)	4.27 %
PFNA (µg/L)					0.21 (0.05;1.39)	4.27 %
PFOA (µg/L)			1.53 (0.50;4.30)	13.73 %	1.19 (0.26;5.87)	4.27 %
PFOS (µg/L)			2.69 (0.80;17.30)	13.73 %	1.11 (0.13;8.37)	4.27 %

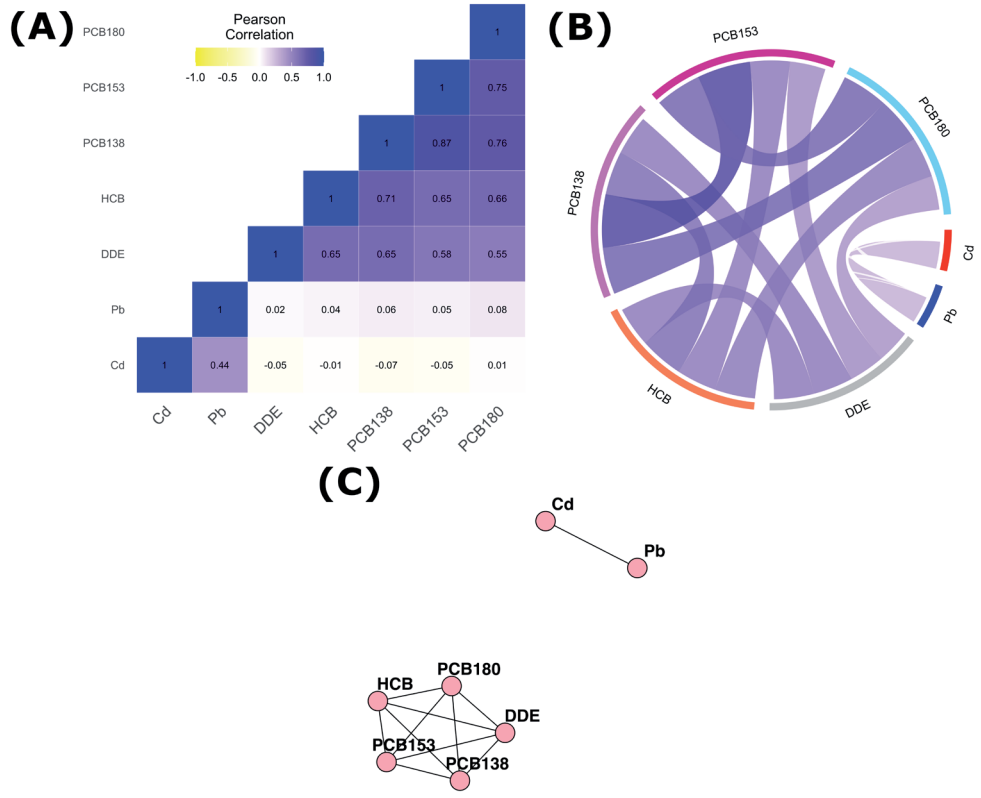


Figure S1. *Heatmap (A), circular correlation globe (B) and network (C) of FLEHS I, seven biomarkers, n=1196. Data is corrected for maternal age, smoking during pregnancy and maternal pre-pregnancy BMI. Within the circular globe each biomarker is presented as a color-block on the circular axis. Within the network, each dot or node represents a biomarker, each edge represents a connection between the biomarkers.*

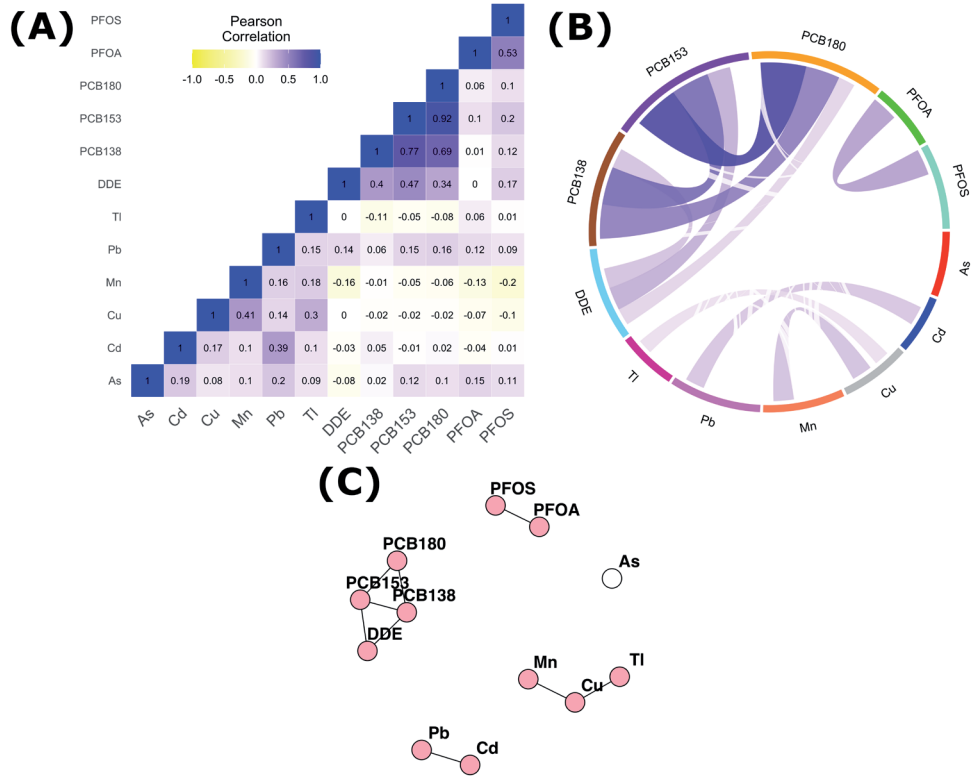


Figure S2. *Heatmap (A), circular correlation globe (B) and network (C) of FLEHS II, twelve biomarkers, n=255. Data is corrected for maternal age, smoking during pregnancy and maternal pre-pregnancy BMI. Within the circular globe each biomarker is presented as a color-block on the circular axis. Within the network, each dot or node represents a biomarker, each edge represents a connection between the biomarkers.*

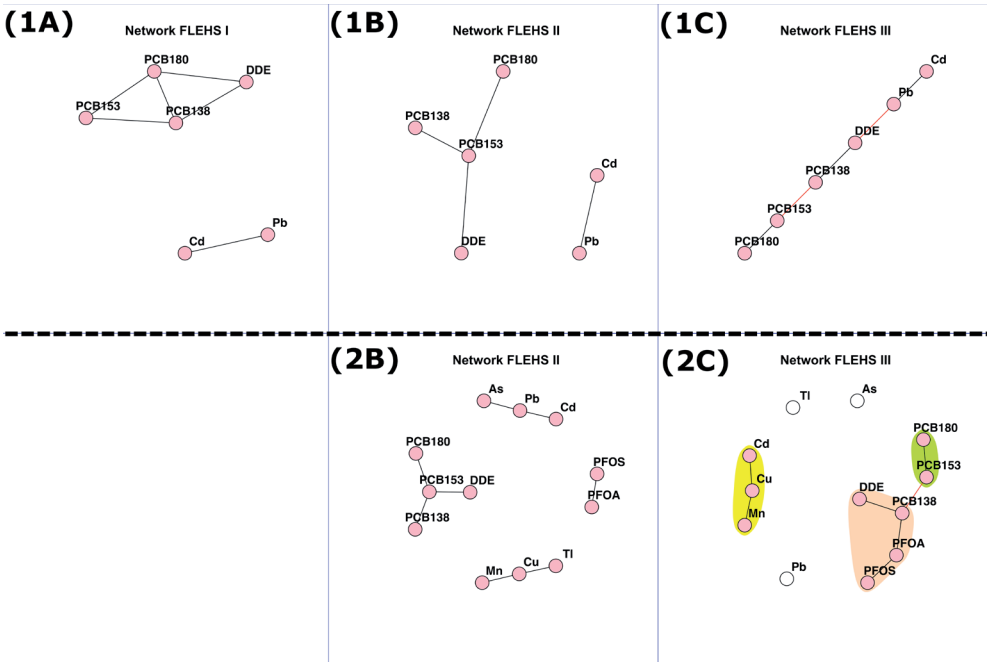


Figure S3. *Networks for each FLEHS sampling campaign, including community detection, with networks consisting of six corresponding biomarkers of (from left to right) FLEHS I, II and III (upper panel, 1A, 1B, 1C) and networks of twelve corresponding biomarkers of FLEHS II and III (lower panel, 2B, 2C). Each dot or node represents a biomarker, each edge represents a connection between the biomarkers. FLEHS I includes 1196 samples, FLEHS II 255 samples and FLEHS III 281 samples. The biomarker data is corrected for maternal age, smoking during pregnancy and BMI. In the upper panel (1) it can be seen that for the FLEHS I and II datasets two (not connected) subnetworks were estimated, one consisting of p,p'-DDE and PCB138/153/180, and the other of cadmium and lead. In FLEHS III a single network is estimated. In the lower panel (2) it can be seen that in the FLEHS II data four subnetworks were detected, with thallium, copper and manganese in one and cadmium, lead and arsenic in another subnetwork. In the FLEHS III data two subnetworks were estimated. One of the subnetworks contains two communities: PCB180 and PCB153 versus p,p'-DDE, PFOS, PCB138, and PFOA. The other subnetwork contained metals cadmium, lead, copper, thallium, and manganese. Arsenic was not included in either subnetwork.*

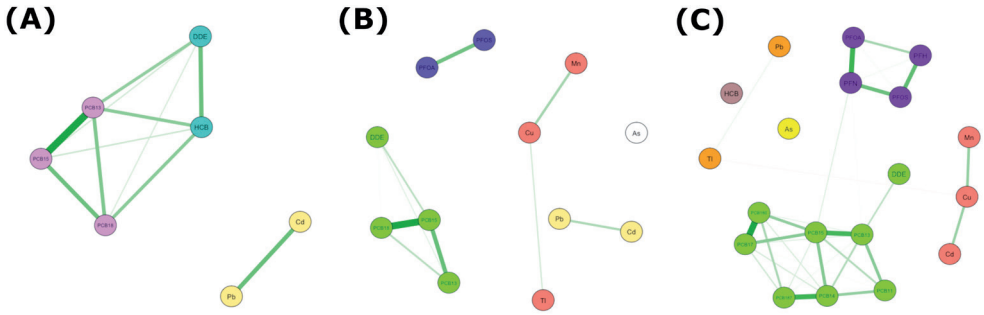


Figure S4. *Weighted networks FLEHS I (A), II (B) and III (C). A parametric bootstrap (1,000 iterations) was used to estimate the median network structure. Communities in the EGA network were estimated using the walktrap algorithm. Nodes are colored according to the community they belong to.*

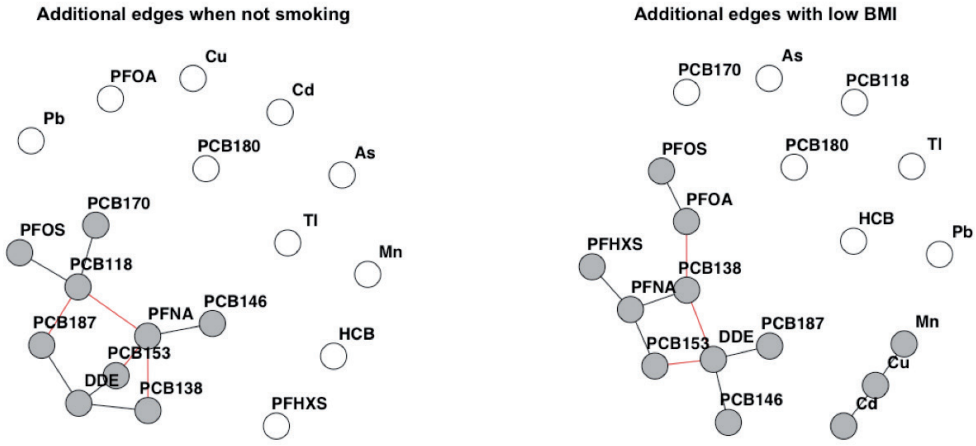


Figure S5. *(A) Results of the CNA, dissimilar, or additional, edges when the mother did not smoke during pregnancy, only nodes part of a subnetwork are colored in gray. Data is corrected for maternal age and BMI during pregnancy. (B) Results of the CNA, dissimilar, or additional, edges when the mother had a low BMI, only nodes part of a subnetwork are colored in gray. Data is corrected for maternal age and smoking status during pregnancy.*

Chapter 3

Identification of Real-Life Mixtures Using Human Biomonitoring Data: A Proof of Concept Study

Laura Rodriguez Martin [1], **Ilse Ottenbros [2,3]**, Nina Vogel [4], Marike Kolossa-Gehring [4], Phillipp Schmidt [4], Katarína Řiháčková [5], Miguel Juliá Molina [6], Elena Varea Jiménez [6], Eva Govarts [1], Susana Pedraza [6], Erik Lebret [2,3], Jelle Vlaanderen [2] and Mirjam Luijten [7]

[1] Health, Flemish Institute for Technological Research (VITO), 2400 Mol, Belgium
[2] Institute for Risk Assessment Sciences (IRAS), Utrecht University, 3584 CM Utrecht, The Netherlands
[3] Center for Sustainability, Environment and Health, National Institute for Public Health and the Environment (RIVM), 3720 BA Bilthoven, The Netherlands
[4] German Environment Agency (UBA), 14195 Berlin, Germany
[5] RECETOX, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic
[6] National Centre for Environmental Health, Instituto de Salud Carlos III, 28220 Majadahonda, Spain
[7] Centre for Health Protection, National Institute for Public Health and the Environment (RIVM), 3720 BA Bilthoven, The Netherlands

Published in Toxics, February 2023
DOI: 10.3390/toxics11030204

ABSTRACT

Human health risk assessment of chemical mixtures is complex due to the almost infinite number of possible combinations of chemicals to which people are exposed to on a daily basis. Human biomonitoring (HBM) approaches can provide inter alia information on the chemicals that are in our body at one point in time. Network analysis applied to such data may provide insight into real-life mixtures by visualizing chemical exposure patterns. The identification of groups of more densely correlated biomarkers, so-called "communities", within these networks highlights which combination of substances should be considered in terms of real-life mixtures to which a population is exposed. We applied network analyses to HBM datasets from Belgium, Czech Republic, Germany, and Spain, with the aim to explore its added value for exposure and risk assessment. The datasets varied in study population, study design, and chemicals analyzed. Sensitivity analysis was performed to address the influence of different approaches to standardize for creatinine content of urine. Our approach demonstrates that network analysis applied to HBM data of highly varying origin provides useful information with regards to the existence of groups of biomarkers that are densely correlated. This information is relevant for regulatory risk assessment, as well as for the design of relevant mixture exposure experiments.

INTRODUCTION

Humans are exposed to a myriad of concurrent and protracted environmental, occupational, dietary, lifestyle, and consumer product exposures. Due to the (increasingly) 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 (Drakvik et al., 2020; EFSA et al., 2021). 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 prioritization of chemical mixtures in policy and research.

There is no broadly accepted operational definition of mixtures. The European Commission communication on "The combination effects of chemicals—Chemical mixtures" (European Commission, 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: (a) intentional mixtures, i.e., manufactured formulated products that are marketed as such; (b) mixtures originating from a single source, also known as 'unintentional mixtures'; and (c) mixtures of chemicals originating from multiple sources and through multiple pathways, also known as 'coincidental mixtures' (Kienzler et al., 2016; Rotter et al., 2018).

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. 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 causal assessment of risks, and to regulation of substances and general risk management policies (Agier et al., 2016; Barrera-Gómez et al., 2017). In this manuscript, the term 'mixture' is used to describe any combination of exposure to chemical substances or of exposure biomarkers that have been measured in one or more biological matrices of a person during a single time point. These biomarkers include both the chemical substances themselves and/or their metabolites.

In the context of the European Joint Programme HBM4EU (hbm4eu.eu) on human biomonitoring (HBM), we evaluated existing HBM data using correlation network analysis to identify real-life exposure patterns to mixtures in the human body. Network analysis is a graphical method to visualize correlations between variables in a dataset. The method allows for the identification of groups of exposure biomarkers that are more densely related amongst each other than with other biomarkers. These groups are referred to as "communities". Building on a successful network analysis exploration based on Flemish data (Ottenbros et al., 2021), we further developed and applied network analysis to HBM datasets from Belgium, the Czech Republic, Germany, and Spain. The objective was to describe the distribution of (patterns in)

biomarkers of exposure and to identify possible determinants that explain observed variation of patterns in biomarkers of exposure. For each of the four studies, results of the network analyses are shown, and findings are discussed.

MATERIAL AND METHODS

In this section, we first describe data selection and preparation steps, followed by the characteristics of the four datasets, the statistical descriptives and network analyses.

Selection of Existing HBM Studies

With the aim to further explore the added value of network analysis, four HBM studies participating in the HBM4EU project were selected. The selection of the studies was based on data availability, as well as on availability of appropriate statistical expertise at the respective institutes.

Data Selection and Preparation

Harmonized data selection and preparation steps were performed with the subsequent network analyses 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 are described in more detail in Ottenbros et al., 2021. In brief, these involve (a) checking the distribution of the variables; (b) transforming the data if needed; (c) imputing the data points below the LOD (limit of detection) or LOQ (limit of quantification); (d) correcting for outliers; (e) standardizing around zero; and (f) scaling of the data.

Concentrations of biomarkers were natural log transformed because HBM distributions are typically skewed. The network analysis makes use of the partial correlation structure. Therefore, a strategy for dealing with censored and missing data is required. Thus, an (arbitrary) cut-off at a maximum of 40% of HBM levels below LOD/LOQ was applied. Substances with more than 40% of the measured HBM values below LOD/LOQ were excluded from further analysis. For the included substances, 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 (version 3.15; (Buuren et al., 2011)) in R (v3.5.0 or higher). Please refer to the description of the individual studies for details on which determinants this strategy was applied to. Determinants (e.g., age, sex, and smoking) 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 those biomarkers that were completely missing, using linear regression.

For several substances, notably metals, different species were measured. For example, for arsenic, data for total arsenic, organic, and inorganic arsenic were available. Additionally,

in some studies, the same substance was measured in urine as well as in blood, e.g., lead or cadmium. This would lead to relatively high correlations between the different biomarkers for the same substance. In terms of combined exposures to chemical substances, such correlations do not provide relevant information. Furthermore, it may also affect the partial correlations structure with other substances. Therefore, only a single biomarker was selected for inclusion in the network analysis; where possible, the biomarker that best reflects the long-term exposure of the individual was selected. 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 standardization for creatinine content 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 lipophilic substances measured in blood, blood lipid levels were used to standardize measured blood levels. A sensitivity analysis was performed on the German data (see Appendix A), showing the results standardized for creatinine or not.

Characteristics of the Four Existing HBM Datasets

3xG (Belgium)

The 3xG study (Health—Municipalities—Birth, translated from Gezondheid, Gemeenten, Geboorten) is a birth cohort study that monitors and promotes health of the inhabitants of three bordering rural communities (Dessel, Mol, and Retie) in Flanders, Belgium. This study focuses on the effect of the environment and lifestyle on health. This is performed by researching 301 growing children from the region and by processing the disease and mortality registers of the 3 municipalities. 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 the 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. 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, such as the educational level of the household members, smoking habits, information on consumption of local food, and the course of pregnancy, were collected. A urine sample was collected in the second trimester of pregnancy. Birth weight, length, and head circumference of the baby at birth were collected with consent from the mothers. A blood sample of the mother and umbilical cord blood were collected at delivery and a questionnaire was filled in by the mothers at the same time point. Since not all biomarkers were measured in the same group of participants, we selected the biomarkers that ensure a subset with enough participants. Consequently, a subset of 125 mother–child pairs were included in the network analysis. Biomarkers included in the network analysis were corrected for age (in years), body mass index (BMI), and/or smoking status of the participant. Networks

were stratified by education status; low ISCED (International Standard Classification of Education) is defined by participants belonging to educational levels 0–4, and high ISCED is defined by participants belonging to educational level ≥5.

CELSPAC—FIREexpo (Czech Republic)

The CELSPAC—FIREexpo study, conducted in the Czech Republic, aimed to determine the health risks resulting from the occupational exposure of Czech firefighters and to implement measures to minimize such risks. All participants were males between the age of 18 and 35 years and non-smokers. All participants expressed and signed their informed consent before their participation in the study. The sampling campaign took place from January 2019 to June 2020. Samples of venous blood and morning urine were collected and analyzed for the presence of biomarkers. More information is publicly available on the study website (https://www.recetox.muni.cz/hear/projects/celspac-fireexpo (accessed on 27 January 2023)) and in Řiháčková et al., 2023 Because of the case-control study design, the analysis of the detection frequency and the imputation of values <LOQ was carried out separately for the two population groups (firefighters and controls); therefore, the list of biomarkers used for the network analysis slightly differed between the two groups. The biomarker levels were corrected for age (in years) and BMI. Stratification for sex and smoking status was not relevant for this study (all participants were male and non-smokers), and data on education level were not collected.

GerES V (Germany)

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 and their sources. 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, including HBM samples with more than 80 biomarkers. The study was performed in a stratified randomly selected sample design. In GerES V, a subsample (n = 2294) of the 3- to 17-year-old participants of the German Health Interview and Examination Survey for Children and Adolescents (KiGGS Wave 2) by the Robert Koch Institute (RKI; Berlin, Germany) was examined (Mauz et al., 2017; Schulz et al., 2017). Participants of GerES V from 167 different sampling locations in Germany were visited by a trained interviewer, conducting an interview on exposure-relevant behavior and collecting information on the living environment with the participants and their parents or legal guardians, and collecting inter alia samples of first-morning void urine and blood. For more details on both studies, see Murawski et al., 2020 and Hoffmann et al., 2018.

Different biomarkers were measured in subsets of participants in the nationally representative GerES V. To have the maximum number of chemical substances while avoiding high proportions of missing data, for the current analyses, data from urinary biomarkers were used that were available for a subgroup of GerES V participants (n = 515, aged from 3 to 17 years old). This resulted in a set of 51 different chemicals.

Biomarkers included in the networks were corrected for the determinant's age (in years), sex, BMI, smoking status of the participant creatinine, and education of the household (ISCED).

Networks were stratified by ISCED, median age, and BMI (each only correcting for the remaining determinants). A sensitivity analysis, using different dilution adjustments of creatinine, was conducted.

BIOAMBIENT.ES (Spain)

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 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 older, who were residents in Spain for at least 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). 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. For the purpose of the network analysis, the dataset with the highest number of substances was selected, although this reduced the number of participants, since not everyone had all substances determined.

Statistical Analysis

Descriptive Analysis

The descriptive analysis of the data used for network analysis largely follows the conventions developed in HBM4EU's Work Package on data management and analysis (HBM4EU D10.12; www.hbm4eu.eu/work-packages/deliverable-10–12-update-statistical-analysis-plan-for-the-co-funded-studies-of-wp8/ (accessed on 27 January 2023)). 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 descriptive statistics.

Descriptive statistics were calculated using R (v3.5.0 or higher). The 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 standardized for creatinine or blood lipids (biomarker measured in urine in the case of creatinine or measured in blood in the case of lipid standardization for lipophilic biomarkers). Pearson correlation structures in the datasets were computed and displayed using heatmaps.

Network Analysis

Network analyses were performed as previously described (Ottenbros et al., 2021). After the data selection and preparation steps, partners performed the network analysis using uniform centrally prepared scripts. Network analysis was 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. The output network presents unweighted edges, only providing information on whether the edge connecting nodes is present or absent, depending on a cut-off value (lambda).

For comparison purposes, weighted network analysis, which is more computationally demanding, was applied as well, making use of the package *EGAnet* (v1.2.3 (Golino et al., 2020)) (Christensen et al., 2021; Golino et al., 2021). The output weighted network shows the strength of the edge by thickness of the line and direction of the correlation by color of the line (green for a positive correlation and red for a negative correlation). Both the unweighted and weighted networks were estimated using the graphical lasso (GLASSO), which involves penalized maximum likelihood estimation (Friedman et al., 2008). This method is a simple and fast algorithm for estimation of a sparse inverse covariance matrix using a lambda penalty. The GLASSO cycles through the variables, fitting a modified lasso regression to each variable in turn. Regularization 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 maximum lambda.

For the unweighted networks, the optimal lambda selection was conducted using the stability approach to regularization 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.

For the weighted networks, the optimal graph from the GLASSO was selected with the EBIC tuning parameter (default of 0.5). A parametric bootstrap (1000 iterations) was used to estimate the median network structure, which was then plotted as the final result.

On both the weighted and the unweighted networks, the walktrap clustering algorithm from the *igraph* package was used, which performs random walks (using a default of 4 steps) across the network to merge nodes to so-called communities in a bottom-up manner (Orman et al., 2009; Pons et al., 2005). Nodes were colored according to the community they were assigned to. Edges of the unweighted networks linking different communities were colored in red, and edges within a community were colored in black. Biomarkers within the same community were more closely related to one another than to the other measured biomarkers in the network. To the degree possible, usage of colors is standardized within each dataset, but not across datasets, nor between unweighted and weighted network graphs.

RESULTS

Descriptive statistics for the Chemical substances Included in the Network Analysis

Table 1 shows an overview of the descriptive statistics for the HBM datasets for those chemicals that were measured in more than one country, i.e., the biomarkers for the substances included for the network analysis, the matrix in which the biomarkers were measured, their proportions below LOD or LOQ, and percentiles and geometric mean of the biomarker concentrations. Please note that the concentrations for urinary biomarkers were standardized for creatinine. Country-specific descriptive statistics of biomarker levels as used in the network analyses are presented in Supplementary Tables S1–S4. The correlation structure between biomarkers is graphically represented in the subsequent sections by heatmaps.

3xG (Belgium)

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), thallium (Tl), and lead (Pb), total arsenic (As), hydroxy pyrene (1-PYR), trans-muconic acid (TTMA), phthalates including mono(2-ethylhexyl) phthalate (MEHP), mono(2-ethyl-5-hydroxy- hexyl) phthalate (5OH-MEHP), mono(2-ethyl-5-oxo-hexyl) phthalate (5oxo-MEHP), mono-n-butyl phthalate (MnBP), mono-benzyl phthalate (MBzP), mono-ethyl phthalate (MEP), and mono-isobutyl phthalate (MiBP), and bisphenol A (BPA total) 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, thallium, managenese, and lead) and arsenic were available in cord blood (CB) samples of the newborn; and organochlorine compounds (OCs) including polychlorinated biphenyl 138 (PCB128), polychlorinated biphenyl 153 (PCB153), polychlorinated biphenyl 180 (PCB180), dichlorodiphenyldichloroethylene (p,p'-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 (CBP) samples of the newborn.

The descriptive statistics for the biomarkers included in the network analysis are given in Supplementary Table S1. Nickel measured in the cord blood of the newborn and HCB and PFHxS measured in the cord blood plasma of the newborn show the highest percentage of values below LOD/LOQ, being 37%, 24.8%, and 20.8%, respectively. Furthermore, for nickel measured in the cord blood of the newborns, the P25 value is under the LOD (Supplementary Table S1).

Figure 1 shows the correlation between the biomarkers for the abovementioned substances. Biomarkers belonging to the same chemical groups show higher correlations, such as PCBs, phthalates, PFASs, and heavy metals. Interestingly, specific heavy metals measured in the urine of the mother during pregnancy and the cord blood of the newborn during birth show low correlations. For example, arsenic and lead show a Pearson correlation of 0.34 and 0.35, respectively, while other heavy metals do not show any significant correlation. PFASs and PCBs also show small positive correlations. Significant negative correlations were not observed for any of the biomarkers.

Table 1. *Descriptive statistics for biomarkers included in the network analysis measured in more than one study. PFAS were measured in blood(in µg/L), all others in urine (in µg/g creatinine).*

Substance Group	Biomarker	3XG (Belgium)					CELSPAC—FIREexpo; Controls (Czech Republic)					GerES V (Germany)					BIOAMBIENT.ES (Spain)				
Distribution		% < LOQ	P25	P50	P75	P95	% < LOQ	P25	P50	P75	P95	% < LOQ	P25	P50	P75	P95	% < LOQ	P25	P50	P75	P95
Elements	Cd	0%	0.21	0.28	0.37	0.54						26%	< LOQ	0.06	0.09	0.15	2.5%	0.12	0.2	0.38	0.72
	Cr	1.6%	0.26	0.49	0.82	1.76						7.8%	0.26	0.34	0.49	0.77					
	Hg											5.1%	0.04	0.06	0.1	0.26	0.68%	0.56	0.99	1.58	2.75
	Sb	18%	0.03	0.04	0.06	0.15						20%	0.03	0.05	0.07	0.13					
	As	0%	6.72	13.89	38.79	81.22						0%	4.4	6.9	14.0	55.21					
	Pb	0%	0.64	0.84	1.14	1.7											2.6%	0.43	0.7	1.04	2.36
	Tl	0%	0.18	0.22	0.26	0.35											11.35%	0.08	0.11	0.16	0.26
Phthalate substitute	OH-DINCH											0.19%	0.98	2.13	4.66	14.7	4.91%	0.29	0.7	6.81	19.82
	oxo-DINCH											1.55%	0.39	0.93	2.03	7.16	13.5%	0.11	0.4	1.19	11.87
	cx-MINCH											0.19%	0.49	1.02	2.11	7.8	3.68%	0.26	0.43	1.22	8.21
Phthalates	MEHP	2.4%	1.75	2.53	4.33	8.42						13.0%	0.71	1.22	2.04	4.19	3.68%	2.47	4.09	6.63	14.9
	SOH-MEHP	0%	6.67	10.08	13.75	38.04						0%	5.87	8.98	13.94	28.8	0%	11.54	18.44	26.26	56.6
	Soxo-MEHP	0%	4.39	7.18	9.69	22.98						0%	4.08	6.42	10.49	21.6	0.61%	7.74	11.45	17.03	35.71
	Scx-MEPP											0%	6.1	9.92	16.9	35.8	0%	12.82	18.88	28.15	54.27
	MBzP	0%	3.79	7.1	11.97	22.59						0.39%	1.45	2.38	4.75	17.5	1.23%	3.16	5.09	8.98	28.53
	MnBP	0%	23.36	34	51.35	91.75						0%	12.04	18.18	28.67	54.8	0.61%	9.63	14.74	22.23	41.53
	OH-MnBP											0.78%	1.25	2.12	3.49	7.27	2.45%	1.08	1.71	2.44	5.04
	MiBP	0%	43.16	60	94.4	288.5						0%	13.54	21.36	33.58	87.2	0%	16.33	23.71	34.19	72.73
	OH-MiBP											0%	4.7	7.52	12.17	30.2	0%	6.5	9.17	14.19	25.01
	MEP	0%	17.76	40.38	82.98	203.57						0%	10.96	17.76	32.05	113	0%	87.08	189.47	345.21	1307.09
	OH-MiNP											0%	3.35	5.27	8.73	24.6	1.84%	2.03	3.45	5.91	23.17
	oxo-MiNP											0%	1.39	2.17	3.66	9.65	3.07%	1.19	2.08	3.62	14.93
	cx-MiNP											0%	2.88	4.55	7.5	19.5	0.61%	3.85	6.11	9.91	45.64
	OH-MiDP											0.78%	0.75	1.19	2.06	5.9	1.84%	1.23	1.76	2.83	5.11
	oxo-MiDP											10.5%	0.29	0.54	0.89	2.56	10.4%	0.42	0.62	0.96	1.82
	cx-MiDP											2.14%	0.41	0.7	1.19	3.62	0.61%	1.05	1.43	2.27	4.87
	MMP											1.55%	3.21	5.07	10.44	36.0	4.91%	2.01	2.69	4.2	10.56

Substance Group	Biomarker	3XG (Belgium)					CELSPAC—FIREexpo; Controls (Czech Republic)					GerES V (Germany)					BIOAMBIENT.ES (Spain)				
Distribution		% < LOQ	P25	P50	P75	P95	% < LOQ	P25	P50	P75	P95	% < LOQ	P25	P50	P75	P95	% < LOQ	P25	P50	P75	P95
PAHs	1-OH-Pyr	1.6%	0.11	0.15	0.24	0.49	0%	0.06	0.10	0.13	0.26	1.36%	0.06	0.09	0.14	0.29					
	4-OH-Phe						20%	0.02	0.05	0.12	1.5	0.39%	0.02	0.04	0.08	0.26					
	1-OH-Phe						5.5%	0.08	0.17	0.34	0.70	0%	0.08	0.12	0.2	0.46					
	2-OH-Flu						0%	0.21	0.36	0.56	1.0	10.5%	0.23	0.43	0.69	2.19					
	2-OH-Nap						0%	3.0	5.2	7.1	21	0.19%	1.86	3.15	5.89	15.9					
	1-OH-Nap						0%	1.0	1.7	3.3	6.2	3.5%	0.36	0.68	1.41	4.88					
Bisphenols	BPA	2.4%	0.9	1.29	2.3	4.61	0%					3.69%	1.03	1.6	2.88	6.91					
PFAS	PFNA						0%	0.23	0.3	0.36	0.49						0.61%	0.7	0.95	1.39	2.14
	PFDA						0%	0.11	0.12	0.17	0.25						11.0%	0.26	0.37	0.53	0.84

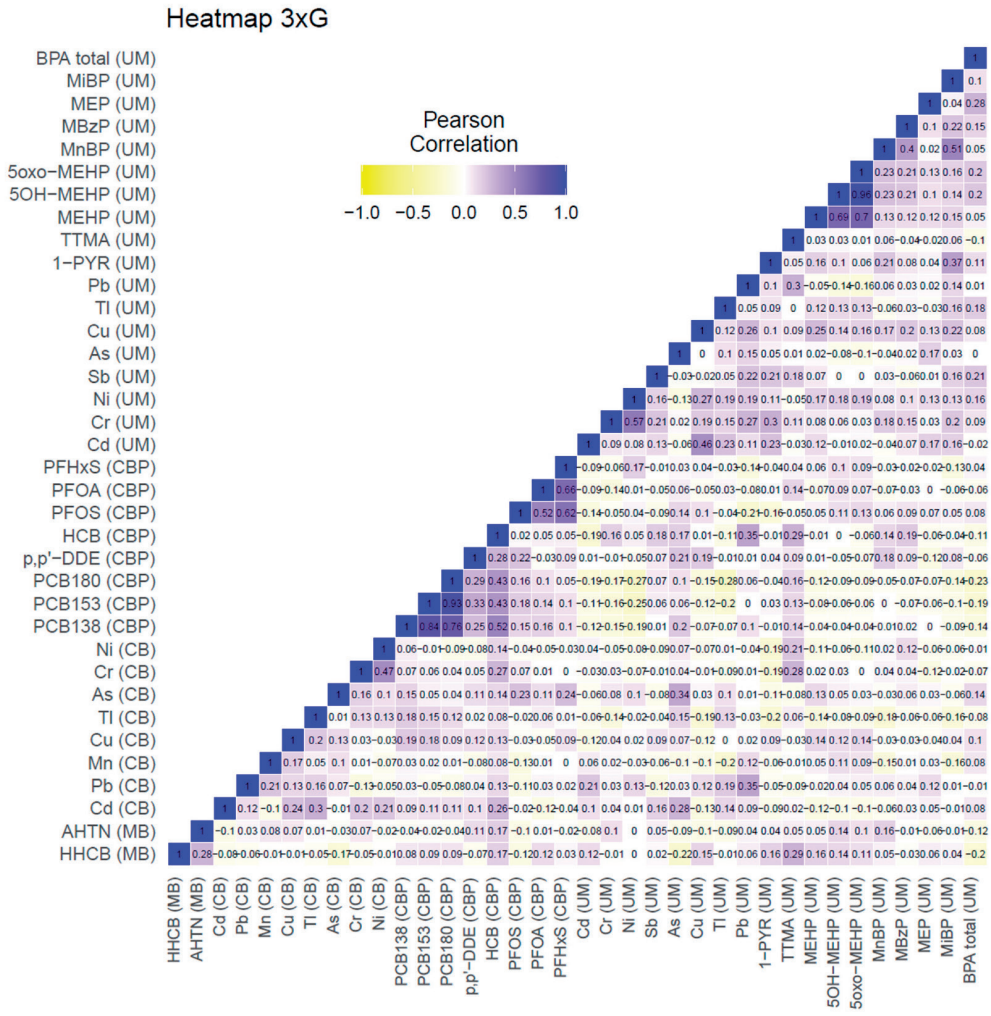


Figure 1. *3xG: heatmap showing the Pearson correlations between all creatinine-standardized and lipid-standardized measured biomarkers measured in urine and blood, respectively, available for the selected subset of participants. Data were corrected for age, BMI, and smoking status of the participants. The matrices in which biomarkers were measured are shown between brackets (MB: maternal blood, CB: cord blood, CBP: cord blood plasma, UM: morning urine).*

CELSPAC—FIREexpo (Czech Republic)

In the CELSPAC—FIREexpo study, data for the following substances were used (please note that the list of substances in the control group and firefighters might slightly differ due to differences in percentage above LOD/LOQ, see Materials and Methods): serum PFASs, i.e., PFPeA, PFHxA, PFOA, PFNA, PFDA, PFUnDA, PFBS, PFHxS, PFHpS, and PFOS, and urine OH-PAHs, i.e., 1-NAPH, 2-NAPH, 2-FLUO, 3-FLUO, ∑(2-PHEN + 3-PHEN), 1-PHEN, 4-PHEN, and 1-PYR.

Supplementary Table S2 shows the descriptive statistics for all biomarkers used in the analysis for the firefighters and the control group. The summed exposure to PFASs is significantly higher in firefighters than in the control group (Mann–Whitney U test, $p < 0.05$). When

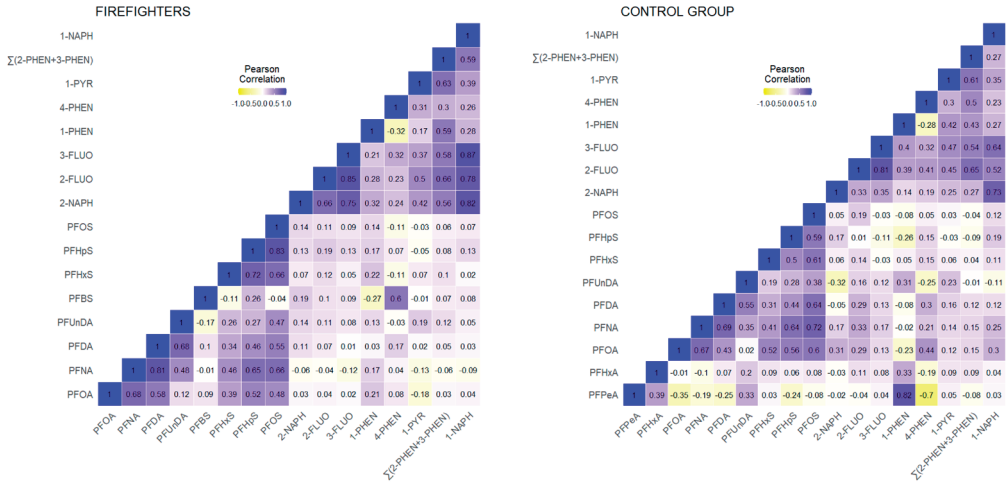


Figure 2. *CELSPAC—FIREexpo: heatmap showing the Pearson correlations between serum PFASs and creatinine-standardized urinary OH-PAHs for firefighters (left) and the corresponding control group (right). Data were corrected for age and BMI.*

assessing individual substances, the levels of all measured PFASs are higher in firefighters than in the control group, except for PFPeA and PFUnDA. No significant difference was observed in the summed exposure levels for OH-PAHs between the firefighters and the control group (Mann–Whitney U test, $p < 0.05$); however, the levels of individual OH-PAHs slightly differ between the firefighters and the control group (Řiháčková et al., 2023).

Figure 2 shows the correlation heatmap for the biomarkers included for the CELSPAC–FIREexpo study. The correlations of biomarkers for substances belonging to the same family of chemicals are generally higher compared to those that belong to different chemical families. This trend is more prominent in firefighters, where the correlations within a chemical family slightly increased, while the correlations between substances from different chemical families remained weak (except for the correlation between PFBS and 4-PHEN). In the control group, the heatmap was more heterogeneous, and the within-family correlations were slightly weaker compared to firefighters, but some moderate correlations were observed for chemicals from different families.

GerES V (Germany)

The following substances were included in first-morning void urine samples in the selected subset of 515 participants: cadmium (Cd), chromium (Cr), mercury (Hg), phthalates, DINCH, bisphenol A (BPA), polyaromatic hydrocarbons (PAHs), acrylamide, pesticides, aprotic solvents (n-ethyl-pyrrolidone; n-methyl-pyrrolidone), UV-filters (benzophenones (BP)), antimony (Sb), selenium (Se), parabens, lysmeral (TBBA), and CIT/MIT (methylchloroisothiazolinone/methylisothiazolinone). From the above set, 10 biomarkers were excluded from the network analyses because more than 40% of the measurements were below LOQ: phthalate metabolites MnOP, MnPeP, MCHP, OH-MPHP, and cx-MPHP; the aprotic solvents metabolite 5-HNEP; the pesticide glyphosate and its metabolite AMPA; and the UV-filter metabolites of BP-1 and BP-3. As a

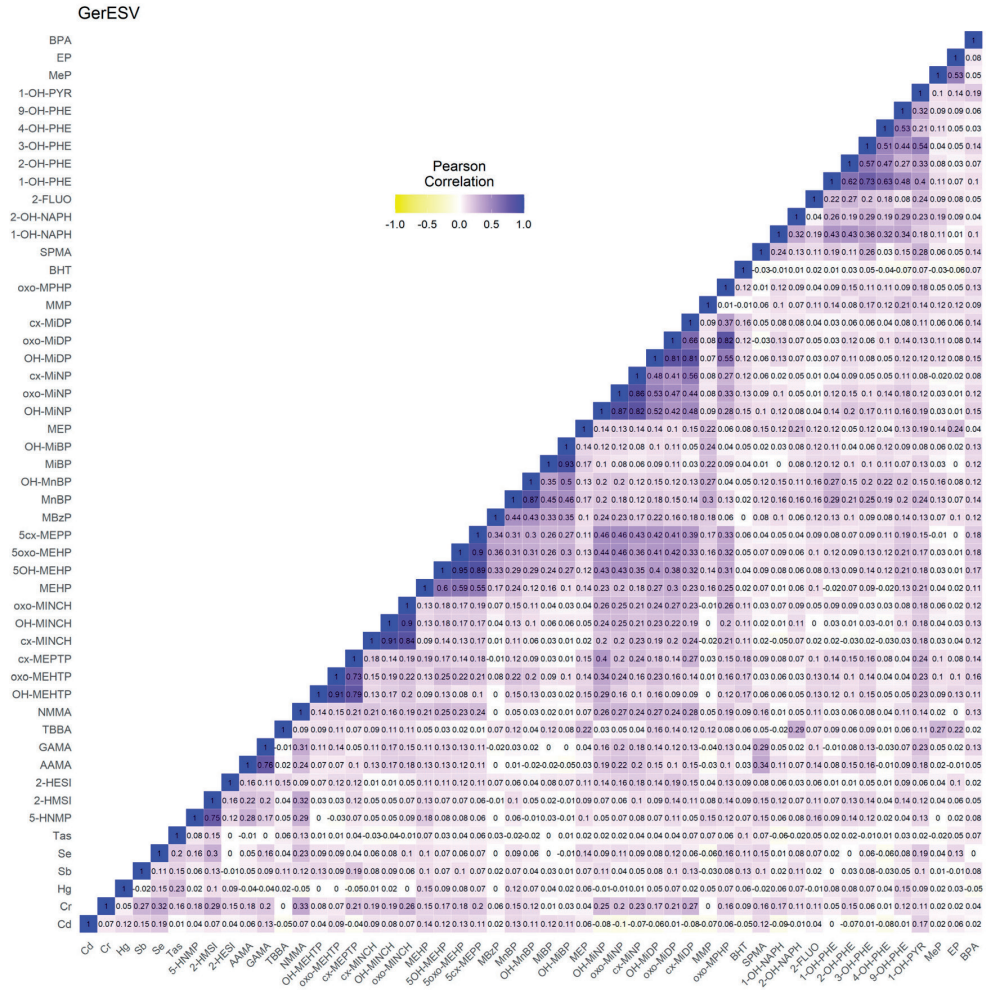


Figure 3. *GerES V: heatmap showing the Pearson correlations between all measured creatinine-standardized biomarkers available for the selected subset of participants. Data were corrected for age, sex, body mass index (BMI), smoking status of the participant, and education of the household.*

result, a total of 51 biomarkers were included in the analyses (see Table 1 and Table S3). Missing data in biomarker data were imputed as described in the Materials and Methods section (Data Selection and Preparation).

Supplementary Table S3 shows all substances included for network analyses in GerES V, their proportions below LOQ, and percentiles and geometric mean of the creatinine-standardized biomarker concentrations. Figure 3 shows the correlation heatmap for the biomarkers included for GerES V, using data standardized for creatinine and corrected for the determinants age, sex, BMI, smoking status of the participant, and education of the household. The heatmap shows mostly positive, small to medium correlations. For example, chromium and NMMA show correlations around 0.3 with several metabolites from other substance groups such as acrylamide, aprotic solvents, and some phthalates, whereas the lowest correlations

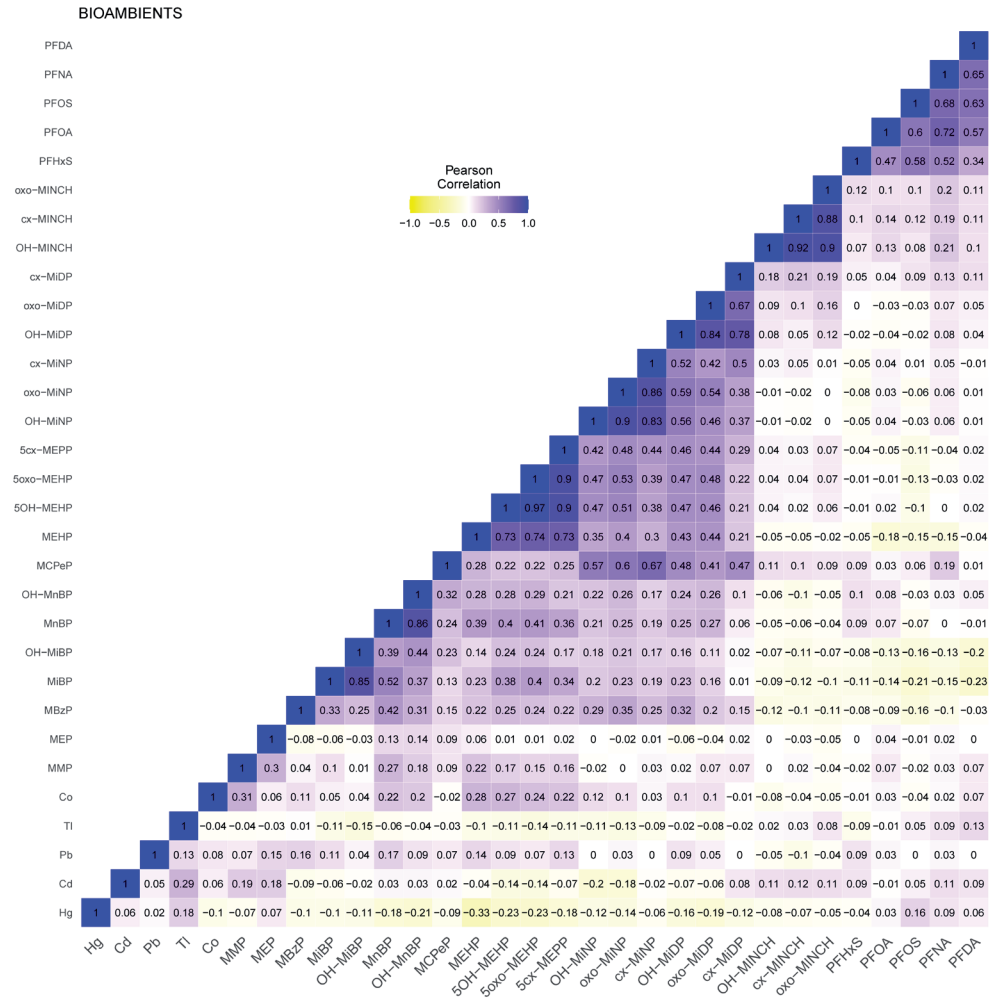


Figure 4. *BIOAMBIENT.ES: heatmap showing the Pearson correlations between all measured creatine-standardized biomarkers available for the selected subset of participants. Data were corrected for sex, age, body mass index (BMI) and smoking status of the participants.*

with other substance groups ($r \approx 0–0.27$) are observed for phthalate substitute DINCH, arsenic, mercury, and parabens. In contrast, correlations between metabolites of the same substance showed the highest correlations (up to $r \approx 0.95$), e.g., acrylamide and glycidamide, and phthalates and their substitute DINCH and DEHTP.

BIOAMBIENT.ES (Spain)

The selected subset of 163 participants had data on biomarkers for the following substances: metals, i.e., mercury (Hg), cadmium (Cd), lead (Pb), thallium (Tl), and cobalt (Co), phthalates (DMP, DEP, BBzP, DiBP, DnBP, DEHP, DiNP, and DiDP), DINCH, and PFASs (PFHxS, PFOA, PFOS, PFNA, and PFDA). As a result, a total of 31 biomarkers were included in the analyses. Metals and phthalates were measured in valid morning void urine and PFAS in blood. Missing values in

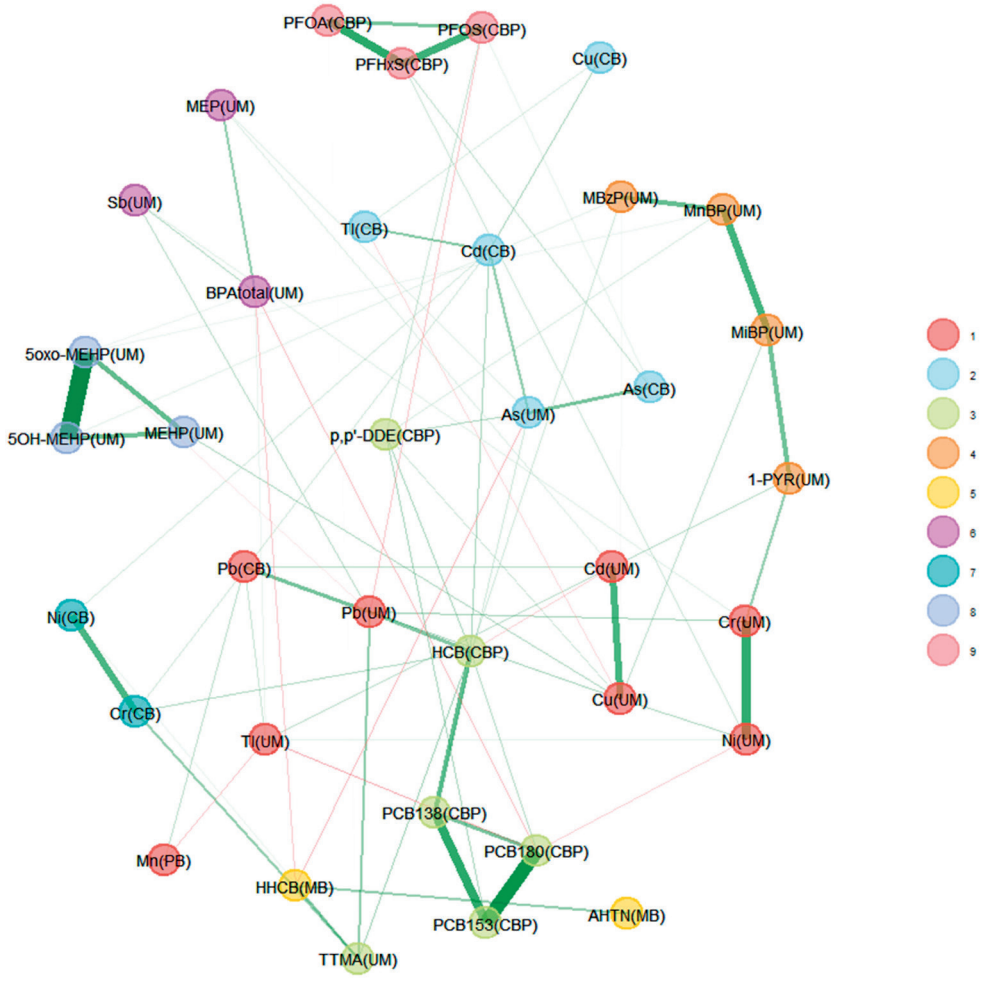


Figure 5. *Weighted network for 3xG. The data were corrected for age, smoking, and BMI. Urinary markers were standardised for creatinine and lipid soluble blood markers were standardised for lipids. Matrices in which biomarkers are measured appear between brackets. Green lines represent a positive dependency between nodes (biomarkers) while red lines represent a negative dependency.*

biomarker data were imputed as described in the Materials and Methods section. Descriptive statistics for this set of biomarkers are shown in Table 1 and Table S4. Figure 4 shows the correlation heatmap for the biomarker included for BIOAMBIENT.ES, using data standardization for creatinine and corrected by age, sex, body mass index (BMI), and smoking status of the participant. The heatmap showed positive and negative, mainly small to medium correlations. The correlation among metabolites of the same group of substances showed higher positive correlations, except for metals and some phthalates such as MEP. In addition, some negative correlations were observed among PFAS or DINCH, and most of the phthalates. Mercury and thallium showed negative correlations with most biomarkers, except for PFAS.

Network Analysis

The network analyses produce a graphical representation of the conditional independence between the observed biomarker levels. Different colours in the networks indicate the clustering structure or communities and which biomarkers are more closely related to one another compared to the rest of the network. The sensitivity analysis of the networks consisted of two parts. The first part comprised a comparison of two weighted network estimation approaches. This was performed on the Belgium 3xG data. Secondly, the impact of different approaches on correcting biomarker levels against creatinine levels (as a measure for the level of dilution of the urine sample) was evaluated using the German GerES V data. The results of both comparisons are presented in Appendix A.

3xG (Belgium)

Figure 5 shows the weighted network for the 3xG subset of participants (n = 125). Biomarkers measured in urine are standardised for creatinine and lipid-soluble biomarkers in blood are standardised for lipids. Nine different communities were identified (represented by the different colours), with the strongest relations within the communities (thick lines). Negative correlations (red lines) were minimal. Green lines represent positive associations while red signify negative associations between biomarker levels. Communities with biomarkers originating from the same chemical group were detected, such as the musks (HHCB and AHTN, community 5 in yellow) or the heavy metals. The heavy metals were, however, split into three separate communities (numbers 1, 2, and 7 in Figure 5).

In line with what was observed in the heatmap of 3xG (Figure 1), As and Pb measured in the urine of the mother during pregnancy and in the cord blood of the newborn at birth are highly related, which is in agreement with previous studies on the migration of hazardous heavy metals through the placenta to the fetus (Rísová, 2019; Vahter, 2008). Other interesting communities can be observed in Figure 5. For example, community number 6 shows a relationship between total BPA, MEP, and Sb measured in the urine of the mother during pregnancy. Both BPA and phthalates have been found in packaging for cosmetic and personal care products and food packaging materials (Benjamin et al., 2017; Schettler, 2006), and the use of make-up has been previously associated with an increase in BPA and MEP in urine (Fisher et al., 2019). The relationship between Sb and total BPA could be explained due to their presence in plastic containers that leach plasticizers and plastic additives into water or other food products (Andra et al., 2013). Interestingly, this association is not seen in Figure 6 in a subset of participants with high educational level compared to a subset of participants with low educational level, which may be due to the fact that women with a higher educational level are more aware of the leaching of chemicals from plastic containers to water or food products. The relation of total BPA with MEP was not detected in either network once the data was stratified. Overall, the networks observed for the higher educated subset appear to be more connected with larger communities, having more (red) connections between nodes across communities.

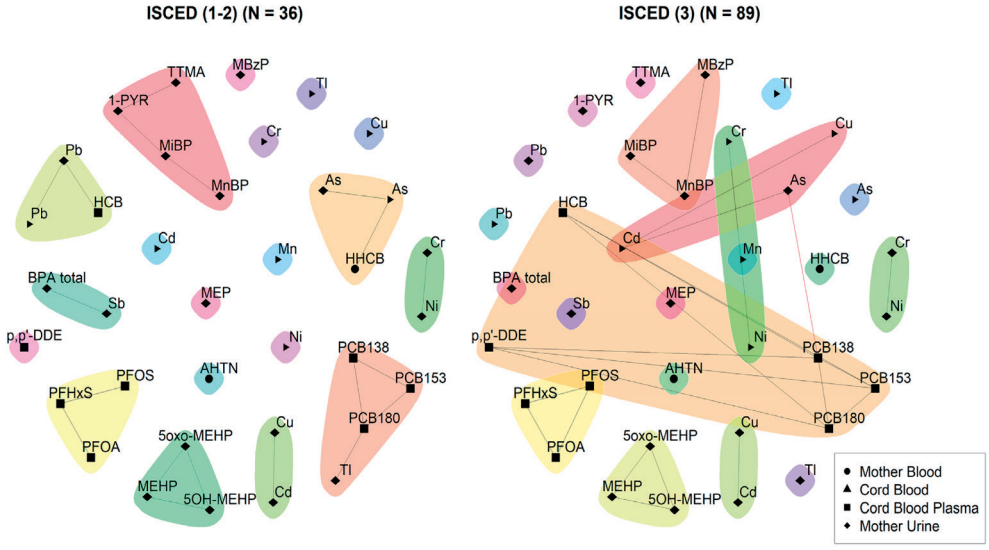


Figure 6. *Unweighted network for 3xG for participants with low educational level (left) and participants with high educational level (right). The data were corrected for age and BMI. Urinary markers were standardised for creatinine and lipid soluble blood markers were standardised for lipids. Low ISCED is defined by participants belonging to educational levels 0–4 according to the ISCED (International Standard Classification of Education) and high ISCED is defined by participants belonging to educational level ≥5. Black lines indicate dependency between nodes (biomarkers) within a community; red lines indicate dependency between nodes in different communities.*

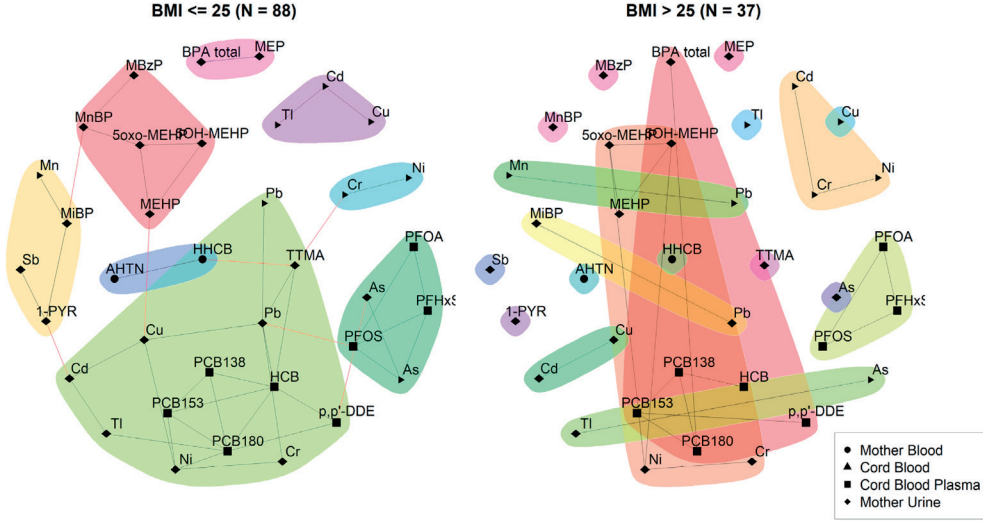


Figure 7. *Unweighted network for 3xG for participants with a BMI ≤ 25 kg/m² and participants with a BMI > 25 kg/m². The data were corrected for age and smoking. Urinary markers were standardised for creatinine and blood markers were standardised for lipids. Black lines indicate dependency between nodes (biomarkers) within a community; red lines indicate dependency between nodes in different communities.*

In Figure 5, another interesting community is the one consisting of 1-PYR and MiBP, MnBP, and TTMA. The most important route of exposure for 1-PYR is through smoking; however, living in a highly polluted environment also has an influence on the 1-PYR levels (Llop et al., 2008). No common route of exposure for 1-PYR and MiBP has been found in the literature. It is intriguing to notice that the link is no longer found in the network for participants with a high level of education, but it is conserved in those with a low educational level, as seen in Figure 6. Furthermore, we also noticed that the link is no longer conserved in participants with a BMI > 25 kg/m^2 while it is in participants with a BMI ≤ 25 kg/m^2 (Figure 7). Figure 7 also shows more dependencies in the low BMI category, where all substances are part of a community, with some communities comprising multiple chemical families. Additionally, some dependencies across communities can be observed. Moreover, we observe again a community of BPA and MEP. In contrast, the high BMI category displays smaller communities and many substances not part of a community.

Further stratifications were explored in Figure 8 where networks are explored for participants with a low fish consumption (less than 1–3 times per week) and relatively high fish consumption (equal or more than 1–3 times per week). While some communities are conserved, such as the PFASs, DEHP metabolites, and urinary heavy metals (Cu, Cd, Cr, and Ni), some others show slight changes, especially regarding other heavy metals measured in the cord blood of the newborn at birth.

CELSPAC—FIREexpo (Czech Republic)

Figure 9 shows the weighted network of the firefighters (n = 52) and the control group (n = 55) of the CELSPAC—FIREexpo study. The set of biomarkers differ between the two groups due to differences in percentage detected above LOQ (Supplementary Table S2).

In the firefighters' network, most PFASs and OH-PAHs clustered together in a community of the same chemical group. Two communities were created in the PFASs group (numbers 2 in blue and 4 in orange), and two in the OH-PAHs group (naphthalenes and fluorenes in community 3 in green, and other OH-PAHs in 1 in red). The exception was PFBS which was strongly linked to 4-PHEN, and therefore included in the community of OH-PAHs, rather than PFASs. In the control group network, three communities were detected: a community of naphthalenes and fluorenes (1, red), a community of seven PFAS (2, blue), and the rest of the compounds (other OH-PAHs, PFHxA, and PFPeA in the green community, 3).

In the firefighters' network, the intra-community links were strong, and there were weak inter-community links, resulting in more strictly separated PFASs and OH-PAHs communities, while in the control group, more inter-community links were present, resulting in communities with substances from different chemical families. This might be caused by the firefighting occupation being the predominant exposure factor contributing to the PAHs and PFASs exposure in firefighters. In the controls, the levels of PFASs and PAHs are, in general, lower than in firefighters and there might not be a predominant exposure source contributing to stronger communities of PFASs and PAHs.

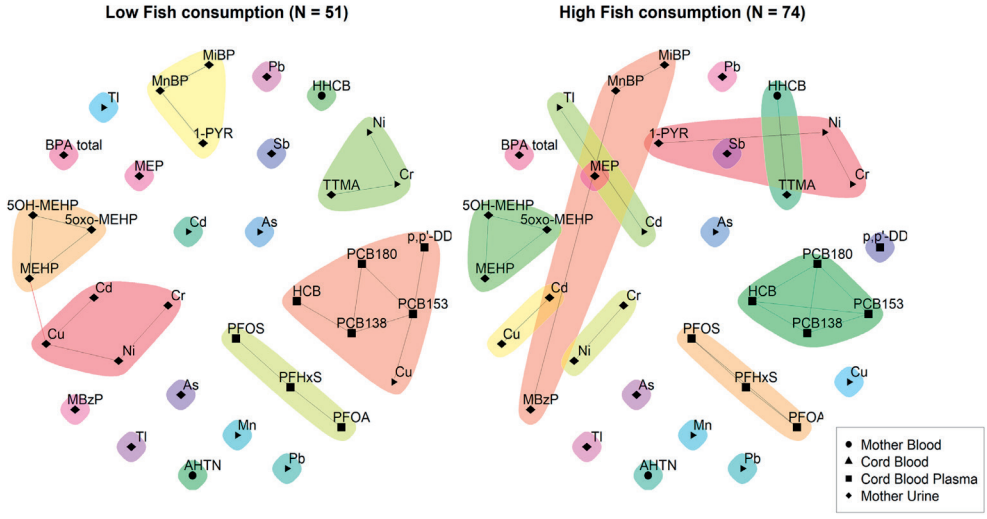


Figure 8. *3xG: unweighted network for participants with a low fish consumption (left) and participants with a relatively high fish consumption (right). The data were corrected for age and smoking. Urinary markers were standardised for creatinine and lipid soluble blood markers were standardised for lipids. Low fish consumption is defined as consumption of fish less than 1–3 times per week and high fish consumption is defined as fish consumption of at least 1–3 times per week. Black lines indicate dependency between nodes (biomarkers) within a community; red lines indicate dependency between nodes in different communities.*

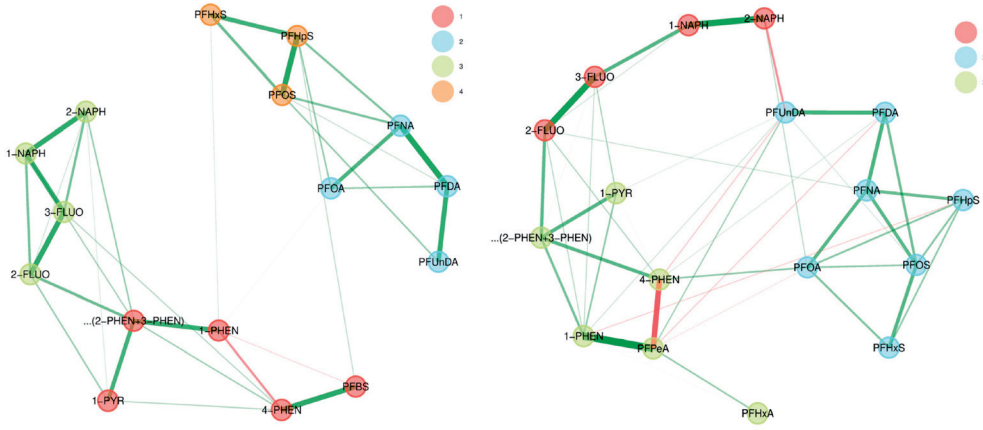


Figure 9. *Weighted network for CELSPAC—FIREexpo firefighters (left) and the control group (right). The data were corrected for age and BMI. Urinary markers (OH-PAHs) were standardized for creatinine. Green lines represent positive associations while red signify negative associations between biomarker levels.*

GerES V (Germany)

The weighted network for GerES V, allowing for assessment of the strength of the links between substances, is shown in Figure 10. Ten communities were identified. Links were stronger (i.e., thicker lines) within substance groups and among metabolites from the same parent compound; the strongest links were observed within acrylamide, aprotic solvents, parabens EP and MeP, DINCH, DEHTP, and several phthalates.

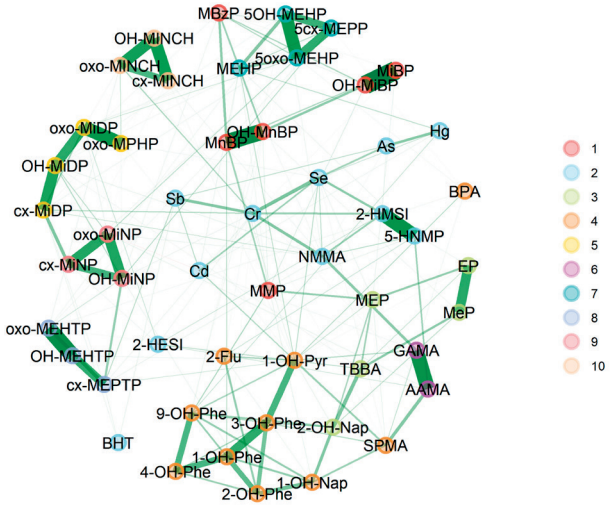


Figure 10. *Weighted network for GerES V subsample, using creatinine-standardised and creatinine-adjusted data. Data were corrected for age and BMI. Green lines represent a positive dependency between nodes (biomarkers).*

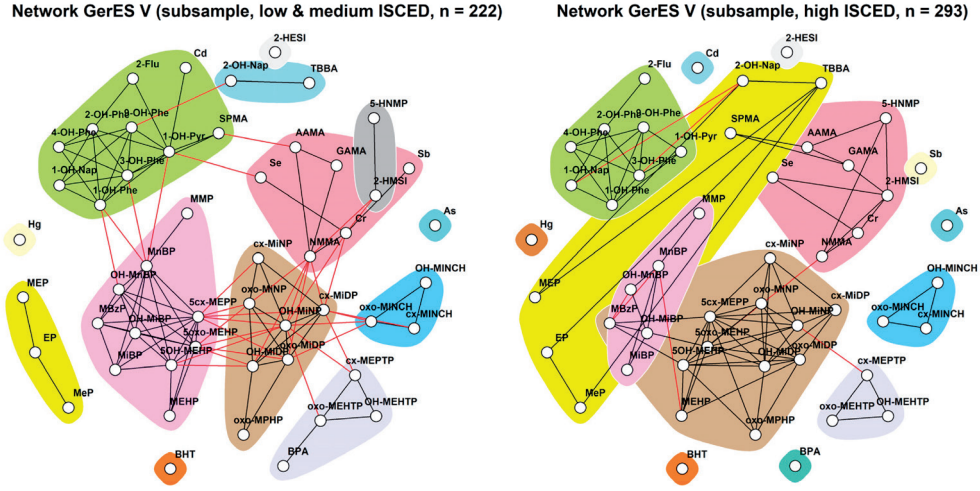


Figure 11. *Stratification of the network for the GerES V subsample by education (ISCED), using creatinine-standardized and creatinine-adjusted data. Data were corrected for sex, smoking status, age, and BMI. Low ISCED reflects educational levels 0–4 from the ISCED (International Standard Classification of Education) and high ISCED reflects educational level ≥5. Black lines indicate dependency between nodes (biomarkers) within a community; red lines indicate dependency between nodes in different communities.*

Comparison of the two networks stratified by education (Figure 11) revealed more differences than similarities. Few communities can be identified as similar between both groups, namely those of DINCH metabolites (blue), DEHTP metabolites (lavender), and PAHs (green). However, even within these communities, some remarkable differences can be observed between the groups. In the subset of participants from households with low to medium education (left panel), DEHTP co-occurs together with BPA, which is not the case for the higher

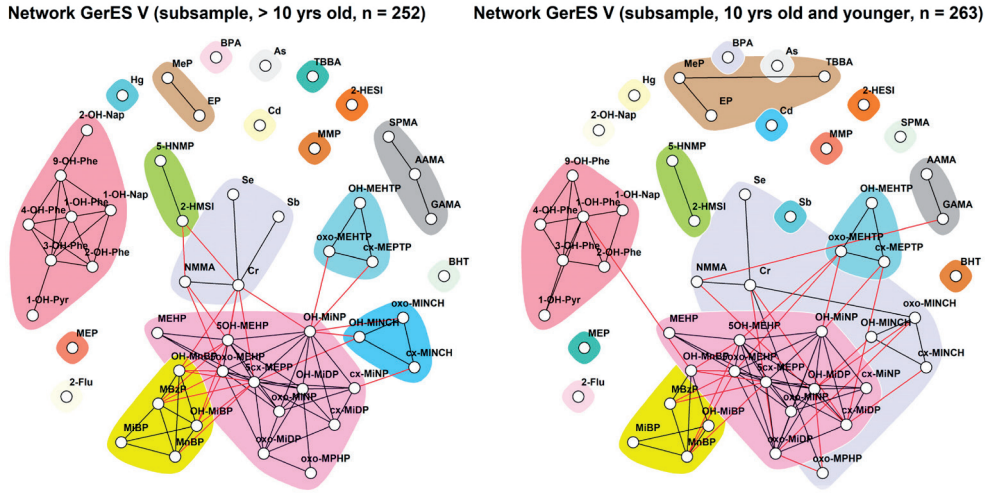


Figure 12. *Stratification of the network for the GerES V subsample by median age (10 years old), using creatinine-standardized and creatinine-adjusted data. Data were corrected for sex, smoking status, age, and BMI. Black lines indicate dependency between nodes (biomarkers) within a community; red lines indicate dependency between nodes in different communities.*

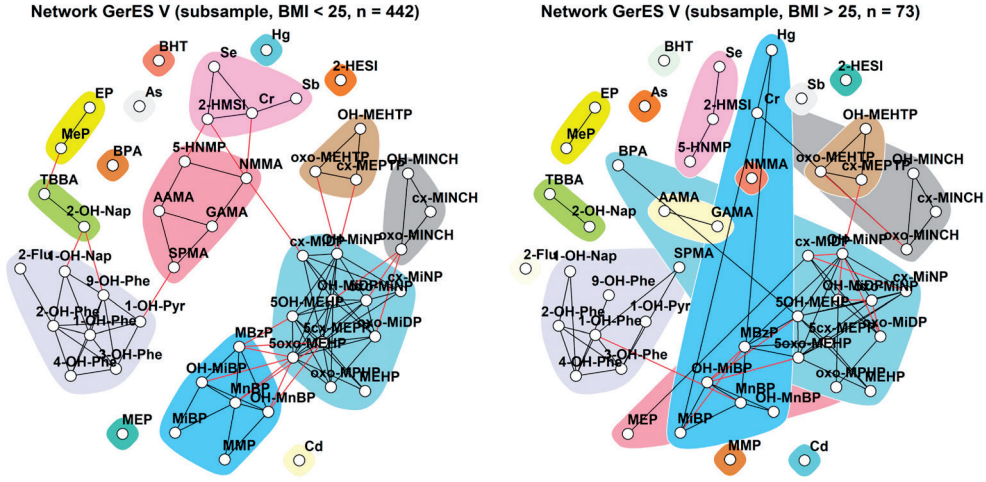


Figure 13. *Stratification of the network for the GerES V subsample by BMI, using creatinine-standardized and creatinine-adjusted data. Data were corrected for sex, smoking status, age, and BMI. Black lines indicate dependency between nodes (biomarkers) within a community; red lines indicate dependency between nodes in different communities.*

educated subset. Similarly, PAHs co-occur with cadmium and benzene (SPMA) in the lower educated subset, while this co-occurrence was not observed for participants with a higher level of education. Additionally, the networks for phthalates are different between the groups, with the major difference being DEHP: this substance is part of a different community of phthalates in each education group. Furthermore, phthalate substitutes are more inter-related with communities of phthalates among participants of low to medium educated households but occur more distinctly in children and adolescents from higher educated households. In

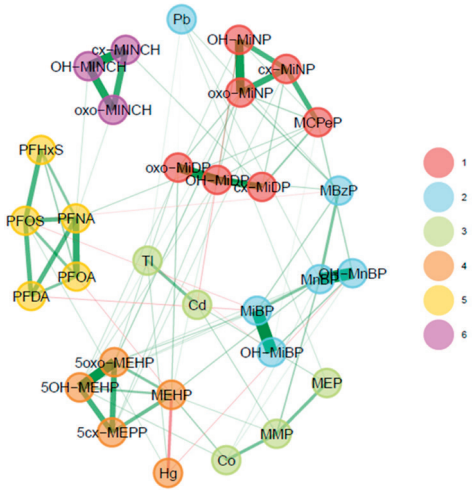


Figure 14. *Weighted network for BIOAMBIENT.ES subsample, using creatinine-standardised and creatinine-adjusted data. Data were corrected for sex, smoking status, age, and BMI. Green lines represent positive associations; red lines signify negative associations between biomarker levels.*

contrast to the 3xG observation, in GerES V, connections between nodes across communities are more prominent in the lower education group.

Figure 12 shows stratified networks by the median age of the GerES V subset, which was 10 years. Both children older than 10 and 10 years old and younger show a community each for PAHs (light pink), two aprotic solvents (HNMP and HMSI, green), and DEHTP (light blue) 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.

When comparing participants with a BMI ≤ 25 lower versus participants with a BMI > 25 (Figure 13), we observed that for participants with a higher BMI (right panel), communities are more likely to include substances from other substance groups or substances which usually stand alone. For example, the PAHs community includes in addition SPMA (salmon), the phthalate community of DnBP and DiBP co-occurs with mercury (blue), the phthalate community of DiNP, DEHP, DiDP, and BBzP co-occurs with BPA, and DINCH metabolites co-occur with chromium.

BIOAMBIENT (Spain)

Figure 14 shows the weighted network for the BIOAMBIENT.ES dataset (n = 163). The graph shows seven communities, with mostly positive dependencies between substances. As in the other studies, the strongest dependencies were observed in communities of substances from the same chemical family. Nonetheless, in addition to communities from the same chemical family, dependencies across chemical families were also observed. For example, the PFAS form one community, together with mercury (Hg) through a link to PFOS. Several metals form a community with phthalates (community 1). Correlations across communities also exist, e.g., lead (Pb) with PFAS.

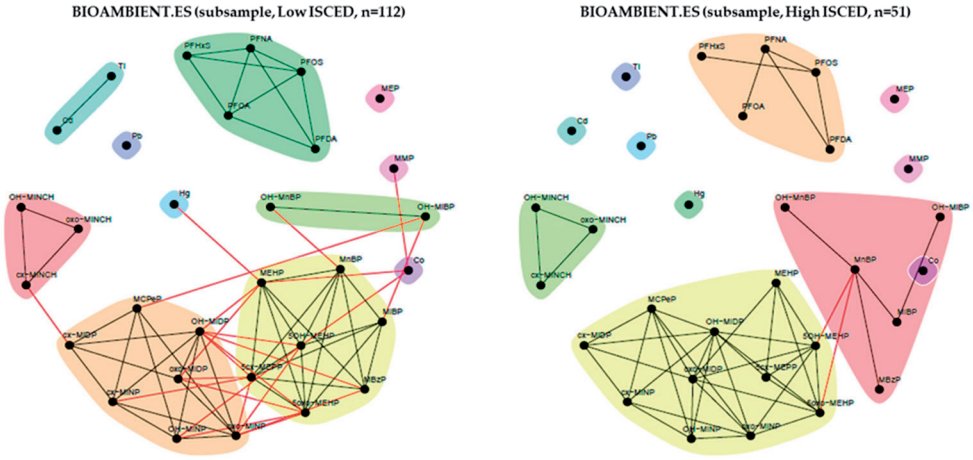


Figure 15. *Stratification of the network for the BIOAMBIENT.ES dataset by education (ISCED level), using creatinine-standardized data. Low ISCED (left panel) is defined by participants belonging to educational levels 0–4 according to the ISCED (International Standard Classification of Education) and high ISCED (right panel) is defined by participants belonging to educational level ≥5. Data were corrected for sex, smoking status, age, and BMI. Black lines indicate dependency between nodes (biomarkers) within a community; red lines indicate dependency between nodes in different communities.*

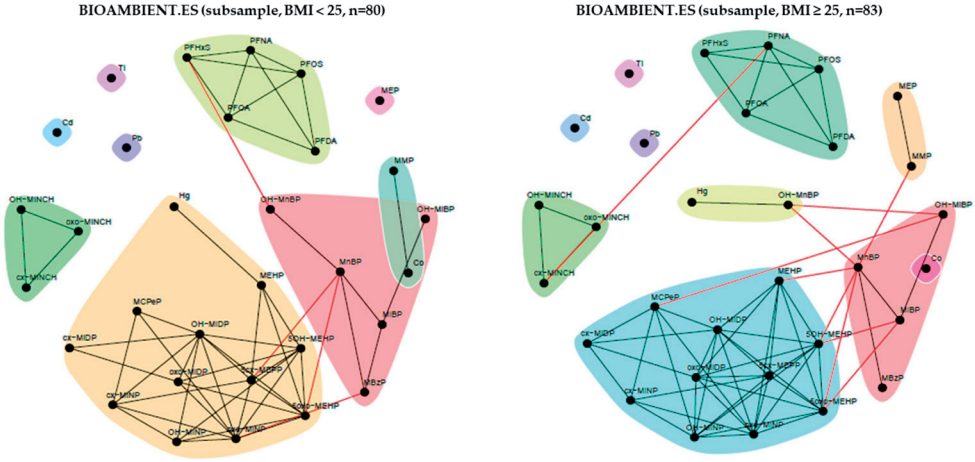


Figure 16. *Stratification of the network for the BIOAMBIENT.ES subsample by BMI, using creatinine-standardized data. Unweighted network for participants with a normal weight (defined as BMI ≤ 25) is shown in the left panel, while the network for participants with overweight (BMI > 25) is shown in the right panel. Data were corrected for sex, smoking status, age, and BMI. Black lines indicate dependency between nodes (biomarkers) within a community; red lines indicate dependency between nodes in different communities.*

Figure 14 also shows separate grouping within parent compounds in the case of phthalates: DiBP metabolites (MiBP and OH-MiBP), DEHP metabolites (MEHP, OH-MEHP, oxo-MEHP, and cx-MEPP), DiNP metabolites (OH-MiNP, oxo-MiNP, and cx-MiNP), and DiDP metabolites (OH-MiDP, oxo-MIDP, and 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.

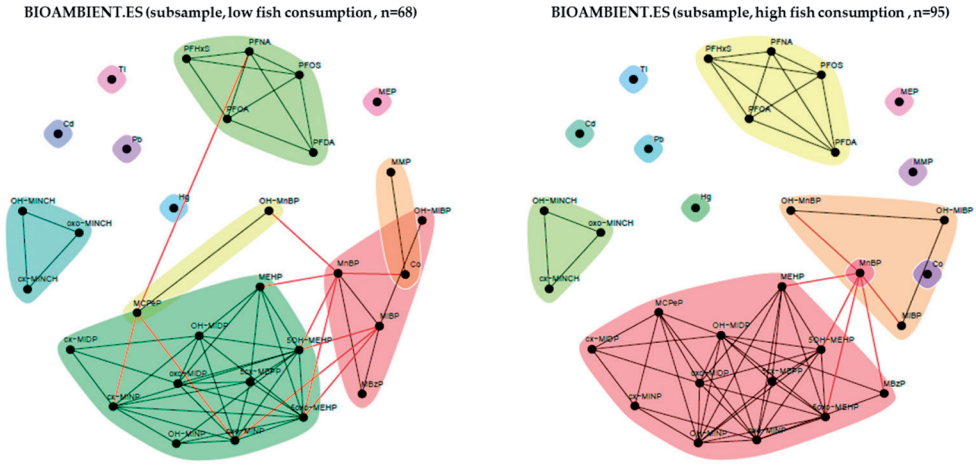


Figure 17. *BIOAMBIENT.ES: unweighted network for participants with a low fish consumption (**left**) and participants with a relatively high fish consumption (**right**), using creatinine-standardized data. The data were corrected for age and smoking. Low fish consumption is defined as consumption of fish less than 1–3 times per week and high fish consumption is defined as fish consumption of at least 1–3 times per week. Black lines indicate dependency between nodes (biomarkers) within a community; red lines indicate dependency between nodes in different communities.*

We performed stratified unweighted network analysis for relevant determinants, including educational level (ISCED), BMI, and fish consumption. The networks identified with unweighted analysis showed fewer communities than the weighted network, possibly because of the smaller number of observations within strata (Figures 15–17). Here, metals tend to appear as standalone compounds, DINCH metabolites form a distinct community, as generally so 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 stratification. MEP, a metabolite of the phthalate substance DEP, always appears separate from other phthalates and substances.

Figure 15 shows the stratified networks by education level. We found differences in the dependencies amongst PFASs, which appear as standalone substances in the lower education group (left panel), whereas PFOA–PFNA–PFDA form a community in the high ISCED group (right panel). For phthalates, in the lower ISCED level, each substance appears in a separate community, with the exception of MCPP, a metabolite of DnBP, which, as seen earlier, appears in the same community as the metabolites of DiNP. In contrast, in the high ISCED group, there are two main communities showing dependencies between them. Similar to the GerES V study, the community of DINCH metabolites was not different between the groups.

In the stratification by BMI (Figure 16), we observed that for participants with lower BMI (BMI < 25), mercury is included in DEHP, DiNP, and DiDP community, whereas in the high BMI group (BMI ≥ 25), mercury and OH-MnBP form a separate community.

When evaluating the effect of fish consumption (Figure 17), we observed communities, mainly comprising substances from a single chemical family, in both groups. In participants with a relatively low fish consumption, MCPP – OH-MnBP and MMP – Co were grouped in independent communities.

DISCUSSION

In this study, we applied network analysis to HBM datasets from Belgium, Czech Republic, Germany, and Spain, with the aim to further explore its added value for mixture risk assessment. The network approach combined with a clustering algorithm (community detection) proved to be an intuitive graphical manner to describe the correlation structure in a dataset, taking into account all exposure markers in the mixture. Application of the network analysis in this study revealed some new insights in inter-dependencies within each dataset. Importantly, pan-European application of these methods and their interpretation would require harmonization across Europe in terms of study design, biomarker media, chemical analysis, and the substances that are assessed. Overall, the four studies yielded diverse correlations, with more positive than negative associations (Figure 1, Figure 2, Figure 3 and Figure 4). With the exception of parent–metabolite relations, correlations were generally below 0.8, while negative correlations were generally below 0.3. It should be noted that in this study, the focus was rather on the dependencies between biomarkers (correlation structure), and not so much on the absolute levels of exposure. Nonetheless, when interpreting differences or commonalities in community patterns across studies, one should be aware that sometimes marked differences exist between studies in biomarker levels, sometimes up to one or two orders of magnitude. These may reflect differences in study population, in design, chemical analytical procedures, and actual differences in exposure patterns between study populations. In case the output should be used for prioritizing mixtures of concern, of course the absolute levels should be considered as well.

The network analysis identified in all four studies, as expected, several communities of chemical families, e.g., phthalates and PAHs. Additionally, links between parent substances and metabolites were observed, e.g., for acrylamide and glycidamide. However, also exposure patterns involving substances from different chemical families were observed. Examples include the dependency between 1-PYR (biomarker for PAHs), TTMA (biomarker for benzene), and the phthalates MiBP and MnBP in the 3xG study, and the dependency between acrylamide, its metabolite glycidamide, SPMA (biomarker for benzene), and aprotic solvents (NMMA, HNMP, and HMSI) in the GerEs V study. In the CELSPAC—FIREexpo study, the network analysis revealed both positive (e.g., 4-PHEN and PFBS in firefighters) and negative (e.g., 4-PHEN and PFPeA in controls) dependencies between PAHs and PFASs. Such communities, comprising substances from different chemical families, possibly reflect a commonality in exposure patterns and thus reflect real-life mixture patterns. The communities observed may also be impacted by similarities in physicochemical properties of the substances involved. Our findings also show that in the German and Spanish data, metals (e.g., arsenic and mercury) were not always part of communities, in contrast to the Belgian data. Additionally, in the German weighted network (Figure 10), BPA was not part of a community, while a relatively strong correlation between BPA and MEP was observed in the Belgian weighted network (Figure 5). In contrast to the mostly positive links observed in the weighted networks in the three larger studies, in the smaller CELSPAC-FIREexpo control group network, a negative dependency could be observed (between 4-PHEN and PFPeA).

The unweighted network analysis stratified by covariates demonstrated differences in the community patterns. These may reflect differences in exposure patterns and pathways between strata, although no clear interpretation can be given at this point. The differences between strata may also reflect some sample differences between strata. The stratified unweighted networks also show many dependencies across communities, as indicated by the red lines in the graphs. Even though the unweighted network analysis showed differences between strata, no obvious immediate clues about sources or exposure pathways were observed. Nonetheless, the communities in the network analysis may hold some indications about relevant exposure routes. For example, the community of parabens (MeP and EP), preservatives in cosmetics, with lysmeral (TBBA), a fragrance in cosmetics, in the German network results would point at the role of cosmetics.

The above (and other) differences between studies may deserve further investigation; however, we here explicitly abstain from doing so because of the differences in study designs. Firstly, the populations sampled highly differ across the four studies. The German study focused on exposure in adolescents, the Spanish study on subjects aged 16 or over, while the Belgian study combined data from mothers and newborns, with different time points of sampling, and the Czech study focused on occupational exposure in firefighters. Additionally, the biomarkers, and thus the substances, included in the four studies vary. The same applies to the matrices in which biomarkers were determined: in the German study, only urine samples were used, while in the other three studies also blood samples were included. Hence, differences observed between the studies may stem not only from differences in exposure patterns, but also from differences in various aspects of the study designs. For a better interpretation of cross-country differences, a harmonized sample collection and laboratory analysis would be beneficial.

The analyses applied comprised both weighted and unweighted network analyses. The weighted and unweighted network analyses yielded generally similar results (data not shown). While weighted network analysis is more computationally intensive and less fit for high dimensional data in comparison to the unweighted networks, a clear advantage is the indication of the relative strength of the links and the direction of the association (Horvath, 2011). For a comparison between determinants within a study, only unweighted networks were used for their ease in interpretation (occurring or not-occurring edges between the biomarkers). Future work could also include a comparison between determinants based on weighted correlation networks.

The results of our study clearly show that network methods become more informative when biomarkers for a larger number of substances are included in the HBM dataset, as demonstrated, e.g., by the findings for the GerES V study versus the CELSPAC—FIREexpo study. Existing HBM studies typically have a limited number of individuals in which a wide range of chemical substances has been measured. This hampers the potential to identify patterns of chemical mixtures, and even more so to study the role of determinants, with fewer observations per stratum. For future studies, we therefore recommend to expand, where possible, the number of observations with a wide(r) range of chemicals, to improve the ability to identify real-life mixtures and to study determinants of the patterns observed.

Regarding the methodology applied, some aspects certainly deserve further improvement. Firstly, better insight into the stability and consistency of the identified networks and communities is needed (Bodinier et al., 2021). Further work should also include characterization of the uncertainty in the networks, and the decision for the community detection algorithm (Orman et al., 2009). Better insight into aspects such as the impact of measurement errors on the networks and communities identified will enhance the appreciation of the possibilities and limitations of network analysis of HBM data for mixture risk assessment. This is crucial for its acceptance and implementation in regulatory risk assessment. Further work should also be conducted on the interpretation of the communities and the possible impact for regulatory risk assessment. We consider it crucial to take into account the toxicological properties and mechanisms of the chemical substances included in a community, because this may indicate which communities might be of more toxicological concern compared to others. Furthermore, in cases where chemicals from different families appear together in the same community, the different families may fall under different legislations and/or regulations. Such a situation would give rise to the question of how to deal with this in regulatory risk assessment.

Taken together, our study demonstrates that network analysis of HBM data allows for the identification of real-life exposure patterns to chemical mixtures occurring at a single point in time in the human body. Network analysis can be a good addition to other data explorative methods, such as heatmaps or principal component analysis. The derived networks and accompanying communities should, therefore, not replace existing methods, but rather complement and assist researchers in the description of complex mixtures in HBM data.

Graphical visualization of the networks and communities identified greatly aids the interpretation of the output. Weighted network analysis reveals the strength and direction of the links between substances identified as co-occurring, while stratification provides insight into the impact of determinants on the exposure patterns. These features make network analysis of HBM data a useful, valuable tool for mixture risk assessment.

Author Contributions

Conceptualization, **J.V., I.O., E.G.** and **E.L**.; methodology, **I.O., J.V., E.G., E.L**. and **M.L**.; data curation, **L.R.M., N.V., K.Ř., M.J.M**. and **S.P-D.;** analysis, **L.R.M., N.V., K.Ř., M.J.M**. and **E.V.-J**.; writing—original draft preparation, **L.R.M., I.O., N.V., P.S., K.Ř., E.G., S.P.-D., E.L., J.V**. and **M.L**.; writing—review and editing, **L.R.M., I.O., N.V., M.K.-G., P.S., K.Ř., M.J.M., E.V.-J., E.G., S.P.-D., E.L., J.V.** and **M.L**.; supervision, **M.L**.; funding acquisition**, M.K.-G**. and **E.L**. All authors have read and agreed to the published version of the manuscript.

Funding

This research was carried out under the HBM4EU project. The HBM4EU project has received co-funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 733032. The work was supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 857560. This publication reflects only the author's views, and the European Commission is not responsible for any use that may be made of the information it contains. K.Ř. thanks the RECETOX RI (No LM2018121)

financed by the Ministry of Education, Youth and Sports, and OP RDE—project CETOCOEN EXCELLENCE (No CZ.02.1.01/0.0/0.0/17_043/0009632) and the CETOCOEN PLUS project (CZ.0 2.1.01/0.0/0.0/15_003/0000469) for supportive background. GerES V received funding by the German Ministry for the Environment, Nature Conservation, Nuclear Safety and Consumer Protection. The 3xG study is commissioned and co-financed by NIRAS/ONDRAF (Belgian National Agency for Radioactive Waste and enriched Fissile Material), STORA (Study and Consultation Radioactive Waste Dessel) and MONA (Mols Overleg Nucleair Afval). BIOAMBIENT.ES received funding from the Spanish Ministry of Agriculture, Food and Environment (MAGRAMA) and the Instituto de Salud Carlos III (ISCIII) (SEG 1251/07 and SEG 1210/10).

Institutional Review Board Statement

The 3xG study was approved by the Ethics Committee of the University hospital in Antwerp (UZA) on 8 November 2010. The CELSPAC-FIREexpo study was approved by the ELSPAC Ethics Committee, ethical approval number No: ELSPAC/EK/1/2019. All participants received an information brochure and participated in personal interviews to be fully informed about the study and their participation. All data were pseudonymised to protect the identity of the participants. Regarding the GerES V study, the Ethics Committee of the Berlin Chamber of Physicians (Eth-14/14) and the Federal Officer for Data Protection and Freedom of Information (III-425/009#0018) had approved the project. The BIOAMBIENT.ES study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Ibermutuamur (protocol code 2009/1; date of approval 13 January 2009).

Informed Consent Statement

Informed consent was obtained from all subjects and or their legal guardians involved in the reported studies.

Data Availability Statement

Summary data are listed in Table 1 and Tables S1–S4. As the individual data are considered pseudonymized data, those cannot be made public according to the European General Data Protection Regulation (GDPR). Raw data from the CELSPAC—FIREexpo study are available in anonymized form upon request and upon the approval of the steering committee. Aggregated data from the BIOAMBIENT study can be found at IPCHEM (https://ipchem.jrc.ec.europa.eu/, accessed on 27 January 2023).

Acknowledgments

The authors gratefully acknowledge all participants of the studies reported as well as the constructive contribution of the data owners of the HBM databases used, and the use of their computational facilities.

Conflicts of Interest

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

REFERENCES

Agier, L., Portengen, L., Chadeau-Hyam, M., Basagaña, X., Giorgis-Allemand, L., Siroux, V., … Vermeulen, R. (2016). A Systematic Comparison of Linear Regression–Based Statistical Methods to Assess Exposome-Health Associations. *Environmental Health Perspectives*, *124*(12), 1848–1856. https://doi.org/10.1289/EHP172

Andra, S. S., & Makris, K. C. (2013). Incorporating potable water sources and use habits into surveys that improve surrogate exposure estimates for water contaminants: the case of bisphenol A. *Journal of Water and Health*, *12*(1), 81–93. https://doi.org/10.2166/wh.2013.068

Barrera-Gómez, J., Agier, L., Portengen, L., Chadeau-Hyam, M., Giorgis-Allemand, L., Siroux, V., … Basagaña, X. (2017). A systematic comparison of statistical methods to detect interactions in exposome-health associations. *Environmental Health*, *16*(1), 74. https://doi.org/10.1186/s12940-017-0277-6

Benjamin, S., Masai, E., Kamimura, N., Takahashi, K., Anderson, R. C., & Faisal, P. A. (2017). Phthalates impact human health: Epidemiological evidences and plausible mechanism of action. *Journal of Hazardous Materials*, *340*, 360–383. https://doi.org/https://doi.org/10.1016/j.jhazmat.2017.06.036

Bodinier, B., Filippi, S., Nost, T. H., Chiquet, J., & Chadeau-Hyam, M. (2021). Automated calibration for stability selection in penalised regression and graphical models. Retrieved from http://arxiv.org/abs/2106.02521

Buuren, S. van, & Groothuis-Oudshoorn, K. (2011). mice : Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*(3). https://doi.org/10.18637/jss.v045.i03

Christensen, A. P., & Golino, H. (2021). Estimating the Stability of Psychological Dimensions via Bootstrap Exploratory Graph Analysis: A Monte Carlo Simulation and Tutorial. *Psych*, *3*(3), 479–500. https://doi.org/10.3390/psych3030032

Csárdi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *Inter. J. Comp. Syst.*, *1695*, 1–9.

Drakvik, E., Altenburger, R., Aoki, Y., Backhaus, T., Bahadori, T., Barouki, R., … Bergman, Å. (2020). Statement on advancing the assessment of chemical mixtures and their risks for human health and the environment. *Environment International*, *134*, 105267. https://doi.org/10.1016/j.envint.2019.105267

EFSA, More, S. J., Bampidis, V., Benford, D., Bragard, C., Hernandez-Jerez, A., … Hogstrand, C. (2021). Guidance Document on Scientific criteria for grouping chemicals into assessment groups for human risk assessment of combined exposure to multiple chemicals. *EFSA Journal*, *19*(12). https://doi.org/10.2903/j.efsa.2021.7033

European Commission. (2012). *The combination effects of chemicals – Chemical mixtures*.

Fisher, M., Arbuckle, T. E., MacPherson, S., Braun, J. M., Feeley, M., & Gaudreau, É. (2019). Phthalate and BPA Exposure in Women and Newborns through Personal Care Product Use and Food Packaging. *Environmental Science & Technology*, *53*(18), 10813–10826. https://doi.org/10.1021/acs.est.9b02372

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432–441. https://doi.org/10.1093/biostatistics/kxm045

Golino, H., & Christensen, A. P. (2020). EGAnet: Exploratory Graph Analysis: A framework for estimating the number of dimensions in multivariate data using network psychometrics. Retrieved from https://cran.r-project.org/package=EGAnet

Golino, H., Moulder, R., Shi, D., Christensen, A. P., Garrido, L. E., Nieto, M. D., … Boker, S. M. (2021). Entropy Fit Indices: New Fit Measures for Assessing the Structure and Dimensionality of Multiple Latent Variables. *Multivariate Behavioral Research*, *56*(6), 874–902. https://doi.org/10.1080/00273171.2020.1779642

Govarts, E., Portengen, L., Lambrechts, N., Bruckers, L., Den Hond, E., Covaci, A., … Vermeulen, R. (2020). Early-life exposure to multiple persistent organic pollutants and metals and birth weight: Pooled analysis in four Flemish birth cohorts. *Environment International*, *145*, 106149. https://doi.org/10.1016/j.envint.2020.106149

Hoffmann, R., Lange, M., Butschalowsky, H., Houben, R., Schmich, P., Allen, J., … Gößwald, A. (2018). KiGGS Wave 2 cross-sectional study – participant acquisition, response rates and representativeness. *J Health Monit*, *3*(1), 79–91. https://doi.org/10.17886/RKI-GBE-2018-032

Horvath, S. (2011). *Weighted Network Analysis: Applications in Genomics and Systems Biology*. Springer New York: New York, NY, USA.

Kienzler, A., Bopp, S. K., van der Linden, S., Berggren, E., & Worth, A. (2016). Regulatory assessment of chemical mixtures: Requirements, current approaches and future perspectives. *Regulatory Toxicology and Pharmacology*, *80*, 321–334. https://doi.org/10.1016/j.yrtph.2016.05.020

Liu, H., Roeder, K., & Wasserman, L. (2010). Stability Approach to Regularization Selection (StARS) for High Dimensional Graphical Models. Retrieved from http://arxiv.org/abs/1006.3316

Llop, S., Ballester, F., Estarlich, M., Ibarluzea, J., Manrique, A., Rebagliato, M., … Iñiguez, C. (2008). Urinary 1-hydroxypyrene, air pollution exposure and associated life style factors in pregnant women. *Science of The Total Environment*, *407*(1), 97–104. https://doi.org/https://doi.org/10.1016/j.scitotenv.2008.07.070

Lubin, J. H., Colt, J. S., Camann, D., Davis, S., Cerhan, J. R., Severson, R. K., … Hartge, P. (2004). Epidemiologic Evaluation of Measurement Data in the Presence of Detection Limits. *Environmental Health Perspectives*, *112*(17), 1691–1696. https://doi.org/10.1289/ehp.7199

Mauz, E., Gößwald, A., Kamtsiuris, P., Hoffmann, R., Lange, M., von Schenck, U., … Hölling, H. (2017). New data for action. Data collection for KiGGS Wave 2 has been completed. *J. Health Monit.*, *2*. https://doi.org/https://doi.org/10.17886/RKI-GBE-2017-105

Murawski, A., Roth, A., Schwedler, G., Schmied-Tobies, M. I. H., Rucic, E., Pluym, N., … Kolossa-Gehring, M. (2020). Polycyclic aromatic hydrocarbons (PAH) in urine of children and adolescents in Germany – human biomonitoring results of the German Environmental Survey 2014–2017 (GerES V). *International Journal of Hygiene and Environmental Health*, *226*, 113491. https://doi.org/https://doi.org/10.1016/j.ijheh.2020.113491

Orman, G. K., & Labatut, V. (2009). A Comparison of Community Detection Algorithms on Artificial Networks (pp. 242–256). https://doi.org/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 . *Frontiers in Public Health* . Retrieved from https://www.frontiersin.org/article/10.3389/fpubh.2021.590038

Pérez-Gómez, B., Pastor-Barriuso, R., Cervantes-Amat, M., Esteban, M., Ruiz-Moraga, M., Aragonés, N., … BIOAMBIENT.ES, on behalf of. (2013). BIOAMBIENT.ES study protocol: rationale and design of a cross-sectional human biomonitoring survey in Spain. *Environmental Science and Pollution Research*, *20*(2), 1193–1202. https://doi.org/10.1007/s11356-012-1320-3

Pons, P., & Latapy, M. (2005). Computing Communities in Large Networks Using Random Walks (pp. 284–293). https://doi.org/10.1007/11569596_31

Řiháčková, K., Pindur, A., Komprdová, K., Pálešová, N., Kohoutek, J., Šenk, P., … Čupr, P. (2023). The exposure of Czech firefighters to perfluoroalkyl substances and polycyclic aromatic hydrocarbons: CELSPAC – FIREexpo case-control human biomonitoring study. *Science of The Total Environment*, *881*, 163298. https://doi.org/https://doi.org/10.1016/j.scitotenv.2023.163298

Rísová, V. (2019). The pathway of lead through the mother's body to the child. *Interdisciplinary Toxicology*, *12*(1), 1–6. https://doi.org/10.2478/intox-2019-0001

Rotter, S., Beronius, A., Boobis, A. R., Hanberg, A., van Klaveren, J., Luijten, M., … Solecki, R. (2018). Overview on legislation and scientific approaches for risk assessment of combined exposure to multiple chemicals: the potential EuroMix contribution. *Critical Reviews in Toxicology*, *48*(9), 796–814. https://doi.org/10.1080/10408444.2018.1541964

Schettler, T. (2006). Human exposure to phthalates via consumer products. *International Journal of Andrology*, *29*(1), 134–139. https://doi.org/10.1111/j.1365-2605.2005.00567.x

Schulz, C., Kolossa-Gehring, M., & Gies, A. (2017). German Environmental Survey for Children and Adolescents 2014-2017 (GerES V) – the environmental module of KiGGS Wave 2. Robert Koch-Institut, Epidemiologie und Gesundheitsberichterstattung. https://doi.org/10.17886/RKI-GBE-2017-108

Vahter, M. (2008). Health Effects of Early Life Exposure to Arsenic. *Basic & Clinical Pharmacology & Toxicology*, *102*(2), 204–211. https://doi.org/10.1111/j.1742-7843.2007.00168.x

Zhao, T., Liu, H., & Roeder, K. (2012). The huge Package for High-dimensional Undirected Graph Estimation in R. *The Journal of Machine Learning Research*, *13*, 1059–1062.

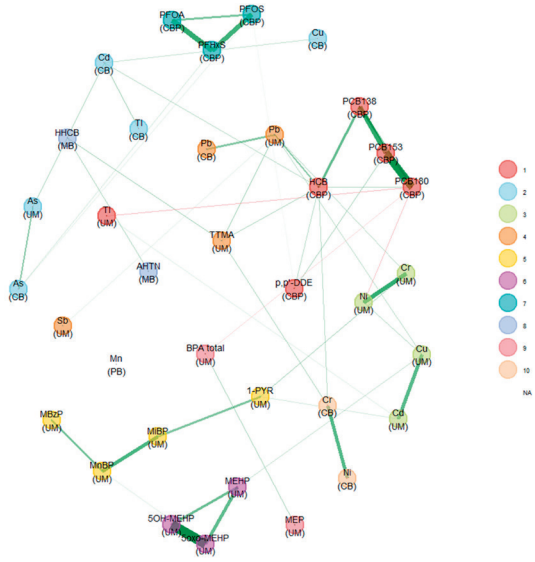


Figure A1. *Weighted network for 3xG using the graphical lasso method. Data were corrected for age, smoking, and BMI. Urinary markers were standardized for creatinine; lipid soluble blood markers were standardized for lipids. Matrices in which biomarkers are measured appear between brackets. Green lines represent a positive dependency between nodes (biomarkers) while red lines represent a negative dependency.*

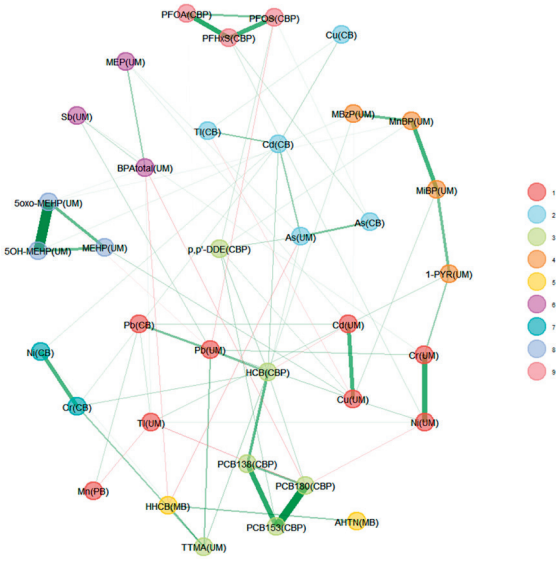


Figure A2. *Weighted network for 3xG using the graphical lasso method with bootstrap of 80 iterations. Data were corrected for age, smoking, and BMI. Urinary markers were standardised for creatinine; lipid soluble blood markers were standardised for lipids. Matrices in which biomarkers are measured appear between brackets. Green lines represent a positive dependency between nodes (biomarkers) while red lines represent a negative dependency.*

SUPPLEMENTARY MATERIAL

Appendix A

Comparison of the Two Unweighted and Weighted Network Estimation Approaches

Two methods to visualize weighted networks were explored with the R Package EGAnet (v1.2.3 (Golino et al., 2020)). For the first method, the EGA() function was applied to the correlation matrix of the data. This function estimates the number of dimensions of the correlation matrix using graphical lasso with extended Bayesian information criterion to select optimal regularization parameters. Figure A1 shows the resulting weighted network from this method.

The second method uses the function bootEGA() from the EGAnet R package which estimates the number of dimensions of n bootstraps using the empirical (partial) correlation matrix (parametric) or resampling from the empirical dataset (non-parametric). It also estimates a typical median network structure, which is formed by the median or mean pairwise (partial) correlations over the n bootstraps. Here, a parametric bootstrap (1000 iterations) was used to estimate the median network structure, which was then plotted as the final result (shown in Figure A2).

Networks obtained using the two different methods maintain the communities constituted by PFASs, PCBS with HCB and p,p'-DDE, DEHP metabolites, and DiBP metabolites with 1-PYR. The network of MEP and BPA total is conserved; however, Sb is included in the bootstrap network, while in the GLASSO network, it constitutes its own community. The composition of other, less strong communities seems to vary slightly between the two methods; nevertheless, the overall relationships do not seem to differ heavily between the two approaches.

Impact of Different Approaches to Correcting Biomarker Levels Against Creatinine Levels

The impact of different approaches for standardisation of creatinine (or lack thereof) in the network analyses was studied in the GerES V sample (Figures A3–A5) with networks as described above (using a parametric bootstrap of 1000 iterations). We distinguished between the following terms when taking into account dilution. 'Standardisation' means that each individual's raw concentration for the biomarkers studied is divided by its individual dilution level (e.g., creatinine). 'Adjustment' for dilution reflects that the dilution was included as a control variable into multivariate regression (see also (Horvath, 2011)). Finally, 'correction' is used as the general term of taking into account dilution levels as standardisation, adjustment, or the combination of both. To illustrate the effect of correction of urinary dilution with creatinine, Figure A3 shows the resulting communities when standardising raw concentrations for creatinine and adjusting for creatinine in multivariate analyses (recommended by HBM4EU). A total of eight communities containing three or more substances was observed. 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, and acrylamide (plum), selenium, chromium, antimony, and aprotic solvent HMSI (lavender), and

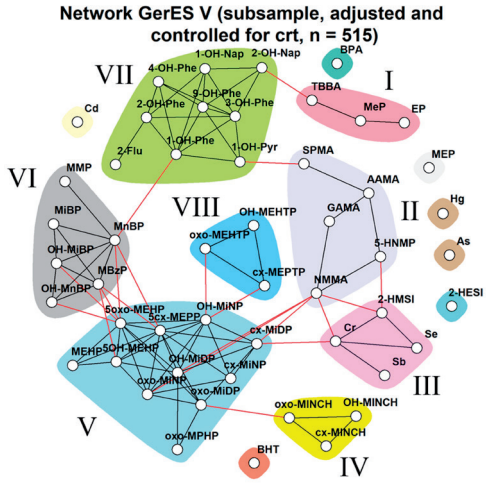


Figure A3. *Network of GerES V subset, using creatinine-standardised and creatinine-adjusted data. Black lines indicate dependency between nodes (biomarkers) within a community; red lines indicate dependency between nodes in different communities.*

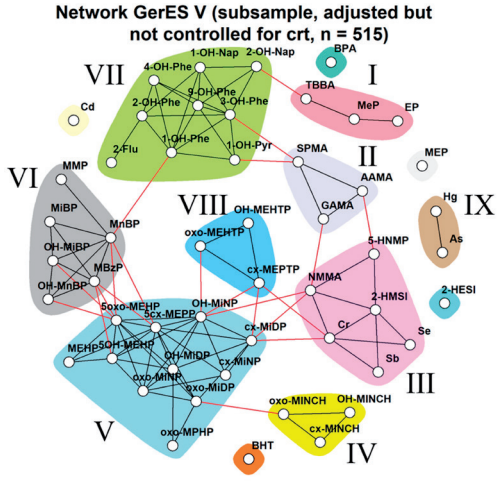


Figure A4. *Network of GerES V subset, using creatinine-standardised data. Black lines indicate dependency between nodes (biomarkers) within a community; red lines indicate dependency between nodes in different communities.*

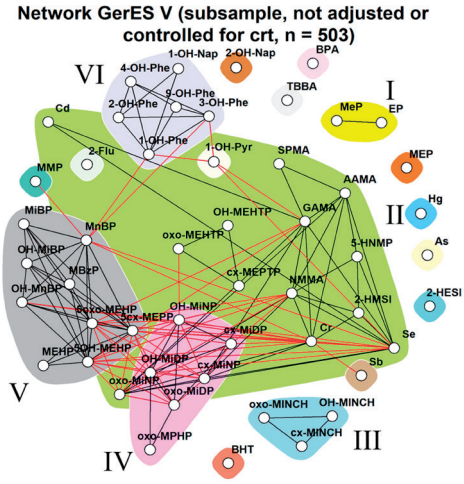


Figure A5. *Network of GerES V subset, without application of any correction for creatinine. Black lines indicate dependency between nodes (biomarkers) within a community; red lines indicate dependency between nodes in different communities.*

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). Several substances were not part of any community, such as some elements (mercury, arsenic, and cadmium), BPA, and the phthalate MEP.

The differences between the network obtained using creatinine-standardised and creatinine-adjusted data (Figure A3) versus using only creatinine-standardised data (Figure A4) are limited. The major differences include an additional community (comprising mercury and arsenic), and the split of the community that consists of acrylamide, aprotic solvents, selenium, chromium, antimony, NMMA, and SPMA into two different communities (plum and yellow in Figure A4). However, not correcting for creatinine in any form results in considerably different communities. As can be seen in Figure A5, three heavily inter-related communities (green, pink, and grey) mask the detailed communities detected in the network when standardising and adjusting for creatinine, 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 standardise the biomarker concentrations for this parameter; nevertheless, this needs to be confirmed by further studies. In conclusion, not correcting for dilution effects may create spurious results. The two methods for correction for dilution effects showed little difference.

Supplementary Table S1. *Descriptive statistics for biomarkers included in the network analysis for 3xG study. Urinary markers expressed as µg/g creatinine, blood values as µg/L.*

Substance group	Substance	Biomarker		% < LOQ	P25	P50	P75	P95
Elements	Cadmium	Cd	Mothers, morning urine (µg/L)	0%	0.15	0.21	0.36	0.65
	Cadmium	Cd	Newborns, cord blood (µg/L)	8%	0.02	0.03	0.04	0.05
	Chromium	Cr	Mothers, morning urine (µg/L)	1.6%	0.24	0.39	0.66	1.4
	Chromium	Cr	Newborns, cord blood (µg/L)	18%	0.14	0.24	0.57	1.3
	Antimony	Sb	Mothers, morning urine (µg/L)	18%	0.02	0.04	0.06	0.09
	Nickel	Ni	Mothers, morning urine (µg/L)	0%	1.3	1.7	2.6	4.7
	Nickel	Ni	Newborns, cord blood (µg/L)	37%	<LOD	0.12	0.2	1.2
	Copper	Cu	Mothers, morning urine (µg/L)	0%	9.3	12	17	29
	Copper	Cu	Newborns, cord blood (µg/L)	0%	540	576	626	698
	Arsenic	As	Mothers, morning urine (µg/L)	0%	0.8	1.2	2.1	4.5
	Arsenic	As	Newborns, cord blood (µg/L)	0%	0.5	0.89	1.6	3.3
	Lead	Pb	Mothers, morning urine (µg/L)	0%	0.51	0.76	1	1.7
	Lead	Pb	Newborns, cord blood (µg/L)	0%	5	6	7	10
	Manganese	Mn	Newborns, cord blood (µg/L)	0%	28	35	44	67
	Thallium	Tl	Mothers, morning urine (µg/L)	0%	0.11	0.18	0.26	0.36
	Thallium	Tl	Newborns, cord blood (µg/L)	0%	0.01	0.02	0.02	0.04
Phthalates	DEHP	MEHP	Mothers, morning urine (µg/L)	2.4%	1.3	1.9	3.8	8.5
		OH-MEHP	Mothers, morning urine (µg/L)	0%	5.3	8.2	13	34
		oxo-MEHP	Mothers, morning urine (µg/L)	0%	3.9	5.7	9.3	20
	BBzP	MBzP	Mothers, morning urine (µg/L)	0%	3	5.8	11	26
	DnBP	MnBP	Mothers, morning urine (µg/L)	0%	18	28	45	119
	DiBP	MiBP	Mothers, morning urine (µg/L)	0%	34	46	79	276
	DEP	MEP	Mothers, morning urine (µg/L)	0%	12	34	77	189
Bisphenols	Bisphenol A	BPA	Mothers, morning urine (µg/L)	2.4%	0.8	1.2	2.1	4.5
PCBs	PCB 138	PCB 138	Newborns, cord blood plasma (µg/L)	6.4%	0.02	0.03	0.03	0.05
	PCB 153	PCB 153	Newborns, cord blood plasma (µg/L)	1.6%	0.03	0.04	0.05	0.09
	PCB 180	PCB 180	Newborns, cord blood plasma (µg/L)	5.6%	0.02	0.02	0.04	0.07
Persistent organic pollutants	Dichlorodiphenyldi-chloroethylene	DDE	Newborns, cord blood plasma (µg/L)	0%	0.09	0.15	0.26	0.53
	Hexachlorobenzene	HCB	Newborns, cord blood plasma (µg/L)	24.8%	0.01	0.01	0.02	0.03
PFAS	PFOA	PFOA	Newborns, cord blood plasma (µg/L)	0%	0.77	1.1	1.4	2.1
	PFHxS	PFHxS	Newborns, cord blood plasma (µg/L)	20.8%	0.21	0.34	0.46	0.76
	PFOS	PFOS	Newborns, cord blood plasma (µg/L)	0%	1.1	1.6	2.2	3.6
Musks	Galaxolide	HHCB	Mothers, blood (µg/L)	0%	230	282	359	543
	Tonalide	AHTN	Mothers, blood (µg/L)	0.8%	50	63	81	186

Supplementary Table S2. *Descriptive statistics for biomarkers included in the network analysis for the CELSPAC – FIREexpo study. Urinary markers expressed as µg/g creatinine, blood values as µg/L.*

Study population	Substance group	Biomarker	% <LOQ	LOQ	P10	P25	P50	P75	P90
Firefighters	**Blood serum (ng/ml)**								
	PFASs	PFOA	0	0.07	0.66	0.92	1.2	1.5	1.9
		PFNA	0	0.004	0.23	0.29	0.40	0.54	0.63
		PFDA	0	0.004	0.10	0.14	0.19	0.25	0.30
		PFUnDA	12	0.012	<LOQ	0.04	0.05	0.07	0.10
		PFBS	27	0.04	<LOQ	<LOQ	0.15	0.18	0.25
		PFHxS	0	0.004	0.3	0.38	0.49	0.67	0.76
		PFHpS	3.9	0.005	0.04	0.06	0.08	0.10	0.14
		PFOS	0	0.03	1.7	2.4	3.2	4.8	6.4
	Morning urine (µg/g CRT)								
	PAHs	1-NAPH	0	0.006	0.52	0.99	1.6	2.6	3.6
		2-NAPH	0	0.006	1.4	2.9	4.2	6.2	9.5
		2-FLUO	0	0.006	0.14	0.19	0.26	0.32	0.48
		3-FLUO	5.8	0.006	0.02	0.04	0.06	0.11	0.16
		1-PHEN	38	0.006	<LOQ	<LOQ	0.02	0.04	0.10
		Σ(2-PHEN+3-PHEN)	0	0.006	0.07	0.10	0.14	0.20	0.30
		4-PHEN	1.9	0.006	0.04	0.21	0.32	0.50	0.59
		1-PYR	0	0.006	0.04	0.05	0.07	0.10	0.14
Control group	**Blood serum (ng/ml)**								
	PFASs	PFPeA	16	0.036	<LOQ	0.18	0.22	0.26	0.31
		PFHxA	7.3	0.04	0.05	0.07	0.08	0.10	0.11
		PFOA	1.8	0.07	0.49	0.69	0.90	1.1	1.4
		PFNA	0	0.004	0.18	0.23	0.30	0.36	0.41
		PFDA	0	0.004	0.08	0.11	0.12	0.17	0.23
		PFUnDA	3.6	0.012	0.03	0.05	0.07	0.10	0.11
		PFHxS	0	0.004	0.27	0.33	0.43	0.52	0.65
		PFHpS	36	0.005	<LOQ	<LOQ	0.04	0.07	0.09
		PFOS	0	0.03	1.1	1.7	2.2	2.7	3.5
	Morning urine (µg/g creatinine)								
	PAHs	1-NAPH	0	0.006	0.34	0.58	0.93	1.3	2.3
		2-NAPH	0	0.006	1.3	1.7	2.8	4.1	5.3
		2-FLUO	0	0.006	0.09	0.13	0.18	0.25	0.31
		3-FLUO	3.6	0.006	0.01	0.02	0.03	0.05	0.07
		1-PHEN	5.5	0.006	0.02	0.05	0.09	0.12	0.20
		Σ(2-PHEN+3-PHEN)	0	0.006	0.06	0.09	0.12	0.20	0.28
		4-PHEN	20	0.006	<LOQ	0.01	0.02	0.04	0.35
		1-PYR	0	0.006	0.02	0.03	0.04	0.06	0.09

Supplementary Table S3. *Descriptive statistics for biomarkers included in the network analysis for the GerEs V study. Urinary markers expressed as µg/g creatinine, blood values as µg/L.*

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	8.4
Aprotic solvents		HNMF	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
		5OH-MEHP	0	0 %	0.2	3.12	3.94	5.87	8.98	13.94	21.69	28.84	9.25
		5oxo-MEHP	0	0 %	0.2	1.96	2.52	4.08	6.42	10.49	15.89	21.63	6.47
		5cx-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
	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

Substance group	Substance	Biomarker	N < LOQ	% < LOQ	LOQ	P05	P10	P25	P50	P75	P90	P95	GM
	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

Supplementary Table S4. *Descriptive statistics for biomarkers included in the network analysis for the BIOAMBIENT.ES study. Urinary markers expressed as µg/g creatinine, blood values as µg/L.*

Substance group	Substance	Biomarker	N < LOQ	% < LOQ	LOQ	P05	P10	P25	P50	P75	P90	P95	GM
	Mercury	Hg	1	0.68	0.10	0.21	0.31	0.56	0.99	1.58	2.3	2.75	0.88
	Cadmium	Cg	4	2.45	0.05	0.05	0.07	0.12	0.2	0.38	0.59	0.72	0.2
Metals	Lead	Pb	4	2.6	0.10	0.19	0.29	0.43	0.7	1.04	1.67	2.36	0.69
	Thallium	Tl	16	11.35	0.05	0.05	0.06	0.08	0.11	0.16	0.21	0.26	0.11
	Cobalt	Co	2	1.46	0.05	0.2	0.23	0.36	0.57	0.84	1.28	2.2	0.58
Phtalates	DMP	MMP	8	4.91	1	1.12	1.46	2.01	2.69	4.2	7.08	10.56	3.14
	DEP	MEP	0	0	0.5	25.81	39.64	87.08	189.47	345.2	802.21	1307.09	189.2
	BBzP	MBzP	2	1.23	0.2	1.77	2.27	3.16	5.09	8.98	16.24	28.53	5.69
		MiBP	0	0	1	8.37	10.89	16.33	23.71	34.19	51.71	72.73	24.04
		OH-MiBP	0	0	0.25	3.71	4.77	6.5	9.17	14.19	19.14	25.01	9.58
		MnBP	1	0.61	1	4.76	6.04	9.63	14.74	22.23	33.4	41.53	14.54
	DnBP	OH-MnBP	4	2.45	0.25	0.56	0.72	1.08	1.71	2.44	3.52	5.04	1.65
		MCPP	22	13.5	0.50	0.42	0.51	0.71	0.95	1.52	2.2	2.9	1.07
		MEHP	6	3.68	0.50	1.08	1.43	2.47	4.09	6.63	11.32	14.9	3.97
		OH-MEHP	0	0	0.20	6.24	7.89	11.54	18.44	26.26	39.22	56.6	17.75
		oxo-MEHP	1	0.61	0.20	3.63	4.86	7.74	11.45	17.03	27.82	35.71	11.56
		cx-MEPP	0	0	0.20	7.24	8.29	12.82	18.88	28.15	43.43	54.27	19.19
		OH-MiNP	3	1.84	0.20	0.99	1.28	2.03	3.45	5.91	10.97	23.17	3.62
	DiNP	oxo-MiNP	5	3.07	0.20	0.50	0.71	1.19	2.08	3.62	6.71	14.93	2.2
		cx-MiNP	1	0.61	0.20	0.63	0.76	1.05	1.43	2.27	3.61	4.87	1.59
		OH-MiDP	3	1.84	0.20	0.7	0.9	1.23	1.76	2.83	4.09	5.11	1.84
	DiDP	oxo-MIDP	17	10.43	0.20	0.25	0.29	0.42	0.62	0.95	1.49	1.82	0.64
		cx-MIDP	1	0.61	0.20	0.63	0.76	1.05	1.43	2.27	3.61	4.87	1.59
DINCH	DINCH	OH-DINCH	8	4.91	0.05	0.1	0.15	0.29	0.7	1.9	6.81	19.82	0.85
		cx-MINCH	6	3.68	0.05	0.09	0.14	0.26	0.43	1.22	4.3	8.21	0.64
		oxo-DINCH	22	13.5	0.05	0.03	0.04	0.11	0.35	1.19	4.6	11.87	0.4
PFAS	PFHxS	PFHXS	34	20.86	0.34	0.24	0.24	0.39	0.68	1.13	1.99	2.39	0.70
	PFOA	PFOA	0	0	0.16	0.81	0.96	1.39	2.03	2.94	3.92	5.09	1.98
	PFOS	PFOS	2	1.23	0.33	2.48	3.53	5.30	8.09	11.02	15.18	17.11	7.25
	PFNA	PFNA	1	0.61	0.16	0.48	0.59	0.70	0.95	1.39	1.74	2.14	0.98
	PFDA	PFDA	18	11.04	0.2	0.14	0.14	0.26	0.37	0.53	0.75	0.84	0.37

Chapter 4

Characterization of multiple pesticides using silicone wristbands – An exposure assessment of residents living near agricultural fields in the Netherlands

I.B. Ottenbros[1, 2], R.C.H. Vermeulen[1], E.J.M. Krop[1], H. Beeltje[3,4], S. Fuhrimann[5], D.M. Figueiredo[1 *]

[1] Institute of Risk Assessment Sciences, Utrecht University, Yalelaan 2, 3584CM, Utrecht, the Netherlands
[2] Center for Sustainability, Environment and Health, National Institute for Public Health and the Environment (RIVM), Antonie van Leeuwenhoeklaan 9, 3721MA, Bilthoven, the Netherlands
[3] AQUON, De Blomboogerd 12, 4003BX Tiel, the Netherlands
[4] TNO Circular Economy & Environment, P.O. Box 80015, 3508, TA, Utrecht, the Netherlands
[5] Swiss Tropical and Public Health Institute, Basel, Switzerland

In preparation

ABSTRACT

Public health concerns exist regarding pesticide exposure of residents living near agricultural fields. Still, knowledge is limited in part due to the difficulties of assessing cumulative personal exposure to pesticides over time. Silicone wristbands are a low-cost and non-invasive passive sampling tool to assess exposure to multiple pesticides over time.

In this study, 19 residents living close to flower bulb fields in the Netherlands wore wristbands for an average of 60 days (range: 38-155). 31 different pesticides were quantified in the wristbands via liquid chromatography with tandem mass spectrometry (LC-MS/MS). Pesticides were categorized by application status: 1) applied during the course of the study, 2) registered for usage on flower bulbs but not applied, 3) not applied and not registered.

Measured concentrations reflected long-term, highly individualized exposure profiles. The minimum number of pesticides that were detected in a wristband was 6, with an average of 19 (maximum: 31). Azoxystrobin, carbendazim and pymetrozine were detected in all wristbands. While carbendazim was not authorized for agricultural spraying, it is an environmental degradation product of thiophanate-methyl (authorized at that time). No distinction could be made between days of wearing, vapor pressure and soil half-life of the pesticides.

Wristbands efficiently assessed pesticide mixture exposure profiles. The co-occurrence of pesticides in the wristbands allowed the identification of realistic chemical mixtures in residents living near agricultural fields. This study demonstrates the potential of wristbands to assess a large number of pesticides over a long period of time, informing future toxicology and exposure studies.

INTRODUCTION

Exposure characterization and quantification of pesticides among residents living close to agricultural fields remains of great interest, mostly due to a large amount of different compounds available on the market, but also due to their potential health effects (Dereumeaux et al., 2020). In the Netherlands, a recent study showed that homes close to flower bulb fields have increased pesticide concentrations in dust and air (Figueiredo et al., 2022; Oerlemans et al., 2021). These findings, however, are limited by the measurement period and don't necessarily reflect long-term personal exposure. Characterization of exposure through human biomonitoring is also limited, mainly due to personal variability, rapid excretion, metabolism and analytical challenges to reliably quantify a large variety of pesticides. On the other hand, silicone wristbands provide a possible solution to measure personal pesticide exposures, with the applicability to measure large numbers of pesticides over longer periods of time (Dixon et al., 2019; Doherty et al., 2021). Wristbands are passive samplers with low-impact and low-costs, and can be used for long-term personal exposure assessments in the general population (O'Connell et al., 2014). While most studies have a maximum wearing time of 30 days (Samon et al., 2022), for this study, the wristbands were worn for longer periods (average 60 days). Wristbands will continuously bind and sequester pesticides, providing a time-weighted average during the time of wearing (Dixon et al., 2020). This allows for detection of less often applied pesticides and provides aggregated cumulative exposure estimates. This study describes exposure to a wide range of pesticides measured by silicone wristbands, providing insight in the applicability of using wristbands for pesticide mixture exposure assessment.

MATERIAL AND METHODS

Sample collection

In the previously described OBO study (Figueiredo et al., 2021), exposure to pesticides among residents living close to flower bulb fields was measured via different sampling methods (a.o. active air sampling, biomonitoring/urine, collection of dust samples). Samples were collected through different measurement campaigns and linked to the actual spraying activities of the farmer.

Silicone wristband samplers were worn by a subgroup of the OBO participants to passively measure personal exposure between measurement campaigns. Subjects participated voluntarily. Wristbands were worn continuously by 20 participants (even during showering and sleeping) between two measurement campaigns (between non-spraying and spraying season) from January to October 2017, on average for 60 days. One participant was a farmer by profession and was excluded from further comparisons.

Pesticide analyses

Pesticide compounds were extracted from the silicone wristbands using acetonitrile (about 200 hours), the pesticide analyses were performed with LC-MS/MS. Exact analyses steps and

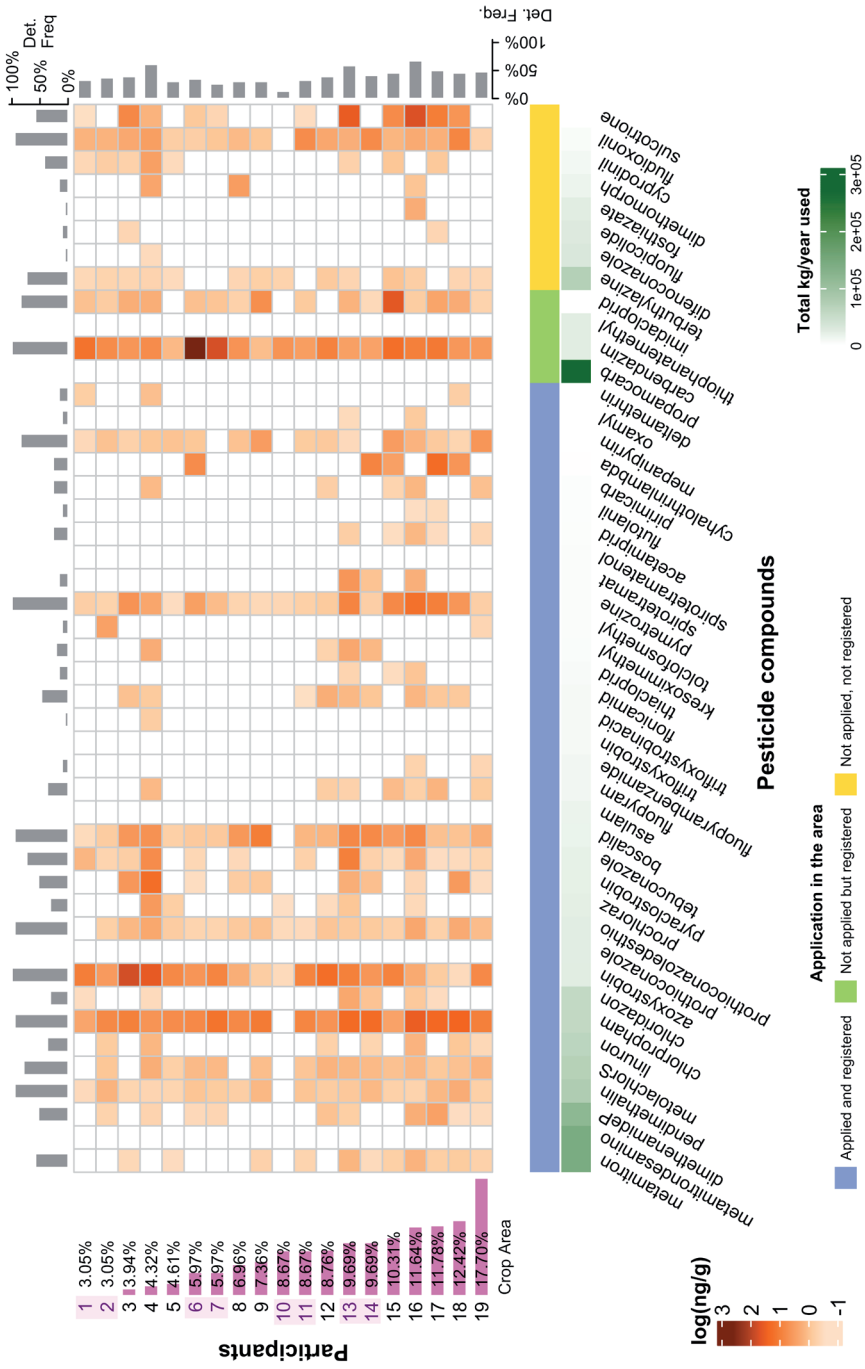


Figure 1. *Pesticide concentrations in the silicone wristbands presented as log10(ng/g), per participant (y-axis) and pesticide compound (x-axis). Non-detects are presented in white. Detection frequencies are presented along the right-x- and top-y-axis (grey bar plots). Participants are ordered by the percentage of crop area in a 250m buffer around the household (left-x-axis, pink bar plot), participants from the same household are indicated by pink brackets. Pesticides are ordered by actual application in the bulb field (applied, not applied but compound registered for use in bulb industry, or not applied not registered for use in bulb industry)), and by the total amount of pesticides used (kg/year 2016) in the Netherlands.*

details are provided in Supplementary Info 1. The detected concentrations ranged from 0.27-414 ng/g. The list of pesticide compounds analyzed (n=46) is the same as those measured in vacuum floor dust (VFD), described previously in Figueiredo et al., 2022.

Statistical analyses

The log10 transformed concentrations detected in the wristbands were presented in a heatmap (Gu et al., 2016), sorted by actual application of the pesticides in that area and the kg of pesticide/year used in 2016 in the Netherlands (Gu et al., 2016). The participants were sorted by the percentage of crops within a 250-meter buffer around their household.

Wristbands were compared to VFD samples from the same study, taken at the beginning and end of wearing the wristbands (Figueiredo et al., 2022). For 8 participants, both measurement types were collected, and possible trends were explored graphically.

For pesticides detected in at least 8 participants (42%), data below the limit of detection was imputed, using the same approach as in Figueiredo et al., 2022. Imputed data was only used to graphically explore patterns with days of wearing, vapor pressure, and chemical soil half-life.

All data analyses were performed using R, version 4.2.1.

RESULTS

A total of 19 participants wore the wristbands for an average of 60 days (range: 38-155). These participants came from 15 different households, which were located on average 117 (range: 24-236) meters from the nearest flower bulb field (Table S1). The majority (63%) was female, and the mean age was 58 years (range: 24-73).

The 46 different pesticides measured in the wristbands are listed in Figure 1 and Table S2. Figure 1 shows the log10-transformed concentrations of all pesticides per wristband/participant, sorted by application and kg used in 2016. The application of pesticides was divided into three categories: 1) applied on the OBO flower bulb fields during the course of the study, 2) registered for usage on flower bulbs but not applied during the course of the study, 3) not applied and not registered for use in flower bulb industry.

In each wristband at least 6 different pesticides were detected, with an average of 19 and maximum of 31 pesticides. In general, participants with a larger crop area around their house had more pesticides detected in their wristband, with the exception of ID 4. Consistencies between participants from the same household (IDs 1:2, 6:7, 10:11 and 13:14) were low with intraclass correlations below 0.2.

The pesticides azoxystrobin, carbendazim and pymetrozine were detected in all 19 wristbands. These three together with chlorpropham and sulcotrione had the highest concentrations detected. Carbendazim and sulcotrine were not applied on flower bulb field during the course of the study. Others frequently detected (>70%) but not reported to be sprayed were imidacloprid, fludioxonil, and terbuthylazine (plus the latter two were not registered for use

in bulb fields). Metamitron-desamino, prothioconazole, asulam, trifloxystrobin, spirotetramat-enol, propamocarb and thiophanate-methyl were not detected at all, although all of them were registered for use in the flower bulb industry (latter two not applied during the study).

The least frequently detected pesticides (of those reported to have been sprayed) were also used in smaller quantities per year (kg in the Netherlands), except for pymetrozine and mepanipyrim. While for the pesticides not registered for use, the most frequently detected were the ones applied in the smallest quantities in the Netherlands (kg/year).

A graphical comparison was made between VFD and wristband samples originating from the same household (maximum n=8, Figure S4). Overall, no clear pattern was seen between both measurement types, likely due to the low number of pairs.

The wristband concentrations versus the days of wearing did not show a clear pattern across all pesticides (Figure S1). Similarly, no clear pattern was seen in relation to the vapor pressure per pesticide and the soil half-life per pesticide (Figures S2-S3).

DISCUSSION AND CONCLUSION

Silicone wristbands were worn to detect a wide range of pesticides among individuals living near flower bulb fields in the Netherlands, providing a quantitative insight in the mixture of pesticides. While this study is limited by the number of wristbands, we were able to quantify a large set of pesticides. In total 31 out of 46 measured pesticides were detected.

Carbendazim was one of the pesticides, that although not applied, was frequently detected. Carbendazim was also frequently detected in dust, handwipes and urine samples from the OBO study (Figueiredo et al., 2022; Oerlemans et al., 2021). Carbendazim is the environmental degradation product from thiophanate-methyl, which could explain the measured concentrations, since thiophanate-methyl rapidly transforms and was not detected at all in the wristbands.

Like wristbands, VFD reflects more long-term exposure (Samon et al., 2022). Two pesticides were frequently detected in the wristbands, but not in VFD samples (Figueiredo et al., 2022): fludioxonil (fungicide) and terbuthylazine (herbicide). A possible hypothesis is exposure caused by field applications (long-range transport). When comparing wristbands with VFD data of the entire OBO study (Figueiredo et al., 2022), six pesticides were frequently detected in both matrices: pendimethalin, S-metalochlor, azoxystrobin, prothioconazole-desthio, tebuconazole, boscalid. A more in-depth exploration of differences between the exposure matrices would require more samples per household.

This study describes the exposure profile over a long-period of time (average of 60 days), in between seasons (spraying and non-spraying) of the OBO study (Figueiredo et al., 2021). This

measure provides an aggregated exposure level and cannot distinguish on which day the actual exposure took place. In literature wearing periods of maximum one month (typically a few days) have been reported (Samon et al., 2022). These differences in wearing time would require further analyses per compound (e.g. when equilibrium was reached). The variation between household members in our study was high, however due to a lack of statistical power we could not explore the drivers of these differences.

Other wristband exposure studies also detected highly individual exposure profiles (Dixon et al., 2019; Fuhrimann et al., 2022). The longer period of wearing in our study could have resulted in more pesticides per wristband (at least 6 pesticides detected in 100% of the samples), compared to another study in agricultural areas in South Africa (at least 2, detected in 92%) (Fuhrimann et al., 2022). Here, wristbands were worn for six days; the most frequently detected overlapping pesticides between both studies were deltamethrin and boscalid. Differences might occur due to differences in crops and agricultural practices. In Belgium, boscalid was detected at lower frequency compared to our study (7% versus 95%), possibly due to the shorter exposure time (5-days) and higher LOD (Aerts et al., 2018).

In conclusion, wristbands can be used to efficiently (low impact, and easy to implement) assess exposure to pesticide mixtures, capturing both inhalation and dermal individual exposures. The co-occurrence of pesticides in the wristband allows the identification of realistic chemical mixtures, in personal samples over longer periods of time. Our example showed the applicability to capture a large number of pesticides in a quantitative manner using personal samples, informing future toxicology and exposure studies.

Contributors

IO - Analysis, Writing, visualization; **RV** - Conceptualization, funding acquisition; **EK** - Investigation, Data curation; **HB** - Investigation, Methodology; **SF** - Validation; **DF** - Writing, Conceptualization, Supervision. All authors reviewed and approved the manuscript.

Acknowledgment

The authors thank all members of the OBO consortium. We also thank all the field workers that were involved in the collection of the data used in this manuscript, as well as the TNO laboratory personnel that analyzed the samples. We are grateful for the assistance of the National Institute for Public Health and Environment during all phases of the study.

Funding

This work was conducted within the OBO Project (Dutch acronym for "Research on Exposure of residents to pesticides"), funded by the Dutch ministry of Infrastructure and Water Management and the ministry of Economic Affairs and Climate Policy. This work was commissioned by the Dutch National Institute for Public Health and the Environment (RIVM).

Competing interests
None declared.

Data availability statement
We cordially invite other researchers to propose noncommercial research based on the available data in OBO or requests for additional chemical analyses with associated funding. Any such requests can be submitted to exposome.office@uu.nl with subject: OBO-research.

REFERENCES

Aerts, R., Joly, L., Szternfeld, P., Tsilikas, K., De Cremer, K., Castelain, P., … Van Nieuwenhuyse, A. (2018). Silicone Wristband Passive Samplers Yield Highly Individualized Pesticide Residue Exposure Profiles. *Environmental Science & Technology*, *52*(1), 298–307. https://doi.org/10.1021/acs.est.7b05039

Dereumeaux, C., Fillol, C., Quenel, P., & Denys, S. (2020). Pesticide exposures for residents living close to agricultural lands: A review. *Environment International*, *134*, 105210. https://doi.org/10.1016/j.envint.2019.105210

Dixon, H. M., Armstrong, G., Barton, M., Bergmann, A. J., Bondy, M., Halbleib, M. L., … Anderson, K. A. (2019). Discovery of common chemical exposures across three continents using silicone wristbands. *Royal Society Open Science*, *6*(2), 181836. https://doi.org/10.1098/rsos.181836

Dixon, H. M., Poutasse, C. M., & Anderson, K. A. (2020). Silicone Wristbands and Wearables to Assess Chemical Exposures. In *Total Exposure Health* (pp. 139–160). CRC Press. https://doi.org/10.1201/9780429263286-11

Doherty, B. T., Koelmel, J. P., Lin, E. Z., Romano, M. E., & Godri Pollitt, K. J. (2021). Use of Exposomic Methods Incorporating Sensors in Environmental Epidemiology. *Current Environmental Health Reports*, *8*(1), 34–41. https://doi.org/10.1007/s40572-021-00306-8

Figueiredo, D. M., Krop, E. J. M., Duyzer, J., Gerritsen-Ebben, R. M., Gooijer, Y. M., Holterman, H. J., … Vermeulen, R. C. H. (2021). Pesticide Exposure of Residents Living Close to Agricultural Fields in the Netherlands: Protocol for an Observational Study. *JMIR Research Protocols*, *10*(4), e27883. https://doi.org/10.2196/27883

Figueiredo, D. M., Nijssen, R., Krop, E. J. M., Buijtenhuijs, D., Gooijer, Y., Lageschaar, L., … Vermeulen, R. C. H. (2022). Pesticides in doormat and floor dust from homes close to treated fields: Spatio-temporal variance and determinants of occurrence and concentrations. *Environmental Pollution*, *301*, 119024. https://doi.org/10.1016/j.envpol.2022.119024

Fuhrimann, S., Mol, H. G. J., Dias, J., Dalvie, M. A., Röösli, M., Degrendele, C., … Vermeulen, R. (2022). Quantitative assessment of multiple pesticides in silicone wristbands of children/guardian pairs living in agricultural areas in South Africa. *Science of The Total Environment*, *812*, 152330. https://doi.org/10.1016/j.scitotenv.2021.152330

Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, *32*(18), 2847–2849. https://doi.org/10.1093/bioinformatics/btw313

O'Connell, S. G., Kincl, L. D., & Anderson, K. A. (2014). Silicone Wristbands as Personal Passive Samplers. *Environmental Science & Technology*, *48*(6), 3327–3335. https://doi.org/10.1021/es405022f

Oerlemans, A., Figueiredo, D. M., Mol, J. G. J., Nijssen, R., Anzion, R. B. M., van Dael, M. F. P., … Scheepers, P. T. J. (2021). Personal exposure assessment of pesticides in residents: The association between hand wipes and urinary biomarkers. *Environmental Research*, *199*, 111282. https://doi.org/10.1016/j.envres.2021.111282

Samon, S. M., Hammel, S. C., Stapleton, H. M., & Anderson, K. A. (2022). Silicone wristbands as personal passive sampling devices: Current knowledge, recommendations for use, and future directions. *Environment International*, *169*, 107339. https://doi.org/10.1016/j.envint.2022.107339

SUPPLEMENTARY MATERIAL

Supplementary Information I1. Description of analytical steps

Extraction:
Add 75 ml of acetonitrile(ACN) Place on shaker (Gerhardt: type RO 500) for 96 hours (120 rpm)
Transfer the ACN to a flask (250 ml) and add 75 ml of acetonitrile to the silicone rubber.
Let this shake for another 96 hours (120 rpm) (almost 200 hours to extract)
Combine the ACN extracts in the round bottom and rinse the silicone rubber with approximately 25 ml of ACN.
Evaporate the extract back to ca 1-2 ml with a rotary evaporator. (117 mbar, 45˚C)
Add 100 ml Hexane and concentrate to 1 to 2 ml. Transfer the extract to a 15 ml conical point tube and rinse the round bottom 3x with approx. 2 ml hexane.
Evaporate under nitrogen the extract back to < 1000 µl, transfer to an injection vial and make up to exactly 1000 µl.

GPC clean-up:
The 250 µl extract was measured with an Agilent 100 series HPLC system with GPC column (PLgel 5µm, 300 x 7.5 mm). Mobile phases consisted of DCM (100%, A), and Hexane/MTBE (1:1, B).
The chromatographic separation was performed at a flow rate of 2 mL/min, the collected fraction was between 6.5 and 20 minutes. Concentrate under nitrogen
Convert the extract to MeOH and evaporate the extract back to < 250 µl, transfer to an injection vial and make up to exactly 250 µl. For LCMS analyses take 50 µl of extract and dilute it with 50 µl of ultra-pure water.

Chemical LCMS analyses Pesticide:
The analytes and internal standards were detected by an Agilent 1260 series high-performance liquid chromatographer using a 100 × 2.1 mm, 2.6 µm Kinetex column (Phenomenex, Utrecht, the Netherlands) coupled with an Agilent 6460 triple quadrupole LC/MS with Jetstream Electron Spray Ionisation (ESI) and multiple reaction monitoring (MRM). A sample volume of 5 µL was injected with a column temperature of 60 °C and a flow rate of 200 µL min–1. The sample was eluted with a gradient of methanol (eluent A) and 1 mM ammonium fluoride with 0.01% acetic acid in Milli-Q water (eluent B) with flow rates of 0.5 mL min–1. Eluent A was increased from 5% to 90% in 10 min and maintained for 3 min. After this it is decreased to 5% in 0.1 min and maintained for 1.9 min to complete the cycle of 15 min. Mass spectrometry was performed with a gas temperature of 350 °C and a flow rate of 7 L min–1. Sheath gas temperature was set at 350 °C with a flow rate of 12 L min–1. The capillary voltage was set at 3500 V.

The target compounds were determined with two transitions. Calibration was done before measuring the samples with known amounts of the analytes in 6 steps with concentrations ranging between 0.1 and 20 ng mL–1.
The method resulted in values for LOQ ranging between 0.8 – 8 ng/g. (LOD = 0.3 - 2.7 ng/g).

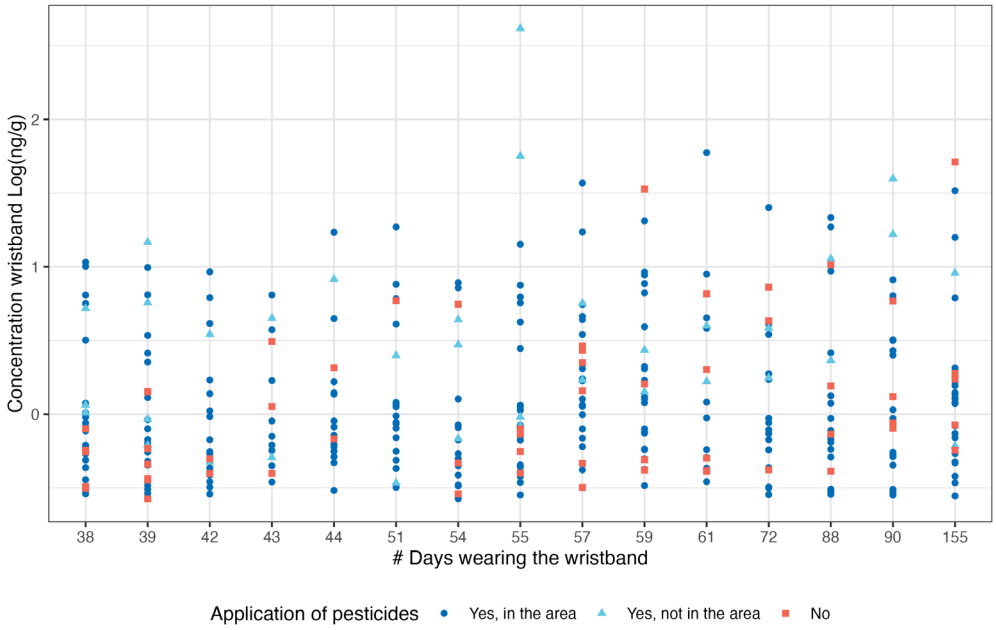
Supplementary Table S1. *Main characteristics study population. subgroup of OBO study*

	Adults (n=19)
Households with 2 participants	4
Gender, female (%)	63
Age, mean (years; min-max)	58 (24 – 73)
BMI, mean	24.7
Self-reported usage of pesticide products at home, 6 months prior to the campaign. % yes	47
Distance to closest active bulb field, mean (meter, min-max)	117 (24 – 236)
Days wearing the wristband, mean (days, min-max)	60 (38 – 155)

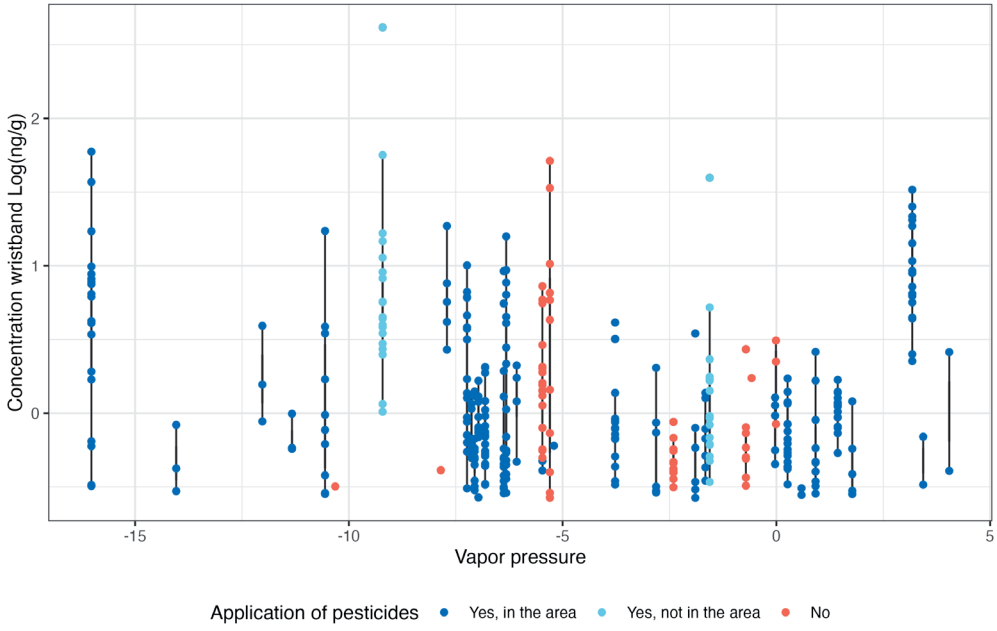
Supplementary Table S2. *Detection frequencies of the detected pesticides in wristbands and vacuum floor dust samples. combined with pesticide specific properties (vapor pressure, half-life, quantity applied, application in the flower bulb fields). Pesticides are ordered by application and quantity applied.*

#	Pesticide name	Parent pesticide or metabolite	Pesticide Type	Wristband	Vacuum Floor Dust (VFD) period 1	Vacuum Floor Dust (VFD) period 2	Vapor pressure (mPa at 25C)	Environmental half-life (Days)	Quantity applied (kg/ applied in 2016)	Application
				ng/g (%detect)	ng/g (%detect)	ng/g (%detect)				
1	Metamitron	Parent	Herbicide	0.8 (58%)	30.2 (11%)	72.7 (26%)	8.6E-04	11.1	144607	Applied
2	Metamitrondesamino	Metabolite	Herbicide	(0%)	39.1 (11%)	25.6 (32%)	4.5E-04	31.1	144607	Applied
3	DimethamidP	Parent	Herbicide	0.8 (53%)	2.2 (5%)	4.6 (11%)	2.5	7	120109	Applied
4	Pendimethalin	Parent	Herbicide	0.9 (95%)	33.2 (42%)	71.2 (32%)	1.3	100.6	77881	Applied
5	metolachlorS	Parent	Herbicide	1.0 (79%)	6.4 (26%)	6.8 (26%)	4.2	21	70394	Applied
6	Linuron	Parent	Herbicide	0.9 (37%)	12.0 (16%)	5.1 (21%)	0.19	48	64693	Applied
7	Chlorpropham	Parent	Herbicide	2.7 (95%)	81.8 (26%)	(0%)	24	13.1	57995	Applied
8	Chloridazon	Parent	Herbicide	0.8 (32%)	48.3 (5%)	37.9 (26%)	6.0E-02	34.7	56829	Applied
9	Azoxystrobin	Parent	Fungicide	2.2 (100%)	382 (42%)	4.5 (32%)	1.1E-07	180.7	24472	Applied
10	Prothioconazoledesthio	Metabolite	Fungicide	1.0 (95%)	12.0 (42%)	11.2 (32%)	1.1E-03	42.2	23858	Applied
11	Prothioconazol	Parent	Fungicide	(0%)	(0%)	(0%)	4.5E-09	0.77	23858	Applied
12	Prochloraz	Parent	Fungicide	0.9 (32%)	28.5 (42%)	30.4 (32%)	0.15	20	22054	Applied
13	Pyraclostrobin	Parent	Fungicide	1.3 (53%)	37.2 (42%)	61.7 (32%)	2.6E-05	33.3	20996	Applied
14	Tebuconazole	Parent	Fungicide	1.0 (74%)	9.3 (42%)	35.0 (32%)	1.7E-03	47.1	20344	Applied
15	Boscalid	Parent	Fungicide	1.4 (95%)	37.1 (42%)	63.1 (32%)	7.2E-04	484	17373	Applied
16	Asulam	Parent	Herbicide	0	24.0 (5%)	6.9 (26%)	0.19	9	16429	Applied
17	Fluopyram	Metabolite	Fungicide	0.9 (37%)	7.6 (42%)	9.7 (32%)	4.2E-03	118.8	11365	Applied
18	Fluopyrambenzamide	Parent	Fungicide	0.7 (11%)	(0%)	(0%)	4.2E-03	8.6	11365	Applied
19	Trifloxystrobinacid	Metabolite	Fungicide	(0%)	(0%)	(0%)	5.5E-03	70	7838	Applied
20	Trifloxystrobin	Parent	Fungicide	0.8 (5%)	(0%)	6.8 (16%)	3.4E-03	1.69	7838	Applied
21	Flonicamid	Parent	Insecticide	0.9 (47%)	7.9 (42%)	6.4 (32%)	9.4E-04	3.1	7672	Applied
22	Thiacloprid	Parent	Insecticide	0.7 (16%)	0	1.1 (5%)	8.0E-07	8.1	6499	Applied
23	Kresoximmethyl	Parent	Fungicide	1.1 (21%)	19.2 (11%)	29.6 (11%)	2.3E-03	1	5444	Applied
24	Toclofosmethyl	Parent	Fungicide	1.1 (11%)	10.4 26%)	1.7 (5%)	57	7.6	4801	Applied
25	Pymetrozine	Parent	Insecticide	1.3 (100%)	1.4 (11%)	7.7 (11%)	1.8E-03	22.6	4758	Applied
26	Spirotetramat	Parent	Insecticide	1.3 (16%)	(0%)	(0%)	6.0E-06	0.7	4303	Applied
27	Spirotetranatenol	Metabolite	Insecticide	(0%)	(0%)	(0%)	6.0E-06	1.9	4303	Applied

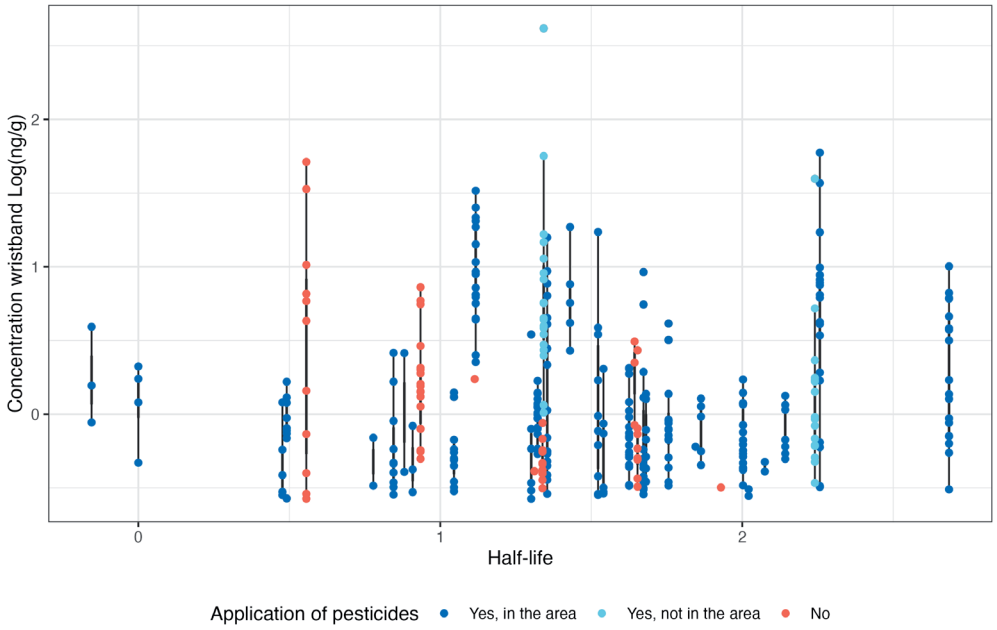
#	Pesticide name	Parent pesticide or metabolite	Pesticide Type	Wristband ng/g (%detect)	Vacuum Floor Dust (VFD) period 1 ng/g (%detect)	Vacuum Floor Dust (VFD) period 2 ng/g (%detect)	Vapor pressure (mPa at 25C)	Environ-mental half-life (Days)	Quantity applied (kg/ applied in 2016)	Application
28	Acetamiprid	Parent	Insecticide	0.7 (26%)	19.2 (11%)	7.4 (5%)	5.9	3	3849	Applied
29	Flutolanil	Parent	Fungicide	0.6 (11%)	2.1 (16%)	14.6 (11%)	1.8	105	3080	Applied
30	Primicarb	Parent	Insecticide	0.9 (26%)	(0%)	(0%)	0.97	73	2986	Applied
31	Cyhalotrinlambda	Parent	Insecticide	2.3 (26%)	(0%)	(0%)	4.5E-04	26.9	1536	Applied
32	Mepanipyrim	Parent	Fungicide	1.0 (84%)	112 (5%)	74.4 (16%)	2.3E-02	57	1117	Applied
33	Oxamyl	Parent	Insecticide	0.7 (11%)	(0%)	1.9 (11%)	31	6	959	Applied
34	Deltamethrin	Parent	Insecticide	0.9 (13%)	(0%)	(0%)	1.2E-05	21	618	Applied
35	Propamocarb	Parent	Fungicide	(0%)	0.8 (42%)	1.8 (32%)	7.3E+03	20	272107	Not applied, but registered
36	Carbendazim	Metabolite	Fungicide	2.9 (100%)	115 (42%)	98.1 (32%)	1.0E-04	22	24418	Not applied, but registered
37	Thiophanatemethyl	Parent	Fungicide	(0%)	14.2 (42%)	31.0 (32%)	9.5E-03	2	24418	Not applied, but registered
38	Imidacloprid	Parent	Insecticide	1.3 (84%)	58.3 (42%)	182 (32%)	0.21	174	588	Not applied, but registered
39	Terbuthylazine	Parent	Herbicide	0.7 (74%)	(0%)	(0%)	0.09	21.8	72199	Not Applied, not registered
40	Difenoconazol	Parent	Fungicide	0.6 (5%)	21.5 (21%)	3.8 (11%)	3.3E-05	85	32061	Not Applied, not registered
41	Fluopicolide	Parent	Fungicide	0.7 (11%)	(0%)	2.2 (11%)	8.0E-04	138.8	28252	Not Applied, not registered
42	Fosthiazate	Parent	Insecticide	1.3 (5%)	(0%)	(0%)	0.56	13	24385	Not Applied, not registered
43	Dimethomorph	Parent	Fungicide	1.3 (16%)	195 (26%)	4.3 (11%)	0.99	44	15002	Not Applied, not registered
44	Cyprodinil	Parent	Fungicide	0.9 (42%)	6.5 (11%)	11.8 (5%)	0.49	45	10237	Not Applied, not registered
45	Fludioxonil	Parent	Fungicide	1.3 (95%)	14.3 (42%)	13.4 (32%)	3.9E-04	20.5	5878	Not Applied, not registered



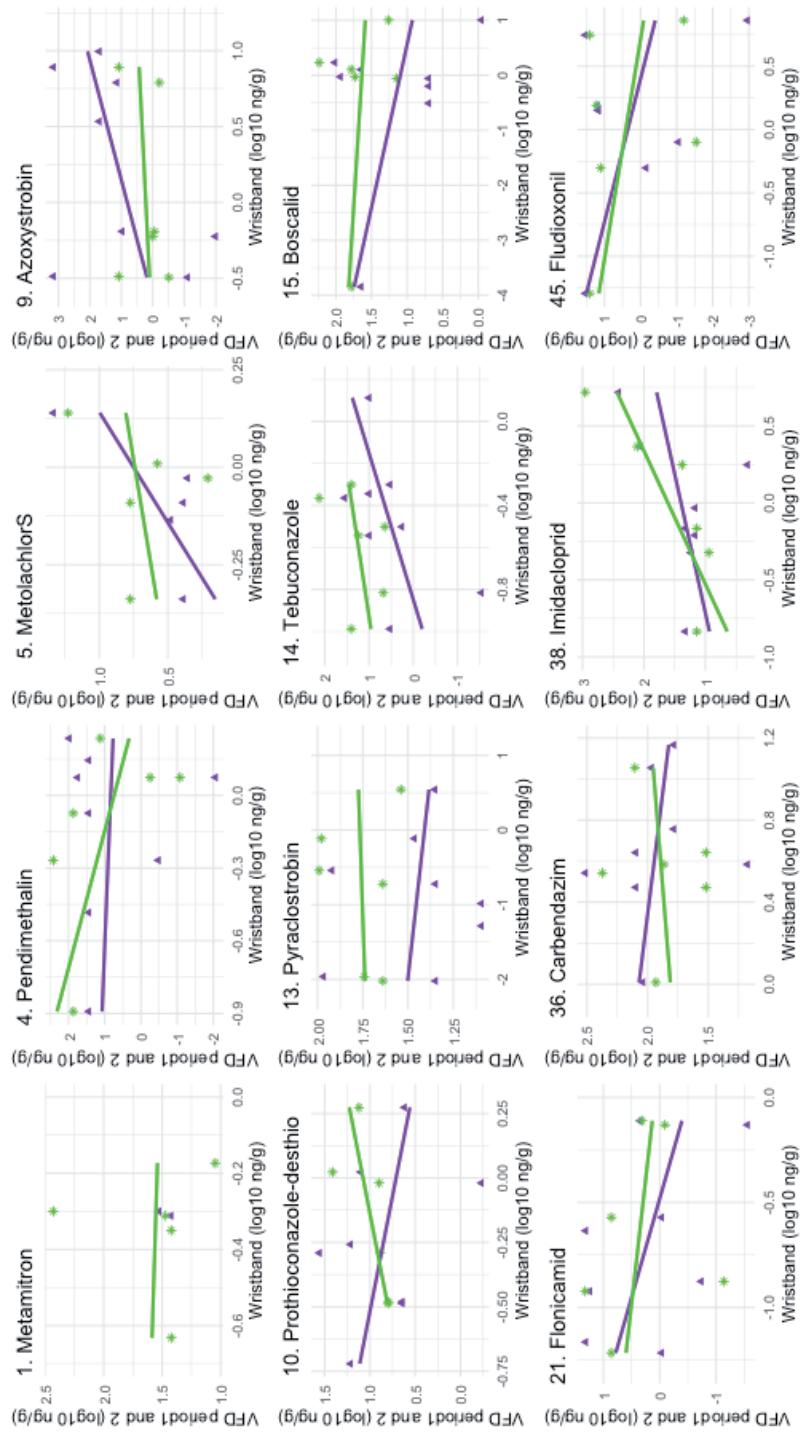
Supplementary Figure S1. *Scatter plots between the concentration of wristbands vs days wearing. distinction is made between the pesticides which were used in the vicinity (flower bulb fields) and which were not.*



Supplementary Figure S2. *Scatter plot between the concentration of wristbands versus vapor pressure.*



Supplementary Figure S3. *Scatter plot between the concentration of wristbands versus soil half-life.*



Supplementary Figure S4. *Scatterplots of the wristband pesticide concentrations (x-axis) versus the VFD measures (y-axis), differentiated by period 1 (prior to wearing the wristband, indicated by a purple triangle) and period 2 (after wearing the wristband, indicated by a green star). Only plots for pesticides with at least 5 overlapping pairs of measurements are shown. Numbers before each pesticide match with Table S2.*

Chapter 5

ASSESSMENT OF EXPOSURE TO PESTICIDE MIXTURES IN FIVE EUROPEAN COUNTRIES BY A HARMONIZED URINARY SUSPECT SCREENING APPROACH

Ilse Ottenbros[a,b], Erik Lebret[a], Carolin Huber[d,e], Arjen Lommen[f], Jean-Philippe Antignac[g], Pavel Čupr[h], Libor Šulc[h], Ondřej Mikeš[h], Tamás Szigeti[i], Szilvia Középesy[i], Inese Martinsone[j], Zanna Martinsone[j], Lasma Akulova[j], Olga Pardo[k,p,'], Sandra F. Fernández[k], Clara Coscollá[k], Susana Pedraza-Diaz[l], Martin Krauss[d], Laurent Debrauwer[m,n], Kévin Wagner[m,n], Rosalie Nijssen[f], Hans Mol[f], Chiara Maria Vitale[h], Jana Klanova[h], Borja Garlito Molina[o], Nuria León[k], Roel Vermeulen[a], Mirjam Luijten[c], Jelle Vlaanderen[a]

[a]Institute of Risk Assessment Sciences, Utrecht University, Yalelaan 2, 3584CM Utrecht, the Netherlands
[b]Center for Sustainability, Environment and Health, National Institute for Public Health and the Environment (RIVM), Antonie van Leeuwenhoeklaan 9, 3721MA Bilthoven, the Netherlands
[c]Center for Health Protection, National Institute for Public Health and the Environment (RIVM), Antonie van Leeuwenhoeklaan 9, 3721MA Bilthoven, the Netherlands
[d]Department of Effect-Directed Analysis, Helmholtz Centre for Environmental Research (UFZ), Permoserstr. 15, 04318 Leipzig, Germany
[e]Institute for Ecology, Diversity and Evolution, Goethe University Frankfurt Biologicum, Campus Riedberg, Max-von-Laue-Str. 13, 60438 Frankfurt am Mein, Germany
[f]Wageningen Food Safety Research, part of Wageningen University and Research, Akkermaalsbos 2, 6708WB Wageningen, the Netherlands
[g]Oniris, INRAE, LABERCA, 44300 Nantes, France
[h]RECETOX, Faculty of Science, Masaryk University, Kotlarska 2, Brno 61137 Czech Republic
[i]National Public Health Center, Albert Flórián út 2-6, 1097 Budapest, Hungary
[j]Laboratory of Hygiene and Occupational Diseases, Rīga Stradiņš University, Latvia
[k]Foundation for the Promotion of Health and Biomedical Research of the Valencia Region, Av. Catalunya 21, Valencia, Spain
[l]National Centre for Environmental Health, Instituto de Salud Carlos III (ISCIII), Majadahonda, Spain
[m]Toxalim (Research Center in Food Toxicology), Toulouse University, INRAE UMR 1331, ENVT, INP-Purpan, Paul Sabatier University, 31027 Toulouse, France
[n]Metatoul-AXIOM platform, National Infrastructure for Metabolomics and Fluxomics: MetaboHUB, Toulouse, France
[o]Environmental and Public Health Analytical Chemistry, Research Institute for Pesticides and Water (IUPA), Universitat Jaumel, Av. Sos Baynat S/N, 12071, Castelló de la Plana, Spain
[p]Present Address: Department of Analytical Chemistry, University of Valencia, Doctor Moliner 50, 46100 Burjassot, Spain
[']Present Address: Public Health Laboratory of Valencia, Av. Catalunya 21, Valencia 46020, Spain

Published in Journal of Hygiene and Environmental Health, March 2023
DOI: 10.1016/j.ijheh.2022.114105

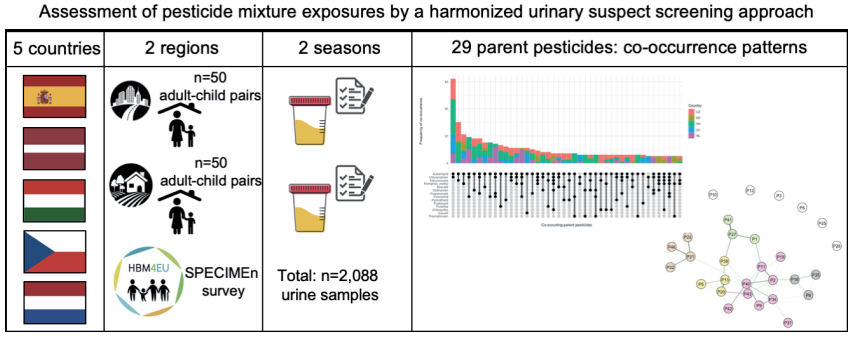
ABSTRACT

Humans are exposed to a mixture of pesticides through diet as well as through the environment. We conducted a suspect-screening based study to describe the probability of (concomitant) exposure to a set of pesticide profiles in five European countries (Latvia, Hungary, Czech Republic, Spain and the Netherlands). We explored whether living in an agricultural area (compared to living in a peri-urban area), being a child (compared to being an adult), and the season in which the urine sample was collected had an impact on the probability of detection of pesticides (-metabolites).

In total 2,088 urine samples were collected from 1,050 participants (525 parent-child pairs) and analyzed through harmonized suspect screening by five different laboratories. Fourty pesticide biomarkers (either pesticide metabolites or the parent pesticides as such) relating to 29 pesticides were identified at high levels of confidence in samples across all study sites. Most frequently detected were biomarkers related to the parent pesticides acetamiprid and chlorpropham. Other biomarkers with high detection rates in at least four countries related to the parent pesticides boscalid, fludioxonil, pirimiphos-methyl, pyrimethanil, clothianidin, fluazifop and propamocarb. In 84% of the samples at least two different pesticides were detected. The median number of detected pesticides in the urine samples was 3, and the maximum was 13 pesticides detected in a single sample. The most frequently co-occurring substances were acetamiprid with chlorpropham (in 62 urine samples), and acetamiprid with tebuconazole (30 samples). Some variation in the probability of detection of pesticides (-metabolites) was observed with living in an agricultural area or season of urine sampling, though no consistent patterns were observed. We did observe differences in the probability of detection of a pesticide (metabolite) among children compared to adults, suggesting a different exposure and/or elimination patterns between adults and children.

This survey demonstrates the feasibility of conducting a harmonized pan-European sample collection, combined with suspect screening to provide insight in the presence of exposure to pesticide mixtures in the European population, including agricultural areas. Future improvements could come from improved (harmonized) quantification of pesticide levels.

Graphical Abstract



INTRODUCTION

Humans are typically exposed to pesticides through multiple sources, including diet, occupational or environmental exposures (Damalas et al., 2011; Deziel et al., 2015). Growing evidence indicates that living in an agricultural area where pesticides are applied contributes to higher exposure than residents living away from agricultural areas (Dereumeaux et al., 2020; Figueiredo et al., 2021; Teysseire et al., 2020, 2021). Determinants contributing to this increased exposure include proximity to agricultural fields where pesticides are applied, crop acreage around the home, and season (Dereumeaux et al., 2020; Teysseire et al., 2021). Pesticide exposure has been linked to various short-term and chronic health effects such as respiratory or neurological development issues (Kim et al., 2017; Ntzani et al., 2013). Therefore a comprehensive characterization of the exposure to real-life mixtures of pesticides, which includes the contribution of living close to agricultural areas where pesticides are applied, is essential for human health risk assessment.

Most non-occupational pesticide exposure studies focus on selected sets of targeted pesticides for human biomonitoring (HBM), often based on *a priori* selected biomarkers related to e.g. the spraying activities in a certain area, the health outcome of interest, or practical considerations such as the commercial availability of standards (Dereumeaux et al., 2020; Teysseire et al., 2021). Currently, HBM for urinary pesticide biomarkers by targeted methods is limited to mostly pyrethroids and non-specific markers of organophosphorus pesticides. However, in real-life pesticide exposure often is already a mixture of multiple co-occurring compounds with repeated exposure timeframes (Crépet et al., 2019). With more than 450 active pesticides currently approved (plus 50 more currently pending) for use in the European Union (EU Database Pest, 2022), there is a growing need for information on the co-occurrence of these compounds in the human body. HBM of pesticides in urine is a useful method to assess the aggregate exposure of pesticides from various exposure sources and routes, by measuring the parent pesticide and/or the corresponding biotransformation products (Bonvallot et al., 2021). However, as the list of registered pesticides is long and they occur often highly metabolized in urine, a large number of targeted assays would be required to assess presence of all urinary pesticides and their metabolites in each sample. This is currently not feasible since many human urinary biomarkers of exposure (typically phase I/II metabolites) are often unknown, and the analytical reference standards are not readily available. Suspect screening (SS) approaches based on full scan High Resolution Mass Spectrometry (HRMS) emerge as an innovative way to assess the presence of a broad range of exposure markers and better capture the complexity of pesticide mixtures (Andra et al., 2017; Pourchet et al., 2020; Huber et al., 2022).

The study presented here, the Survey on PEstiCIde Mixtures in Europe (SPECIMEn), aimed to generate new pesticide exposure data in a harmonized pan-European setting (as part of the European Human Biomonitoring Initiative HBM4EU, www.hbm4eu.eu). This was done by analyzing 2,088 urine samples collected in five countries through a multi-laboratory high-throughput SS approach. This study aimed at exploring co-occurrence (probability of exposure) of pesticide biomarkers across Europe and within each participating country. It also

aimed at assessing differences of exposure patterns by location (living close to agricultural fields *versus* non-agricultural areas), seasons (differences in spraying activities), as well as age groups (adults *versus* children, of which the latter are more sensitive to health effects and usually have higher internal exposure levels, due to e.g. a higher food intake/kg body weight (Eskenazi et al., 1999; Sapbamrer et al., 2019)). The study design therefore provides insight into local contributions, based on a broad combination of pesticides. Higher detection frequencies of pesticide markers might be expected for those pesticides applied on local crops during the spraying season in residents living close to the agricultural fields.

MATERIAL AND METHODS

Sampling strategy

To create geographical coverage across Europe, study sites from five countries were included to provide insight into variations of pesticide exposure patterns across Europe, namely the Czech Republic, Hungary, Latvia, Spain and the Netherlands. Within each country, urine samples were collected simultaneously at two locations: agricultural and non-agricultural areas. Each address in the agricultural area was located within 250 meters from an agricultural field where pesticides were typically applied, mainly focusing on tree-crops or so-called 'overhead cultures' (except Latvia where tree-crops were hardly grown). These 'overhead-cultures' will result in potentially higher exposure concentrations in the air due to machine-drawn air blast or a hand-held overhead spray, which are more prone to drift (Willenbockel et al., 2022). The crop types differed slightly between countries due to e.g. differences in climate. A detailed description of the area selection in all five country can be found in Supplementary Material F. In summary, Spain focused on residential areas close to citrus fruits, Czech Republic on apples, vineyards, peach, plums and apricots, Hungary on apples, the Netherlands on apples and pears, and Latvia mostly on winter and summer rapeseed, summer wheat and barley. Non-agricultural areas were defined as sub-urban areas at least 500 meters away from any agricultural fields.

Per country at each agricultural and non-agricultural area, 50 parent-child pairs (50 households) were included (total of 100 parent-child pairs per country). Each parent-child pair was composed of one child aged 6 to 11 years at the time of inclusion, accompanied by one of their parents or legal guardians living in the same household. Adults who worked in the agricultural sector (i.e. farmers) were excluded from recruitment, since the sample size was too limited to distinguish occupational exposures. The same selection criteria were used in all five countries.

A minimum of 100 parent-child pairs per country (200 individuals) provided a first morning void urine sample, and completed a harmonized questionnaire. The admission of the questionnaire, sample collection procedures and timing of sampling was coordinated, and sampling materials such as cups and tubes were bought in bulk to avoid any batch differences. All collected urine samples were stored and transported refrigerated (at 4°C), until samples were aliquoted and stored at -80°C (within 48 hours of sample collection). Samples were transported to the laboratory of analysis after each season.

All households were visited twice: the first visit was made in winter 2019/2020 (season 1), the second in summer 2020 (season 2). The specific sampling dates (Supplementary Material A) differed slightly between study sites, partly due to differences in spraying season due to climate and the type of crop grown on the field. The sampling of the second season was slightly delayed (end of summer) due to the COVID-19 pandemic and accompanied uncertainties. The recruitment strategy differed between the study sites, a detailed description of the recruitment strategy per country can be found in Supplementary Material F. In summary, the Hungarian partner involved local public health officers to get in touch with the participants, while others sent out letters (the Czech Republic and the Netherlands), contacted colleagues as study participants (Spain and the Netherlands), conducted an online campaign (the Czech Republic and the Netherlands), and/or contacted participants through schools (Spain and Latvia). A detailed questionnaire was completed during the first season by the parent, and a subset of questions was asked again during the second season (Supplementary Material B). The joint questionnaire was developed in English, and subsequently translated to the local languages. The questionnaire covered personal and household characteristics, activities up to three days prior to sampling, potential pesticide exposure scenarios (occupational, usage of products containing pesticides), and the food consumption pattern of the day prior to sampling (origin of consumed foods as well as a food frequency table for food consumption 24 hours prior to sampling).

All partner countries acquired approval from the appropriate local medical ethical committees, and written informed consent was obtained from all participants (parents and children separately). A description of the ethical approval procedure per country can be found in supplementary material F. A harmonized informed consent form was used for all participants, which was evaluated by an internal HBM4EU review board.

Suspect Screening Approach

A SS methodology was applied to analyze the urine samples, of which a detailed description can be found in Huber et al., 2022. Briefly, the applied analytical workflow from sample preparation, instrumental analysis, and data processing was conducted under harmonized conditions in five different laboratories across Europe, in the Netherlands, Germany, France, the Czech Republic and Spain (Vitale et al., 2022). Each laboratory analyzed approximately 400 urine samples originating from one of the five SPECIMEn study sites. Samples were analyzed after each season, and potential batch effects were addressed (Huber et al., 2022). The suspect database generation, MS data analysis and confirmation procedures were performed in a centralized way. Several consolidated quality assurance/quality control (QA/QC) dispositions, parameters and criteria were first implemented to ensure the consistency of the results obtained across the different participating laboratories as well as to document the applied method performances (Vitale et al., 2022). The applied analytical workflow was described in detail by Huber et al., 2022 and consists of i) SPE cleanup/concentration (5-fold) of the urine after pH adjustment, ii) measurement of the extracts by full scan liquid chromatography coupled to HRMS (LC-HRMS), iii) data pre-processing and analysis, iv) prioritization of putative detects, v) generation of a list of representative samples for follow up identification experiments using tandem mass spectrometry

(MS/MS), and vi) final confirmation of putative detects by spectral comparison with reference standards either purchased/synthesized or generated *in vitro* by human liver S9 incubations. The curated suspect list of pesticides may include multiple metabolites originating from the same parent compound, resulting in a final datafile with potentially several metabolites that reflect exposure of the same parent compound. This redundancy is considered enhancing confidence. In the case of SPECIMEn, this list focused on pesticides and one aggregated list of known and predicted pesticide metabolites from all five laboratories was used as suspect database. 'Fully identified' were those with the highest level of confidence: Schymanski level 1 if a reference standard material is commercially available, or Schymanski 2 by diagnostic evidence acquired by human liver S9 incubation experiments (Schymanski et al., 2014). Biomarkers which were identified at a lower tier will end up in lower confidence levels, reflecting the level of uncertainty about the identity of that feature. In the context of the present paper, only biomarkers identified with confidence levels 1 and 2 were considered.

Statistical analysis

In line with the basic principle of the SS approach, the data generated in SPECIMEn are 'semi-quantitative', i.e. quantitative signal intensities for each representative spectrometric mass are reported per sample, yet these intensities cannot be considered as urinary pesticide concentrations and are not standardized across laboratories. The data was analysed by dichotomizing the intensities into 'detected' versus 'non-detected', which allows comparisons across study sites as well as inclusion of biomarkers with low detection rates in the statistical analysis.

The detection rate was calculated as the number of samples in which a particular biomarker was detected and identified with confidence levels 1 and 2 over the total number of samples collected, expressed in percentage. Based on the parent pesticides (if multiple metabolites and the parent pesticide were measured, these were considered as one), the patterns of co-occurrences were explored. First, the total number of pesticides per urine sample was evaluated. Secondly, with the usage of an UpSet plot it was evaluated which parent pesticide combinations co-occurred and how frequent. Thirdly, the correlation pattern in the total set of parent pesticides was evaluated for each study site with a weighted correlation network using the *IsingFit* R package v0.3.1 (van Borkulo et al., 2015). This package estimates the network based on the Ising model: combining L1-regularized logistic regression with EBIC model selection (gamma 0.25). On this network a clustering algorithm was applied (*walktrap)*, to detect communities of closely related features indicated by different colours in the network (Pons et al., 2005).

To assess the influence of co-variates, logistic mixed effects regression models were applied, with participant ID and household ID as random effects. Our main model includes fixed effects for season (season 1/season 2), location (agricultural/non-agricultural) and age category (child/adult). We assessed the sensitivity for further adjustment for potential confounding by including body mass index (BMI) level, education of the parent, consumption of homegrown foods (yearly average percentage), and a summary indicator for pesticide usage in an extended model. The pesticide usage indicator indicates whether pesticide containing products were

used up to three days prior to sampling either for human use, in the garden, indoors and/or for professional use. The estimates for season, location and age groups were transformed to Odds Ratios (OR) with 95% Confidence Intervals (CI) for both the main and extended models.

RESULTS

Population characteristics

The description of the study population for the five study sites of the SPECIMEn study is provided in Table 1. In total 2,088 urine samples were collected, which were equally spread across the five study sites and areas. The loss to follow-up of individuals between seasons was low, varying from 0.9 to 2.9%. Reasons for loss to follow-up were loss of contact, divorce and/or move to another location. The adult samples mainly originated from the mothers, while gender was equally divided across the children's samples. The mean age of the adults was comparable across all study sites, varying from 38 to 44 years. The mean BMI (self-reported) of the adults originating from Latvia and Hungary was slightly higher compared to the adults from other study sites. Most of the participants did not smoke, although in the agricultural areas of Spain and Hungary there was a substantial group of current smokers 35% and 45%, respectively (Supplementary Material B). Based on the total household income categories, participants of agricultural areas mostly earned less money than those living in non-agricultural areas. In all areas except the agricultural area in Hungary, the majority of the participants had a university education level. In Spain and Hungary, about half of the households in agricultural areas used pesticide products during summer season, which includes the use of consumer products, usage indoors, in the garden and/or professional use. These different categories are presented separately in Supplementary Material B. Overall, the homegrown food consumption percentage was higher in households in agricultural areas than those in non-agricultural areas, mostly during summer.

Annotations and detection rates

The application and harmonization of the SS approach was performed on 2,088 urine samples using the method described in detail in Huber et al. 2022. A total number of 498 tentative annotations of pesticide biomarkers was obtained and prioritized, of which 40 pesticide biomarkers were annotated with confidence level 1 or 2 (Table 2). These 40 related to a total of 29 parent pesticides. In addition to these 40, 54 other pesticide biomarkers (either pesticide metabolites or the parent pesticides as such) were detected with a lower confidence level (Schymanski levels 3-5) which are detailed in Supplementary Material C. These 54 are not further described in this paper and not used in the analyses.

For each annotated exposure marker (confidence levels 1 and 2), the overall detection rate per study site was calculated (Table 2). Overall, biomarkers were generally detected below 25% of the samples. The results evidenced a significant variability between study sites, with Latvia having generally the lowest number of detects and Spain the highest one. Overall, the metabolites related to the parent pesticides acetamiprid (N-demethylated metabolite) and

Table 1. *Descriptive characteristics of the SPECIMEn study participants by study site and location.*

Study site		***ES***[1]		***LV***[2]		***HU***[2]		***CZ***[2]		***NL***[2]	
Area		**Agricultural**	**Non-Agricultural**	**Agricultural**	**Non-Agricultural**	**Agricultural**	**Non-Agricultural**	**Agricultural**	**Non-Agricultural**	**Agricultural**	**Non-Agricultural**
Adult-child pairs[2]*, n*		52	53	50	51	51	52	51	60	55	50
Urine samples, n		206	212	200	202	201	208	204	238	219	198
	Season 1	104	106	100	102	102	104	102	120	110	100
	Season 2	102	106	100	100	99	104	102	118	109	98
Gender, female, %	*Adults*	50	87	90	82	94	85	71	60	71	66
	Children	54	49	58	47	49	52	43	43	53	46
Mean age, years	*Adults*	44	44	40	39	38	40	41	42	42	42
	Children	8.2	8.7	8.9	8.4	9.7	9.2	8.8	9.1	8.6	8.6
Mean BMI	*Adults*	25	24	26	26	26	26	24	24	24	23
	Children	17	17	17	17	18	19	16	16	16	16.0
Educational level adult, %											
No or only primary education		0	0	2.0	0	40	5.8	0	1.7	1.8	0
Secondary education		7.8	17	30	12	28	20	2.0	3.3	5.5	2.0
Tertiary education (post-secondary)		25	17	8.0	7.8	23	26	26	10.0	18	18
University studies (BSc, MSc, PhD)		67	66	60	77	8.0	48	71	83	71	76
Don't Know/ NA		0	0	0	3.9	0	0	2.0	1.7	3.6	4.1
Usage of pesticide (-products) up to 3 days prior to sampling[3]*,*											
n households	*Season 1*	9	5	4	7	2	6	4	2	1	6
	Season 2	27	8	12	8	22	7	14	12	10	4
Seasonal homegrown vegetables, fruit and/or herbs consumption,											
% of total consumption	*Winter*	6.7	1.1	30	22	23	4.5	13	10	2.0	0.1
	Spring	10	3.4	28	19	21	9.4	22	12	4.8	2.2
	Summer	12	8.0	63	44	41	25	64	51	15	7.8
	Autumn	8.9	6.0	63	45	39	17	45	40	8.3	4.5

1. ES: Spain, LV: Latvia, HU: Hungary, CZ: Czech Republic, NL: the Netherlands
2. Number of individuals included in season 1
3. Summary indicator which includes: pesticides for human use, use indoors, use in garden, and professional use. For specification of the categories see Supplementary Material B

chlorpropham (4-HSA metabolite) were most frequently detected in samples of all study sites. Other biomarkers that had detection rates of at least 10% (including both locations and both season) relate to the parent pesticides boscalid (not in Hungary), chlorpyrifos (only in Spain and Czech Republic), clothianidin (not in Latvia), cyprodinil (not in Latvia and Hungary), flonicamid (not in Latvia and Czech Republic), fluazifop (not in Latvia), fludioxonil (not in Hungary), imazalil (only in Spain and Latvia), imidacloprid (only in Spain), pirimiphos-methyl (not in Hungary), propamocarb (not in Latvia), pyrimethanil (not in Hungary), tebuconazole (not in Latvia), and thiamethoxam (only in Spain and Hungary). Biomarkers that were detected at low frequencies (<10%) across all study sites include 2,4-dichlorophenoxy acetic acid (2,4-D), ametoctradin, chlorantraniliprole, clopyralid, fluopyram, flupyradifurone, fluvalinate, penconazole, propyzamide, thiabendazole, thiacloprid, trifloxystrobin, as well as the metabolite permethric acid (DCCA) (originated from parent pesticides cypermethrin, cyfluthrin, permethrin or transfluthrin).

Table 2. *Annotated pesticide biomarkers with Schymanski confidence levels 1 and 2 (p = 40) and their overall detection frequency (%) per study site (Schymanski et al., 2014).*

ID	Pesticide type[1]	Parent pesticide	Pesticide (metabolite) annotation[2]	Confidence level[3]	Overall Detection Frequency (%)				
					ES[4]	LV	HU	CZ	NL
P1	H	2,4-Dichlorophenoxy-acetic acid	Parent	1	4.1	0	2.2	2.7	0
P2_a	I	Acetamiprid	-CH2	1	99	33	94	98	93
P3_a	F	Ametoctradin	-C2H6 +2O	1	5.0	2.7	1.2	4.8	2.9
P5_a	F	Boscalid	+O +SO3[5]	2	36	18	3.9	23	33
P5_b			+O +SO3[6]	2	7.2	0	0	0.5	0.2
P6	I	Chlorantraniliprole	+O	2	3.8	0.3	0.2	0	0.2
P8_a	H, GR	Chlorpropham	+O +SO3 (4-HSA)	1	56	32	31	34	75
P9_a	I	Chlorpyrifos (methyl)	TCPy	1	1.7	0	0.2	0.2	0.2
P9_b			-CH2	1	36	0	6.9	21.7	6.5
P10	H	Clopyralid	Parent	1	1.0	0	0	1.4	0.7
P11_a	I	Clothianidin (can come from thiamethoxam)	Parent	1	34	1.7	22	25	20
P11_b			-NO2 +H	1	0.5	0	0.2	0	0.2
P11_c			-CH2	2	21	0.8	9.8	6.6	3.1
P12_a	I	Cypermethrin, cyfluthrin, permethrin, transfluthrin	DCCA	1	0.5	0	0	0	0
P13_a	F	Cyprodinil	+O +SO3	2	14	7.7	2.7	10	26
P18_a	I	Flonicamid	Parent	1	1.7	0.8	2.0	2.7	5.7
P18_b			-C2HN	2	15	0.3	27	0.2	57
P19_a	H	Fluazifop	Parent[6]	1	20	2.5	11	18	21
P19_b			Parent[7]	1	8.1	1.5	4.9	5.2	8.2
P20	F	Fludioxonil	+O +C6H8O6	2	16	15	2.0	14	27
P21_a	F	Fluopyram	+O +SO3	2	3.6	0.5	0.2	1.1	1.0
P21_b			+O +C6H8O6	2	2.4	0.8	0.5	3.2	4.8
P21_c			-2H	2	11	6.7	0.5	3.4	3.1
P22_a	I	Flupyradifurone	Parent	1	2.6	0.3	0.5	0.7	2.2
P25_a	I, Ac	Fluvalinate	-C14H9NO	2	1.0	0	0.7	0.2	0
P27_a	F	Imazalil	+C6H8O6	2	19	11	8.3	4.5	4.6
P28_a	I	Imidacloprid	-NO2 +H	1	17	1.7	4.2	0.7	9.4
P32_a	F	Penconazole	+O +C6H8O6	2	6.5	1.7	2.2	2.0	2.4
P34_a	I, Ac	Pirimiphos-methyl	-CH2	1	85	10	6.6	24	48
P35_a	F	Propamocarb	Parent	1	9.6	1	11	5.0	23
P35_b			+O	2	21	5.5	18	12	43
P37	H	Propyzamide	+H2O3	2	8.6	0	0.5	0.9	1.0
P38_a	F	Pyrimethanil	+O +SO3	2	27	14	4.9	22	32
P38_b			+O	2	0.7	0	2.7	0	0.5
P40_a	F	Tebuconazole	-2H +2O	2	71	5.5	25	52	36
P41_a	F	Thiabendazole	+O +C6H8O6	2	0	0.8	0.2	0	0.5
P42_a	I	Thiacloprid	+O	2	8.4	0.8	2.9	7.9	4.6
P43_a	I	Thiamethoxam	Parent	1	0.7	0	2.4	0	0.5
P43_b			-NO2 +H	1	23	0	15	0	0.2
P46_a	F	Trifloxystrobin	-CH2 -CH2	2	0.7	0.5	0	3.6	3.8

1. H: Herbicide, F: Fungicide, I: Insecticide, GR: Plant Growth Regulator, Ac: Acaricide
2. Metabolite annotation: "-CH2" means the molecular formula of the metabolite is that of the parent minus CH2 (corresponding to demethylation). Similarly, "+O" means the metabolite is the parent compound plus one oxygen atom (hydroxylation). "+SO3" and "+C6H8O6" indicate sulfation and glucuronidation, respectively.
3. Schymanski confidence level, ranging from 1 to 5, (Schymanski et al., 2014)
4. ES: Spain, LV: Latvia, HU: Hungary, CZ: Czech Republic, NL: the Netherlands
5. Positive precursor ion
6. Negative precursor ion

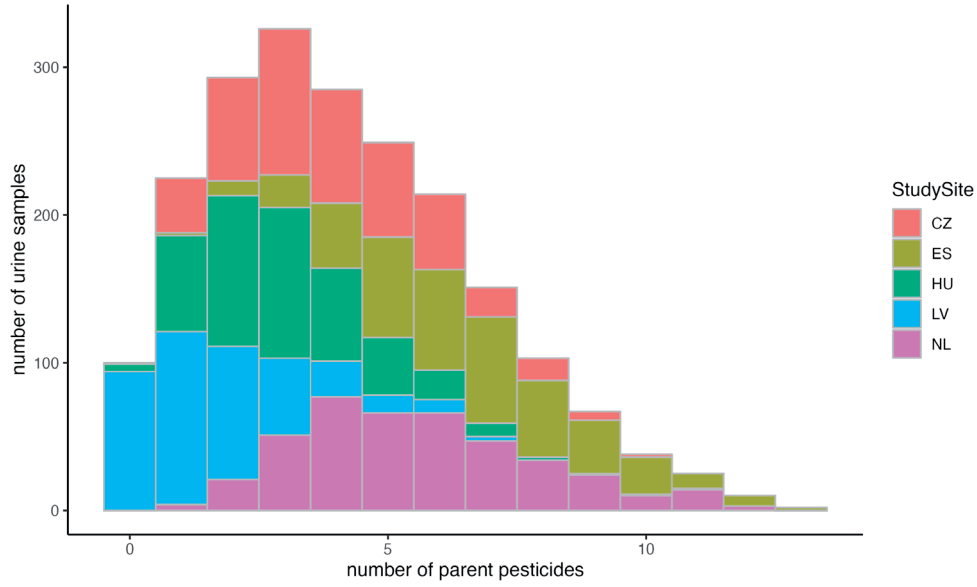


Figure 1. *Number of parent pesticides (p=29) detected per urine sample (n=2,088), with the five different study sites indicated in different colours (CZ=Czech Republic, ES=Spain, HU=Hungary, LV=Latvia, NL=Netherlands). Multiple metabolites and/or parent compounds related to the same parent pesticide were considered as one.*

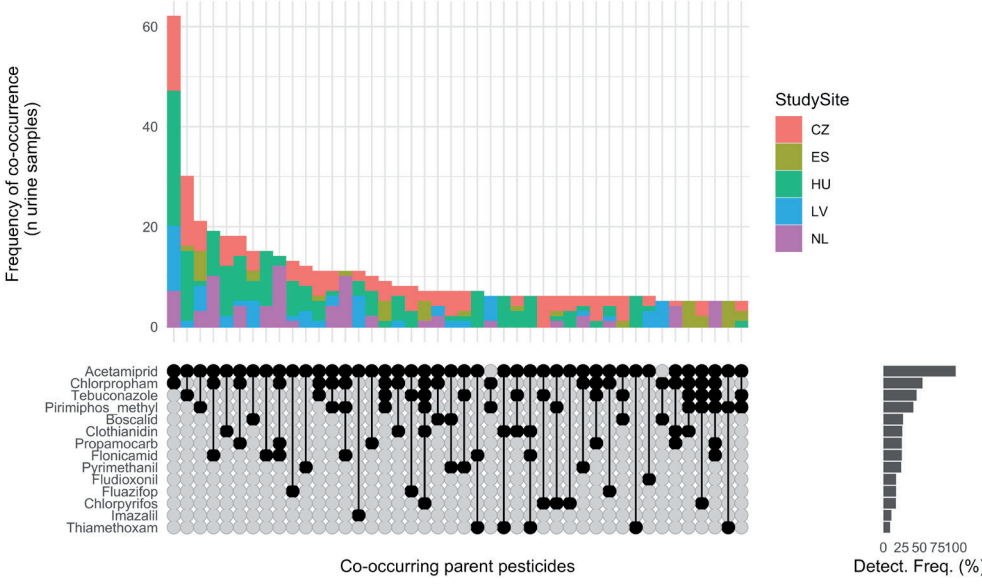


Figure 2. *Frequency (number of urine samples, n=2,088) of co-occurrent parent pesticides; the most frequent (in 5 or more urine samples) co-occurrences are shown. Different study sites are indicated by colours (CZ=Czech Republic, ES=Spain, HU=Hungary, LV=Latvia, NL=Netherlands), the detection frequency (%) of the listed parent pesticides is given on the right. Pesticides are co-occurring in the same sample when both have a black connected dot. Multiple metabolites and/or parent compounds related to the same parent pesticide were considered as one.*

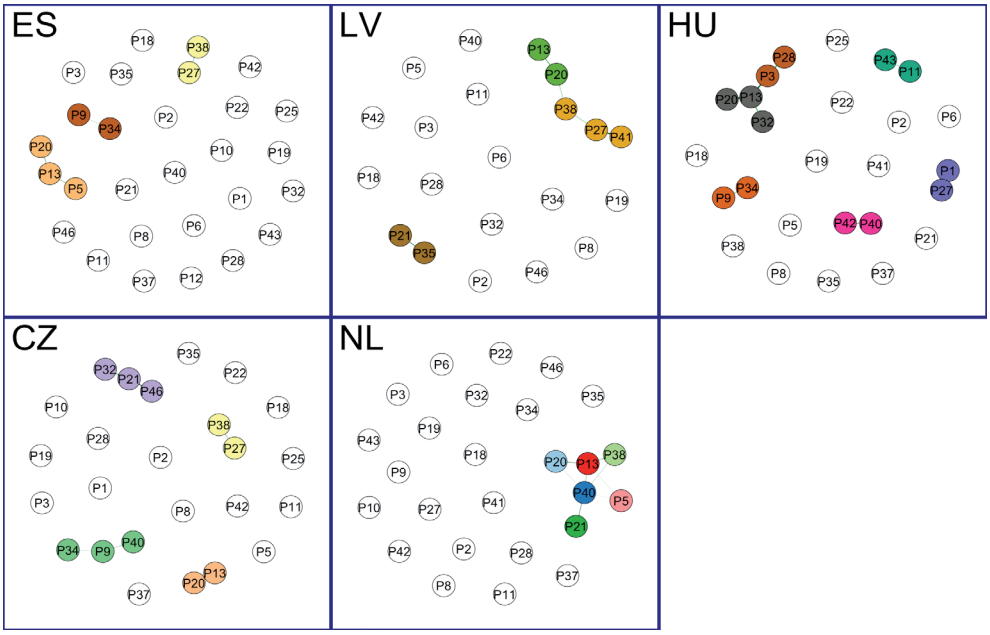


Figure 3. *Weighted correlation networks per study site based on the parent pesticides. Relationships between markers are indicated by a line (green = positive, red = negative). The colours indicate the different communities or groups of more closely related markers.* ***ES)*** *Spain (p=28),* ***LV)*** *Latvia (p=21),* ***HU)*** *Hungary (p=26),* ***CZ)*** *Czech Republic (p=25),* ***NL)*** *the Netherlands (p=26). See Table 2 for a description of the used ID numbers for each pesticide. Multiple metabolites and/or parent compounds related to the same parent pesticide were considered as one.*

Co-occurrence of pesticides

In order to assess how many pesticides were co-occurring within the same individual at a single time point, the number of detected parent pesticides per urine sample are presented in Figure 1. In line with the detection ratios, the lowest number of detected pesticides were in samples originating from Latvia, with mostly less than 3 co-occurring pesticides per urine sample. Samples originating from Spain showed the highest numbers of co-occurring pesticides, with a median value of 7. In the majority of the samples the number of parent pesticides per samples typically ranged from 2 to 5. The maximum number of different pesticides detected in the same urine sample was 13, which was the case for two samples. The samples with no (n=100) or only one (n=225) detected pesticide add up to 16% of the total amount of samples, indicating that in a majority of the samples from the SPECIMEn study at least two different parent pesticides were detected.

The next step was to evaluate which pesticides were co-occurring in each urine sample, for which the most frequently (in 5 or more urine samples) co-occurrent pesticides or mixtures are presented in an UpSet plot in Figure 2. These most frequent co-occurrences consisted of 44 different combinations based on 14 different pesticides. The majority of co-occurrences consisted of 2 or 3 pesticides, with minimal overlap across all study sites. The most common co-occurrence was acetamiprid with chlorpropham, detected in 62 samples although this combination was not detected in any sample originating from Spain. The second most fre-

quently co-occurring pesticides were acetamiprid with tebuconazole, however this combination was not seen in the Netherlands. The only co-occurrence combination detected in all five study sites was acetamiprid with pirimiphos-methyl. The less frequent the co-occurrent pesticides the more variation in combinations were seen, which was even more pronounced when detected in just 2, 3 or 4 urine samples (see Supplementary Material D for the extended UpSet plot).

The stability of the co-occurrences at each study site can be evaluated with correlation networks, which are presented in Figure 3A-E. Similar to the findings of Figure 2, mostly small groups (two to four biomarkers) of co-occurrent pesticides were found. Consistent across all study sites was the positive relation between cyprodinil (P13) and fludioxonil (P20), both fungicides, although sometimes together with other pesticides and/or part of a different community. Also, in both Spain and the Czech Republic, imazalil (P27) was related to pyrimethanil (P38), which are both fungicides. Finally, in Spain and Hungary chlorpyrifos-methyl (P9) was related to pirimiphos-methyl (P34), which are both insecticides. Interestingly, the relations of acetamiprid (P2) with chlorpropham (P8) or tebuconazole (P40) were not detected in the networks.

Changes in occurrence of pesticides by location, season, age category

To explore the differences in occurrence of the pesticide biomarkers by location, season, and age category, logistic mixed effects models were constructed. The main model includes the covariates for location, season and age category, the extended model was also corrected for pesticide usage (self-reported), BMI, level of education and homegrown food consumption. Results of the models of the biomarkers detected in at least four study sites are shown in Table 3, the full table with estimates for all exposure markers associated with confidence levels 1 and 2 can be found in Supplementary Material E.

In Spain, no effect of location was detected in the models, except for clothianidin which was less frequently detected in agricultural areas compared to non-agricultural areas. Chlorpropham, chlorpyrifos, clothianidin, fluazifop, fludioxonil, imazalil, imidacloprid, pyrimethanil, and tebuconazole were most frequently detected during the first sampling season. These effects were not influenced by inclusion of the additional predictors in the extended model. Between the group of parents and children in Spain, the biomarkers related to boscalid, and cyprodinil were most frequently detected among parents, while chlorpropham, chlorpyrifos, chlothianidin, pirimiphos-methyl, tebuconazole, and thiacloprid were more frequently detected among children. The extended models confirmed most of these effects (not for clothianidin and cyprodinil).

In Latvia, propamocarb was the only biomarker more frequently detected at the agricultural area. Acetamiprid, fluopyram, imazalil, and propamocarb were more frequently detected in the first season (winter), while pyrimethanil and tebuconazole were more frequently detected during the second season (summer). Only the effects related to propamocarb and pyrimethanil were confirmed with the extended models. Chlorpropham, pirimiphos-methyl, and propamocarb were more frequently detected among the Latvian children compared to adults, while imazalil was more frequently detected within Latvian parents (not in extended model).

In Hungary, both biomarkers related to clothianidin were more frequently detected at the agricultural areas. On the other hand, chlopyrifos, pirimiphos-methyl, propamocarb, tebuconazole, and thiacloprid were most frequently detected at the non-agricultural areas. Chlorpyrifos, chlothianidin, pirimiphos-methyl, propamocarb, and tebuconazole were most frequently detected during the second season. While, in contrary, chlorpropham and imazalil were most frequently detected during the first season. Acetamiprid, chlorpropham, chlorpyrifos, chlothianidin, fluazifop, pirimiphos-methyl, propamocarb, and tebuconazole, were most frequently detected among the Hungarian children. Of which chlorpropham, fluazifop, pirimiphos-methyl, propamocarb and tebuconazole were confirmed in both models.

In the Czech Republic, the metabolite of ametoctradin was more frequently detected at the agricultural areas, although this effect disappeared in the extended model. The biomarkers related to cyprodinil and fludioxonil were more frequently detected at the non-agricultural locations (only cyprodinil confirmed with the extended model). The chlorpropham metabolite (4-HSA) was more frequently detected during the second season. While the biomarkers related to ametoctradin, imazalil, and pyrimethanil showed an opposite effect, and were more frequently detected during the first season. Of these three, only the effect of pyrimethanil was confirmed with the extended model. Seven different biomarkers were found to be more detected among children compared to adults: boscalid, chlorpropham, chlorpyrifos, flonicamid, pirimiphos-methyl, tebuconazole, and thiacloprid. The extended model confirmed the effects seen for chlorpropham and tebuconazole.

Finally, in the Netherlands, the metabolites of chlorpropham were most frequently detected at agricultural areas. While biomarkers related to cyprodinil and pyrimethanil had higest detection frequencies at the non-agricultural areas. Chlorpropham, fluazifop, and thiacloprid were more frequently detected during the second season, while acetamiprid, chlorpyrifos, clothianidin, imazalil, and pyrimethanil had highest detection rates during the first season. The biomarkers related to ametoctradin, cyprodinil, flonicamid, fludioxonil, pirimiphos-methyl, and tebuconazole were more frequently detected among children. Of these, the effects seen for pirimiphos and tebuconazole were confirmed in the extended models. While on the other hand, propamocarb was more frequently detected among adults (only in extended model).

Overall, almost no biomarkers were more frequently detected in both agriculture areas and (summer) season 2. Only exceptions were chlorpropham (4-HSA metabolite) in the Netherlands, and clothianidin (parent compound and the N-demethylated metabolite) in Hungary.

DISCUSSION

This study reports on the co-occurrence patterns of 40 different pesticide biomarkers at study sites from five European countries, and identifies whether proximity to agricultural fields, season, and age category impacted the probability of detection of these biomarkers. The developed application of a harmonized SS methodology allowed screening for 1000s of suspects (pesticides and their known/predicted phase I/II metabolites), and enabled detection

Table 3. *Results of logistic mixed effects models, main and extended. Results are presented as Odds Ratios (OR) with 95% confidence intervals (CI). Significance levels based on p-value: '***' <0.001, '**' <0.01, '*' <0.05. Random effects are household and participant ID. Main model includes the predictors: location, season, and age category. Extended model includes additional predictors for pesticide usage, BMI, level of education and homegrown food consumption. Results are shown of features detected in at least 4 study sites.*

			SP		LV	
ID	Parent pesticide	Category	Main OR (95% CI)[1]	Extended OR (95% CI)	Main OR (95% CI)	Extended OR (95% CI)
P2_a	Acetamiprid	Agricultural vs Non-Agricultural[2]	NA[3]	NA[3]	1.0 (0.6; 1.6)	1.1 (0.7; 1.9)
		Season 2 vs 1[2]	0.5 (0.1; 2.7)	0.5 (0.1; 2.7)	0.6 (0.4; 1.0)	0.6 (0.4; 1.0) *
		Parent vs Child[2]	0.2 (0.0; 1.7)	0.2 (0.0; 2.8)	0.8 (0.5; 1.3)	0.7 (0.4; 1.5)
P3_a	Ametoctradin	Agri. vs Non-Agri.	0.3 (0.0; 2.8)	0.4 (0.1; 1.0)	1.8 (0.5; 6.3)	1.3 (0.3; 5.3)
		Season 2 vs 1	0.4 (0.1; 1.5)	0.6 (0.2; 1.4)	0.6 (0.2; 2.0)	0.5 (0.2; 1.9)
		Parent vs Child	2.1 (0.3; 17)	3.0 (0.9; 10)	0.8 (0.2; 2.8)	2.3 (0.3; 15)
P5_a	Boscalid	Agri. vs Non-Agri.	1.0 (0.6; 1.9)	1.0 (0.5; 1.8)	1.2 (0.5; 2.6)	1.4 (0.6; 3.1)
		Season 2 vs 1	0.7 (0.4; 1.0)	0.6 (0.4; 1.0)	0.9 (0.5; 1.6)	0.9 (0.5; 1.6)
		Parent vs Child	2.9 (1.8; 4.6) ***	2.5 (1.3; 4.9) **	1.3 (0.7; 2.4)	1.0 (0.4; 2.6)
P8_a	Chlorpropham	Agri. vs Non-Agri.	0.7 (0.4; 1.3)	0.7 (0.4; 1.3)	1.3 (0.7; 2.7)	1.2 (0.6; 2.5)
		Season 2 vs 1	0.4 (0.3; 0.7) ***	0.4 (0.3; 0.7) ***	1.6 (1.0; 2.6)	1.5 (0.9; 2.4)
		Parent vs Child	0.4 (0.2; 0.6) ***	0.3 (0.2; 0.6) ***	0.3 (0.2; 0.6) ***	0.4 (0.2; 1.0) *
P9_a	Chlorpyrifos (/methyl)	Agri. vs Non-Agri.	0.8 (0.5; 1.3)	0.8 (0.5; 1.3)	ND[6]	ND
		Season 2 vs 1	0.2 (0.1; 0.4) ***	0.2 (0.1; 0.4) ***		
		Parent vs Child	0.5 (0.3; 0.7) ***	0.4 (0.2; 0.7) **		
P11_a	Clothianidin (can come from thiamethoxam)	Agri. vs Non-Agri.	0.5 (0.3; 0.8) **	0.4 (0.3; 0.7) ***	1.4 (0.3; 6.3)	0.9 (0.2; 4.6)
		Season 2 vs 1	0.6 (0.4; 0.9) **	0.5 (0.3; 0.8) **	6.3 (0.7; 53)	5.7 (0.7; 50)
		Parent vs Child	0.6 (0.5; 0.9) *	0.6 (0.3; 1.0)	0.2 (0.0; 1.4)	0.3 (0.0; 5.7)
P11_c			1.4 (0.8; 2.8)	1.4 (0.7; 2.8)	ND	ND
			0.9 (0.5; 1.5)	0.8 (0.5; 1.4)		
			0.7 (0.4; 1.3)	0.8 (0.4; 1.8)		
P13_a	Cyprodinil	Agri. vs Non-Agri.	1.4 (0.7; 2.7)	1.4 (0.7; 2.7)	1.1 (0.4; 2.8)	1.1 (0.4; 2.9)
		Season 2 vs 1	1.2 (0.7; 2.1)	1.2 (0.6; 2.2)	0.7 (0.3; 1.5)	0.7 (0.3; 1.5)
		Parent vs Child	2.0 (1.6; 5.6) ***	2.3 (1.0; 5.4)	0.9 (0.4; 2.1)	1.5 (0.5; 5.1)
P18_a	Flonicamid	Agri. vs Non-Agri.	2.6 (0.5; 14)	2.6 (0.4; 15)	ND	ND
		Season 2 vs 1	1.4 (0.3; 6.3)	1.4 (0.3; 7.1)[5]		
		Parent vs Child	6.2 (0.7; 52)	7.8 (0.6; 96)		
P19_a	Fluazifop	Agri. vs Non-Agri.	1.0 (0.6; 1.9)	1.1 (0.6; 2.1)	4.2 (0.9; 20)	3.1 (0.6; 16)
		Season 2 vs 1	0.5 (0.3; 0.8) **	0.5 (0.3; 0.9) *	0.7 (0.2; 2.4)	0.6 (0.1; 2.3)
		Parent vs Child	1.0 (0.6; 1.7)	0.7 (0.3; 1.5)	1.0 (0.3; 3.5)	1.2 (0.2; 7.8)
P19_b			1.6 (0.7; 3.6)	1.6 (0.6; 4.1)	5.2 (0.6; 45)	4.6 (0.5; 44)
			0.8 (0.4; 1.6)	0.7 (0.3; 1.6)	0.5 (0.1; 2.7)	0.4 (0.1; 2.6)
			1.0 (0.5; 2.1)	0.4 (0.1; 1.2)	0.5 (0.1; 2.7)	1.9 (0.1; 29)
P20	Fludioxonil	Agri. vs Non-Agri.	1.1 (0.6; 2.1)	1.0 (0.5; 2.0)	0.8 (0.4; 1.7)	0.9 (0.4; 1.9)
		Season 2 vs 1	0.5 (0.3; 0.9) *	0.5 (0.3; 0.9) *	0.8 (0.4; 1.4)	0.8 (0.4; 1.4)
		Parent vs Child	1.5 (0.8; 2.8)	0.9 (0.4; 2.2)	1.1 (0.6; 2.1)	0.8 (0.3; 2.1)
P21_c	Fluopyram	Agri. vs Non-Agri.	1.4 (0.6; 3.4)	1.5 (0.6; 4.0)	0.6 (0.1; 4.3)	0.7 (0.2; 2.2)
		Season 2 vs 1	1.0 (0.5; 1.9)	1.1 (0.5; 2.2)	0.2 (0.0; 0.8) *	0.4 (0.2; 1.1)
		Parent vs Child	1.5 (0.8; 3.0)	0.9 (0.3; 2.6)	0.9 (0.1; 6.1)	1.1 (0.3; 4.3)
P27_a	Imazalil	Agri. vs Non-Agri.	1.0 (0.5; 2.0)	1.1 (0.6; 2.2)	1.7 (0.6; 4.9)	0.7 (0.2; 2.2)
		Season 2 vs 1	0.2 (0.1; 0.3) ***	0.2 (0.1; 0.4) ***	0.4 (0.2; 0.8) *	0.4 (0.2; 1.1)
		Parent vs Child	1.1 (0.7; 2.0)	0.8 (0.4; 1.9)	2.4 (1.1; 5.2) *	1.1 (0.3; 4.3)
P28_a	Imidacloprid	Agri. vs Non-Agri.	1.5 (0.7; 3.1)	1.2 (0.6; 2.6)	1.4 (0.3; 6.2)	1.1 (0.2; 5.7)
		Season 2 vs 1	0.6 (0.3; 1.0) *	0.5 (0.3; 1.0) *	0.2 (0.0; 1.4)	0.2 (0.0; 1.4)
		Parent vs Child	1.4 (0.8; 2.5)	1.8 (0.8; 4.2)	1.4 (0.3; 6.1)	0.7 (0.1; 6.2)
P32_a	Penconazole	Agri. vs Non-Agri.	0.7 (0.3; 1.7)	0.8 (0.3; 1.9)	1.4 (0.3; 6.1)	1.6 (0.3; 7.7)
		Season 2 vs 1	1.1 (0.5; 2.5)	1.2 (0.5; 2.6)	0.8 (0.2; 3.4)	0.8 (0.2; 3.4)[5]
		Parent vs Child	2.2 (0.9; 5.0)	2.9 (0.9; 9.0)	0.8 (0.2; 3.4)	0.6 (0.1; 6.0)
P34_a	Pirimiphos-methyl	Agri. vs Non-Agri.	0.7 (0.3; 1.4)	1.1 (0.5; 2.5)	1.0 (0.5; 1.9)	1.5 (0.3; 7.4)
		Season 2 vs 1	0.8 (0.4; 1.4)	1.0 (0.5; 2.0)	1.5 (0.8; 2.9)	0.8 (0.2; 3.8)
		Parent vs Child	0.4 (0.2; 0.9) *	0.2 (0.1; 0.7) *	0.4 (0.2; 0.9) *	0.6 (0.1; 5.4)

HU		CZ		NL	
Main OR (95% CI)	Extended OR (95% CI)	Main OR (95% CI)	Extended OR (95% CI)	Main OR (95% CI)	Extended OR (95% CI)
1.2 (0.1; 2.7)	1.4 (0.5; 3.5)	0.5 (0.1; 2.2)	0.6 (0.1; 3.0)	2.0 (0.3; 14)	2.4 (0.9; 6.2)
1.2 (0.5; 2.7)	1.3 (0.5; 3.1)	0.6 (0.1; 2.5)	0.8 (0.2; 3.9)	0.2 (0.0; 0.8) *	0.5 (0.2; 1.2)
0.4 (0.2; 1.0) *	0.5 (0.2; 1.5)	1.0 (0.2; 4.1)	12 (1.0; 149)	1.3 (0.2; 9.0)	1.7 (0.3; 8.0)
1.6 (0.3; 9.5)	NA[4]	3.2 (1.1; 9.5) *	3.0 (1.0; 9.4)	0.8 (0.03; 20)	NA[4]
0.7 (0.1; 4.1)		0.4 (0.1; 1.0) *	0.4 (0.1; 1.1)	0.3 (0.1; 1.6)	
4.0 (0.4; 37)		0.6 (0.2; 1.5)	0.2 (0.0; 1.3)	0.1 (0.02; 0.9) *	
0.8 (0.2; 2.2)	0.5 (0.1; 1.5)	1.4 (0.8; 2.5)	1.4 (0.8; 2.7)	0.6 (0.3; 1.0)	0.5 (0.3; 1.0)
0.6 (0.2; 1.7)	0.6 (0.2; 1.9)	0.7 (0.4; 1.2)	0.8 (0.5; 1.3)	1.4 (0.9; 2.2)	1.4 (0.9; 2.2)
2.1 (1.0; 9.9)	2.4 (0.6; 9.3)	2.1 (1.3; 3.5) **	2.5 (1.0; 6.1)	1.4 (0.9; 2.1)	1.1 (0.4; 2.6)
1.3 (0.7; 2.5)	1.5 (0.7; 3.2)	1.0 (0.6; 2.0)	1.0 (0.5; 2.0)	2.1 (1.1; 3.9) *	2.1 (1.1; 4.1) *
0.5 (0.3; 0.8) **	0.5 (0.3; 0.9) *	2.1 (1.3; 3.3) **	1.9 (1.2; 3.1) *[5]	2.8 (1.7; 4.7) ***	2.7 (1.6; 4.6) ***
0.5 (0.3; 0.7) **	0.4 (0.2; 0.7) **	0.4 (0.2; 0.6) ***	0.3 (0.1; 0.8) *	0.6 (0.4; 1.1)	0.5 (0.2; 1.2)
0.2 (0.1; 0.7) *	0.3 (0.1; 1.0)	1.3 (0.7; 2.4)	1.3 (0.7; 2.4)	1.2 (0.4; 3.2)	1.2 (0.4; 3.3)
2.5 (1.0; 6.1) *	2.7 (1.1; 6.5) *	0.6 (0.4; 1.0)	0.6 (0.4; 1.0)	0.5 (0.2; 1.1)	0.4 (0.1; 1.0) *
0.5 (0.2; 1.1)	0.2 (0.1; 0.8) *	0.5 (0.3; 0.7) **	0.7 (0.3; 1.7)	0.8 (0.3; 1.8)	0.9 (0.2; 5.2)
2.8 (1.6; 4.7) ***	2.8 (1.5; 5.1) **	1.3 (0.8; 2.2)	1.4 (0.8; 2.5)	1.3 (0.7; 2.3)	1.4 (0.8; 2.7)
3.1 (1.8; 5.2) ***	3.5 (2.0; 6.1) ***	0.6 (0.4; 1.0)	0.7 (0.4; 1.1)	0.6 (0.4; 1.0)	0.6 (0.4; 1.0)
0.6 (0.4; 1.0)	0.4 (0.2; 0.7) **	1.0 (0.6; 1.5)	1.1 (0.5; 2.7)	0.9 (0.5; 1.5)	1.7 (0.6; 4.6)
4.4 (1.8; 11) ***	5.5 (2.1; 14) ***	1.3 (0.6; 2.7)	1.3 (0.6; 2.8)	1.5 (0.0; 38)	1.4 (0.7; 2.6)
1.9 (0.9; 3.9)	2.5 (1.2; 5.5) *	0.5 (0.2; 1.1)	0.5 (0.2; 1.1)	0.02 (0.0; 0.3) **	0.6 (0.4; 1.0)[5,7]
0.8 (0.4; 1.5)	0.3 (0.1; 0.8) *	0.9 (0.4; 2.0)	1.2 (0.3; 4.6)	1.6 (0.0; 42)	1.8 (0.7; 4.9)
1.9 (0.1; 44)	1.8 (0.1; 61)	0.3 (0.1; 0.6) **	0.3 (0.1; 0.6) **	0.6 (0.3; 0.9) *	0.5 (0.3; 0.9) *
0.3 (0.1; 2.0)	0.4 (0.1; 2.9)	0.9 (0.5; 1.6)	0.9 (0.5; 1.7)	0.8 (0.5; 1.2)	0.8 (0.5; 1.2)
6.2 (0.9; 40)	2.9 (0.3; 26)	0.8 (0.4; 1.5)	1.0 (0.3; 3.0)	0.5 (0.3; 0.8) **	1.2 (0.5; 2.8)
1.0 (0.3; 4.3)	2.3 (0.5; 11)	2.7 (0.1; 84)	NA[4]	0.4 (0.1; 1.5)	0.4 (0.1; 1.3)
3.1 (0.6; 16)	3.8 (0.7; 20)	1.0 (0.2; 5.3)		1.6 (0.6; 4.0)	1.5 (0.6; 3.8)
0.3 (0.1; 1.6)	0.2 (0.0; 1.5)	0.1 (0.0; 0.4) **		0.3 (0.1; 0.8) *	0.3 (0.0; 2.3)
1.1 (0.5; 2.2)	0.9 (0.4; 2.1)	1.1 (0.6, 2.1)	1.3 (0.7; 2.4)	1.4 (0.7; 3.0)	1.5 (0.7; 3.1)
1.1 (0.6; 2.0)	0.9 (0.4; 1.7)	0.9 (0.5; 1.4)	0.8 (0.5; 1.4) [5]	1.0 (0.6; 1.6)	1.0 (0.6; 1.7)
0.6 (0.3; 1.2)	0.4 (0.2; 1.0) *	0.7 (0.4; 1.2)	0.5 (0.2; 1.3)	0.7 (0.4; 1.2)	0.5 (0.2; 1.5)
2.1 (0.3; 17)	2.4 (0.8; 7.2)	0.9 (0.3; 2.3)	0.9 (0.3; 2.5)	1.4 (0.2; 10)	3.1 (0.3; 29)
1.4 (0.1; 4.5)	1.0 (0.3; 2.6)	2.5 (1.0; 6.3)	2.4 (0.9; 6.3) [5]	7.2 (1.6; 32) **	4.5 (1.1; 17) *
0.2 (0.0; 0.9) *	0.3 (0.1; 0.9) *	1.3 (0.6; 3.2)	1.0 (0.2; 4.6)	0.6 (0.1; 4.1)	0.0 (0.0; 2.9)
0.6 (0.1; 2.6)	0.9 (0.2; 4.8)	0.5 (0.3; 0.9) *	0.6 (0.3; 1.1)	0.9 (0.5; 1.6)	0.8 (0.5; 1.4)
1.7 (0.4; 7.2)	1.7 (0.4; 7.5)	0.7 (0.4; 1.2)	0.7 (0.4; 1.2)	0.6 (0.3; 0.9) *	0.6 (0.4; 0.9) *
1.0 (0.2; 4.0)	0.9 (0.2; 5.5)	0.7 (0.4; 1.2)	1.1 (0.4; 3.1)	0.5 (0.3; 0.9) *	0.8 (0.3; 1.8)
ND	ND	1.3 (0.2; 10)	1.9 (0.2; 20)	1.0 (0.1; 22)	NA[13]
		1.8 (0.5; 6.4)	1.8 (0.5; 7.2)	4.3 (0.5; 36)	
		0.8 (0.2; 2.8)	1.0 (0.1; 12)	0.5 (0.2; 12)	
0.6 (0.2; 2.0)	0.7 (0.2; 2.3)	0.9 (0.1; 11)	1.0 (0.4; 2.6)	1.6 (0.6; 4.2)	1.6 (0.6; 4.1)
0.2 (0.1; 0.5) ***	0.2 (0.1; 0.5) **	0.1 (0.0; 0.6) *	0.4 (0.2; 1.2)[5]	0.3 (0.1; 1.0) *	0.4 (0.1; 1.0)[5]
0.5 (0.2; 1.2)	0.3 (0.1; 1.1)	2.9 (0.2; 44)	3.1 (0.6; 15)	1.4 (0.6; 3.6)	2.2 (0.4; 11)
0.9 (0.2; 3.3)	1.0 (0.3; 3.7)	ND	ND	1.2 (0.6; 2.5)	1.4 (0.6; 2.9)
0.9 (0.3; 2.5)	1.1 (0.4; 3.0)			1.1 (0.5; 2.1)	1.1 (0.5; 2.1)
0.9 (0.3; 2.4)	0.5 (0.1; 2.0)			0.8 (0.4; 1.7)	0.9 (0.2; 3.3)
0.5 (0.1; 2.1)	0.7 (0.0; 38)	2.4 (0.6; 9.7)	2.2 (0.5; 9.3)	0.9 (0.3; 3.2)	0.8 (0.2; 3.0)
3.6 (0.7; 18)	9.5 (1.0; 93)	0.5 (0.1; 2.0)	0.6 (0.1; 2.5)[5,7]	1.0 (0.3; 3.6)	1.0 (0.3; 3.6) [5]
1.3 (0.3; 4.8)	0.7 (0.0; 12)	2.0 (0.5; 8.3)	2.9 (0.3; 30)	1.5 (0.4; 5.5)	0.5 (0.1; 4.5)
0.1 (0.02; 0.4) ***	0.2 (0.1; 0.8) *	1.5 (0.8; 2.7)	1.3 (0.7; 2.5)	1.4 (0.8; 2.7)	1.4 (0.8; 2.7)
3.6 (1.4; 9.4) **	4.1 (1.5; 11) **	0.6 (0.4; 1.0)	0.6 (0.4; 1.1)	0.7 (0.4; 1.0)	0.7 (0.4; 1.0)
0.2 (0.1; 0.5) **	0.1 (0.0; 0.3) ***	0.5 (0.3; 0.8) **	0.6 (0.2; 1.5)	0.3 (0.2; 0.5) ***	0.3 (0.1; 0.8) *

Table 3. *Continued*

ID	Parent pesticide	Category	SP Main OR (95% CI)[1]	SP Extended OR (95% CI)	LV Main OR (95% CI)	LV Extended OR (95% CI)
P35_a	Propamocarb	Agri. vs Non-Agri.	1.3 (0.5; 3.1)	1.0 (0.4; 2.4)	ND	ND
		Season 2 vs 1	1.9 (0.9; 4.1)	1.7 (0.8; 3.5)		
		Parent vs Child	1.1 (0.5; 2.5)	1.2 (0.4; 3.4)		
P35_b			1.3 (0.7; 2.4)	1.2 (0.7; 2.2)	4.2 (1.1; 16) *	4.0 (1.0; 17)
			1.2 (0.7; 1.9)	1.1 (0.7; 1.9)	0.3 (0.1; 0.9) *	0.3 (0.1; 0.8) *
			1.1 (0.7; 1.9)	1.1 (0.5; 2.3)	0.4 (0.1; 1.1)	0.2 (0.0; 1.0) *
P38_a	Pyrimethanil	Agri. vs Non-Agri.	0.8 (0.5; 1.3)	0.8 (0.5; 1.3)	1.2 (0.6; 2.4)	1.1 (0.5; 2.3)
		Season 2 vs 1	0.6 (0.4; 1.0) *	0.7 (0.4; 1.0)	2.1 (1.1; 3.8) *	2.1 (1.1; 3.8) *
		Parent vs Child	0.8 (0.5; 1.2)	0.7 (0.3; 1.3)	1.0 (0.6; 1.8)	1.3 (0.5; 3.2)
P40_a	Tebuconazole	Agri. vs Non-Agri.	0.9 (0.5; 1.7)	0.7 (0.4; 1.4)	1.5 (0.4; 5.3)	1.5 (0.6; 4.0)
		Season 2 vs 1	0.5 (0.3; 0.9) *	0.5 (0.3; 0.8) **	3.3 (1.1; 9.3) *	2.8 (1.1; 7.4)
		Parent vs Child	0.3 (0.2; 0.4) ***	0.2 (0.1; 0.4) ***	0.8 (0.3; 2.1)	2.9 (0.7; 12)
P42_a	Thiacloprid	Agri. vs Non-Agri.	1.6 (0.6; 4.4)	1.6 (0.6; 4.5)	ND	ND
		Season 2 vs 1	0.9 (0.4; 2.0)	1.1 (0.5; 2.4)		
		Parent vs Child	0.4 (0.2; 1.0) *	0.3 (0.1; 1.2)		

[1] OR: Odds Ratio, CI: Confidence Interval.
[2] Underlined is the reference category.
[3] 100% detected in one of the categories, no estimate could be provided.
[4] Due to low detection rate no extended model possible.
[5] Model not corrected for level of Education, separation issue.
[6] ND: Not detected or low detection rate (<1%), no model possible.
[7] Model not corrected for Pesticide usage, separation issue

of many pesticides/metabolites at different levels of confidence in urine. As such, this study should be seen as the first step towards a more complete assessment of the pesticide mixture exposure in the general European population.

Detected pesticides and the impact of location, season and age category

The most frequently detected biomarkers across all study sites were related to the parent pesticides acetamiprid and chlorpropham. Acetamiprid is a neonicotinoid (insecticide), is approved in the EU and commonly used on fruit trees such as apples, pears and citrus, but also on e.g. potatoes and rapeseed (Allema et al., 2017; EU Database Pest, 2022). All study sites included agricultural areas where these crops are grown. However since we did not find a difference between areas for acetamiprid, this high detection frequency is likely due to other factors such as diet. For Latvia and the Netherlands, acetamiprid was less frequently detected during the second season (summer), arguing that additional exploration is needed on for example the change of diet between seasons. Chlorpropham is a plant growth regulator and herbicide, commonly used on e.g. onions and potatoes to prevent sprouting. In the Netherlands only, chlorpropham had a higher probability of detection in the summer season, which is consistent with an earlier study on flower bulb fields in the Netherlands (Gooijer et al., 2019; Oerlemans et al., 2021). Although chlorpropham has no longer been approved as pesticide since 2019 in the EU, still high probabilities of detection were seen in both seasons (EU Database Pest, 2022). This is not unexpected due to periods of grace until October 2020, which overlaps with both sampling periods of the current study. Interestingly, in Spain and Hungary chlorpropham was more frequently detected during the first season, while in Czech Republic

HU		CZ		NL	
Main OR (95% CI)	Extended OR (95% CI)	Main OR (95% CI)	Extended OR (95% CI)	Main OR (95% CI)	Extended OR (95% CI)
0.5 (0.2; 1.0)	0.5 (0.2; 1.2)	0.5 (0.1; 4.5)	0.5 (0.2; 1.9)	1.6 (0.8; 3.0)	1.9 (1.0; 3.6)
1.1 (0.6; 2.1)	1.1 (0.6; 2.2)	1.5 (0.4; 6.0)	1.4 (0.5; 3.5) [7]	1.4 (0.9; 2.4)	1.5 (0.9; 2.5)
0.4 (0.2; 1.0) *	0.3 (0.1; 0.9) *	0.5 (0.1; 3.9)	1.5 (0.2; 10)	1.0 (0.6; 1.6)	3.1 (1.1; 8.5) *
0.5 (0.3; 0.9) *	0.5 (0.2; 1.0) *	0.7 (0.3; 1.9)	0.8 (0.3; 2.1)	1.2 (0.7; 1.9)	1.3 (0.8; 2.2)
2.1 (1.2; 3.6) *	2.3 (1.3; 4.1) **	1.4 (0.7; 2.6)	1.4 (0.7; 2.8)	1.0 (0.6; 1.5)	1.0 (0.7; 1.5)
0.5 (0.3; 0.9) *	0.4 (0.2; 0.8) *	0.5 (0.3; 1.0)	0.9 (0.2; 3.3)	1.2 (0.8; 1.8)	3.2 (1.4; 7.3) **
0.4 (0.1; 2.0)	0.3 (0.0; 3.1)	0.9 (0.5; 1.7)	1.0 (0.5; 1.8)	0.6 (0.4; 1.0) *	0.6 (0.4; 0.9) *
1.3 (0.5; 3.8)	1.6 (0.5; 5.2)	0.4 (0.3; 0.7) **	0.5 (0.3; 0.8) **	0.6 (0.4; 1.0) *	0.6 (0.4; 1.0) *
2.3 (0.8; 7.2)	1.7 (0.8; 7.4)	1.4 (0.9; 2.3)	2.1 (0.9; 5.3)	0.8 (0.5; 1.2)	1.1 (0.5; 2.4)
0.5 (0.3; 1.0)	0.5 (0.2; 1.0) *	1.0 (0.6; 1.5)	1.1 (0.7; 1.7)	0.8 (0.4; 1.5)	0.8 (0.4; 1.6)
2.1 (1.2; 3.5) **	2.2 (1.3; 3.8) **	0.7 (0.5; 1.0)	0.7 (0.4; 1.0)	0.6 (0.4; 1.0)	0.6 (0.4; 1.1)
0.3 (0.2; 0.5) ***	0.5 (0.2; 0.9) *	0.2 (0.2; 0.4) ***	0.4 (0.2; 0.7) **	0.1 (0.1; 0.2) ***	0.2 (0.1; 0.5) ***
0.1 (0.0; 0.7) *	0.1 (0.0; 0.5) *	1.6 (0.6; 4.7)	2.1 (0.8; 5.8)	0.4 (0.1; 3.5)	0.5 (0.1; 4.0)
3.2 (0.8; 12)	3.5 (0.9; 14) [7]	0.9 (0.4; 2.0)	1.2 (0.6; 2.7)	6.5 (1.5; 29) *	6.3 (1.3; 30) * [7]
0.3 (0.1; 1.2)	0.6 (0.1; 3.3)	0.3 (0.1; 0.7) **	0.2 (0.1; 1.1)	0.6 (0.2; 2.1)	2.4 (0.2; 27)

and the Netherlands highest frequencies were seen in the second season. Chlorpropham also had higher probabilities of detection in children compared to adults, which could be related to food consumption: children have a larger food intake per kg of bodyweight; also, biological elimination mechanisms may differ between children and adults (Arena et al., 2017).

Also, high detection rates in SPECIMEn were found for the biomarkers related to pirimiphos-methyl and tebuconazole, which are in good agreement with other targeted studies (Norén et al., 2020; Yusà et al., 2022). For these and other highly detected pesticides, no consistent effect across all countries of season or location was found, in contrast with expectations based on previous findings (Dereumeaux et al., 2020; Teysseire et al., 2020). Differences in study sites might occur due to different crop types. Detected differences are most likely influenced by a set of other covariates not included in the current regression models, such as diet. Dietary habits of participants may differ between the countries, locations within countries, seasons and age groups. Also, there might be differences in percentage of consumption of imported foods, and percentage of homegrown food consumption. These aspects make the variety of exposure due to diet complex and subject to many changes; therefore future work needs to focus on the actual consumed diet and their pesticide residue levels versus the suspect screening patterns. For example, the consumption of organic foods has been linked to lower exposure concentrations of several pesticides such as organophosphates and pyrethroids (Baudry et al., 2019; Hyland et al., 2019).

As a final remark on the detected pesticides, the SS methodology is only recently being applied in large scale studies to assess exposure to pesticides, and only a few HBM studies have previously applied SS approaches to complement for example targeted monitoring

programs (Gerona et al., 2018; Pellizzari et al., 2019; Plassmann et al., 2015; Wang et al., 2018). Within a cohort of approximately 300 pregnant women in France, Bonvallot et al. (Bonvallot et al., 2021) performed a large targeted pesticide exposure study which was extended with the application of suspect screening. This SS approach resulted in the most frequent detection of the parent pesticides azoxystrobin, fenpropimorph, phenmedipham, fluazifop(/butyl) and chlorpyrifos. From these, only the metabolites of fluazifop(/butyl) and chlorpyrifos overlapped and were also detected in the samples of the SPECIMEn study. This is due to among others differences in the suspect database, for example fenpropimorph was not included in our current study because it didn't contain Cl, Br or PO3 (Huber et al., 2022). Another interesting point is the difference in detection frequency between the TCPy and -CH2 biomarkers of chlorpyrifos (methyl) in Spain and Czech Republic. TCPy can originate from both parent compounds chlorpyrifos and methyl-chlorpyrifos. -CH2 is not a human metabolite of chlorpyrifos, and its detection is likely due to exposure through diet. Also, a higher sensitivity for -CH2 compared to TCPy at an individual instrument level might have contributed to this difference.

Co-occurrence

To explore the exposure to pesticide mixtures in the general population, it was assessed which parent pesticides co-occurred in the same urine sample. With the current work we were able to assess the probability of detection of 29 different parent pesticides simultaneously. In a large majority of the samples (84%) two or more different pesticides were detected. Our findings confirm the presence of mixtures and the necessity of assessing co-occurrent exposures, which is a topic of high concern in risk assessment (European Commission, 2020; Socianu et al., 2022; Luijten et al., 2022). The number of co-occurring pesticides typically ranged from 2 to 5, with a maximum of 13 different pesticides (2 urine samples). These two urine samples originate both from the Spanish non-agricultural area, one from a child of the first season, the other of an adult of the second season. Both individuals had a lower number of co-occurring pesticides during the other season, respectively 8 and 10 pesticides.

Based on the 14 most frequent co-occurrent pesticides, 44 different combinations could be made, resulting in highly individualized exposure profiles. The most common combination of acetamiprid with chlorpropham, occurred in just 3% (n=62) of the urine samples. Also, assessment of the co-occurrence patterns at country level (network analysis), did not result in strong relations and hardly any overlap across countries was seen. The underlying correlations between these probabilities of detection were also low, generally below 0.3. These results indicate that, qualitatively, pesticide mixtures might be highly variable between individuals. Nevertheless, the combined exposures may still pose a concern in terms of public health, especially when the different components of a chemical mixture share modes of action underlying toxicity (Rotter et al., 2018). Acetamiprid and chlorpropham seem to induce different toxicological effects (Arena et al., 2017; EFSA, 2016). Acetamiprid has been reported to mainly target the liver (EFSA, 2016), where it may cause, at least in rodents, oxidative stress leading to mitochondrial dysfunction (EL-Hak et al., 2022; Siwen Li et al., 2021). Exposure to chlorpropham rather leads to adverse effects on the hematopoietic system (Arena et al., 2017; Fujitani et al., 2000, 2004). Hemotoxicity such as hemolytic anemia, however, is considered to be due

to oxidative stress (Rokushima et al., 2007; Sivilotti, 2004). Chlorpropham belongs to the family of carbamates, which have been reported to induce oxidative stress in occupationally exposed workers (Saad-Hussein et al., 2022). The other frequently observed combination of co-occurring substances involved acetamiprid and tebuconazole. Tebuconazole is a fungicide that mainly affects the liver and the adrenals (EFSA, 2014). Additionally, it has been reported to induce oxidative stress in the liver and endocrine disruption including anti-androgenic effects (Taxvig et al., 2007; Yang et al., 2018). Follow-up studies involving a larger number of participants and targeted biomarkers for these substances are needed to better assess the composition of the relevant mixtures and associated health risks.

Strengths and limitations

With the uniform design of our study a comparison could be made across Europe between agricultural and non-agricultural areas, seasons, and adults and children. Close collaborations with partners from all five countries resulted in the harmonized data collection, with little loss to follow up. The collection of the urine samples required a minimal invasive protocol, reducing the burden of citizens to participate in this survey, and opening up possibilities for scaling-up studies in future endeavors. A novel SS approach was harmonized and standardized across laboratories, with extensive QA/QC procedures (Vitale et al., 2022). Such harmonization is crucial to compare SS data and results coming from different laboratories and countries, a situation that is often unavoidable in large-scale studies. The applied SS approach allows for a relatively cost-effective way of providing semi-quantitative measurements of a large number of pesticides. A clear strength of the SPECIMEn study is that information is obtained on (putative) internal exposure to pesticides not or hardly monitored before, and on simultaneous exposure to multiple pesticides. Across countries, different pesticides targeted to different controls on different crops are likely to have been applied at the time of sampling, of which the variation is covered with the SS approach. As such, this study should be seen as the first step towards a more complete assessment of the pesticide mixtures that the general European population is exposed to. Further in-depth screening of the collected data and further methodological developments will increase the number of biomarkers that can be detected in the collected urine samples. This allows an increasingly more complete coverage of all pesticides that are present in these samples as well as the detection of other biomarkers that might potentially interact with the pesticide mixture. Also, future more quantitative analysis of signal intensities will allow for a semi-quantitative interpretation, both in co-occurrence patterns and in the role of determinants of pesticide levels.

Although the current study yields many new insights and perspectives on pesticide occurrence and mixtures, several limitations need to be addressed. From an analytical methodology point of view, the suspect screening approach is less sensitive than targeted methods (Pourchet et al., 2020), and the data mining was biased towards halogenated pesticides (Huber et al., 2022). Despite harmonized methods between the involved laboratories, differences in sensitivity between the instruments used by the labs did occur, potentially introducing variability between countries which should be interpreted with care (Vitale et al., 2022, Huber et al., 2022). Importantly, data generated by the SS approach applied in the SPECIMEn can currently

not be related to urinary pesticide concentration levels in the traditional quantitative way as in targeted analysis, but rather as semi-quantitative intensities as indicators of exposure.

With regards to the sample collection, it should be kept in mind that samples of the second season were collected during the COVID-19 pandemic, while the first sample collection was not affected by the pandemic. Activity patterns or diet of participants might have been altered, and differences between seasons should be interpreted with caution. Also, the different seasons cannot be interpreted as 'non-spraying' and 'spraying', since the timing of the actual spraying activities (and spraying techniques) most likely differed between countries and crop types. Since the applied study design was not timed with an actual spraying activity, the detected exposures might be an underestimation as compared to what has been reported in the literature (Dereumeaux et al., 2020; Teysseire et al., 2020). Agricultural areas were selected based on national databases on land-use (see Supplementary Material F for a description of the area selection per country), due to which the application of pesticides during the time of sampling could not be confirmed. Within SPECIMEn, only first morning void urines were collected. Due to the rapid excretion of many pesticides, the detected pesticides in the morning voids likely do not reflect the total daily exposure (Adela Jing Li et al., 2019; Scher et al., 2007). Finally, with respect to the performed logistic regression models, no correction for multiple testing was performed, since we wanted to detect any possible effects, accepting the risk of false-positive results. The inclusion of both location and season could have led to an over-correction, especially since no difference between seasons at the non-agricultural locations would be expected due to any spraying activity (although diet might still differ between the seasons).

CONCLUSIONS

The current survey demonstrates the feasibility of conducting a harmonized pan-European sample collection combined with suspect screening (SS) to provide insight in the co-occurrence of pesticide mixtures in European agricultural areas. The application of a novel LC-HRMS based SS approach harmonized between different laboratories, resulted in detection of 40 biomarkers related to 29 parent pesticides with high levels of confidence. Some effects of living close to agricultural fields or season were detected, but these effects were not common at a European level. This study is a first step in addressing pesticide mixture exposure under real-life conditions. Combined with a suspect screening approach, this approach is a promising strategy for pesticide mixture risk assessment in the European population, that can guide the prioritization of pesticide (metabolites) to be measured using quantitative targeted methods.

Conflict of Interest

Declarations of interest: none.

Credit author statement

Conceptualization and design (**IO, JV, EL, RV, JA**); Investigation (**IO, JV, EL, PČ, LŠ, OM, TS, SK, IM, ZM, LA, OP, SF, CC, SP**); Analytical methodology (**JA, CH, AL, OP, SF, MK, LD, KW, RN, HM, CM, JK, BG, NL**) ; Formal analysis (**IO, JV**); Writing - Original Draft (**IO, JV, EL**); Writing – Review and Editing (all authors); Visualization (**IO**); Supervision (**ML, RV**); Project administration (**IO, JV, EL**); All authors read and approved the final manuscript.

Acknowledgement

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 733032 HBM4EU, No 857560 and No 857340. P.Č., O.M. & L.Š. acknowledge the RECETOX research infrastructure supported by the Ministry of Education, Youth and Sports of the Czech Republic (LM2018121) and the Ministry of Education, Youth and Sports of the Czech Republic (CZ.02.1.01/0.0/0.0/17_043/0009632). This publication reflects only the author's view and the European Commission is not responsible for any use that may be made of the information it contains. BG was supported by the Margarita Salas postdoctoral contract MGS/2021/25 (UP2021-021) financed by the European Union-NextGenerationEU.

Abbreviations

HBM	Human Biomonitoring
HBM4EU	European Human Biomonitoring Initiative
LC-HRMS	Liquid chromatography coupled to High Resolution Mass Spectrometry
SPECIMEn	Survey on PEstiCIde Mixtures in Europe
SS	Suspect Screening

REFERENCES

Allema, B., Hoogendoorn, M., van Beek, J., & Leendertse, P. (2017). *Neonicotinoids in European agriculture*. *CLM-937*.

Andra, S. S., Austin, C., Patel, D., Dolios, G., Awawda, M., & Arora, M. (2017). Trends in the application of high-resolution mass spectrometry for human biomonitoring: An analytical primer to studying the environmental chemical space of the human exposome. *Environment International*, *100*, 32–61. https://doi.org/10.1016/j.envint.2016.11.026

Arena, M., Auteri, D., Barmaz, S., Bellisai, G., Brancato, A., Brocca, D., … Villamar-Bouza, L. (2017). Peer review of the pesticide risk assessment of the active substance chlorpropham. *EFSA Journal*, *15*(7). https://doi.org/10.2903/j.efsa.2017.4903

Baudry, J., Debrauwer, L., Durand, G., Limon, G., Delcambre, A., Vidal, R., … Kesse-Guyot, E. (2019). Urinary pesticide concentrations in French adults with low and high organic food consumption: results from the general population-based NutriNet-Santé. *Journal of Exposure Science & Environmental Epidemiology*, *29*(3), 366–378. https://doi.org/10.1038/s41370-018-0062-9

Bonvallot, N., Jamin, E. L., Regnaut, L., Chevrier, C., Martin, J.-F., Mercier, F., … Le Bot, B. (2021). Suspect screening and targeted analyses: Two complementary approaches to characterize human exposure to pesticides. *Science of The Total Environment*, *786*, 147499. https://doi.org/10.1016/j.scitotenv.2021.147499

Crépet, A., Vanacker, M., Sprong, C., de Boer, W., Blaznik, U., Kennedy, M., … van Klaveren, J. (2019). Selecting mixtures on the basis of dietary exposure and hazard data: application to pesticide exposure in the European population in relation to steatosis. *International Journal of Hygiene and Environmental Health*, *222*(2), 291–306. https://doi.org/10.1016/j.ijheh.2018.12.002

Damalas, C. A., & Eleftherohorinos, I. G. (2011). Pesticide Exposure, Safety Issues, and Risk Assessment Indicators. *International Journal of Environmental Research and Public Health*, *8*(5), 1402–1419. https://doi.org/10.3390/ijerph8051402

Dereumeaux, C., Fillol, C., Quenel, P., & Denys, S. (2020). Pesticide exposures for residents living close to agricultural lands: A review. *Environment International*, *134*, 105210. https://doi.org/10.1016/j.envint.2019.105210

Deziel, N. C., Friesen, M. C., Hoppin, J. A., Hines, C. J., Thomas, K., & Freeman, L. E. B. (2015). A Review of Nonoccupational Pathways for Pesticide Exposure in Women Living in Agricultural Areas. *Environmental Health Perspectives*, *123*(6), 515–524. https://doi.org/10.1289/ehp.1408273

EFSA. (2014). Conclusion on the peer review of the pesticide risk assessment of the active substance tebuconazole. *EFSA Journal*, *12*(1). https://doi.org/10.2903/j.efsa.2014.3485

EFSA. (2016). Peer review of the pesticide risk assessment of the active substance acetamiprid. *EFSA Journal*, *14*(11). https://doi.org/10.2903/j.efsa.2016.4610

EL-Hak, H. N. G., Al-Eisa, R. A., Ryad, L., Halawa, E., & El-Shenawy, N. S. (2022). Mechanisms and histopathological impacts of acetamiprid and azoxystrobin in male rats. *Environmental Science and Pollution Research*, *29*(28), 43114–43125. https://doi.org/10.1007/s11356-021-18331-3

Eskenazi, B., Bradman, A., & Castorina, R. (1999). Exposures of children to organophosphate pesticides and their potential adverse health effects. *Environmental Health Perspectives*, *107*(suppl 3), 409–419. https://doi.org/10.1289/ehp.99107s3409

EU Database Pest. (2022). EU Database Pest. Retrieved June 21, 2022, from https://ec.europa.eu/food/plant/pesticides/eu-pesticides-database/active-substances/?event=search.as

European Commission. (2020). *Progress Report on the Assessment and Management of Combined Exposure to Multiple Chemicals (Chemical Mixtures) and Associated Risks* (Staff Working Document SWD 250 Final). Retrieved from https://ec.europa.eu/environment/pdf/chemicals/2020/10/SWD_mixtures.pdf

Figueiredo, D. M., Duyzer, J., Huss, A., Krop, E. J. M., Gerritsen-Ebben, M. G., Gooijer, Y., & Vermeulen, R. C. H. (2021). Spatio-temporal variation of outdoor and indoor pesticide air concentrations in homes near agricultural fields. *Atmospheric Environment*, *262*, 118612. https://doi.org/10.1016/j.atmosenv.2021.118612

Fujitani, T., Tada, Y., Fujii, A., Kimura, M., & Yoneyama, M. (2000). Subchronic toxicity of chlorpropham (CIPC) in ICR mice. *Food and Chemical Toxicology : An International Journal Published for the British Industrial Biological Research Association*, *38*(7), 617–625. https://doi.org/10.1016/s0278-6915(00)00043-0

Fujitani, T., Tada, Y., & Yoneyama, M. (2004). Chlorpropham-induced splenotoxicity and its recovery in rats. *Food and Chemical Toxicology : An International Journal Published for the British Industrial Biological Research Association*, *42*(9), 1469–1477. https://doi.org/10.1016/j.fct.2004.04.008

Gerona, R. R., Schwartz, J. M., Pan, J., Friesen, M. M., Lin, T., & Woodruff, T. J. (2018). Suspect screening of maternal serum to identify new environmental chemical biomonitoring targets using liquid chromatography–quadrupole time-of-flight mass spectrometry. *Journal of Exposure Science & Environmental Epidemiology*, *28*(2), 101–108. https://doi.org/10.1038/jes.2017.28

Gooijer, Y. M., Hoftijser, G. W., Lageschaar, L. C. C., Oerlemans, A., Scheepers, P. T. J., Kivits, C. M., … Sauer, P. J. J. (2019). *Research on Exposure of Residents to Pesticides in The Netherlands OBO Flower Bulbs*. Retrieved from https://edepot.wur.nl/475219

Hyland, C., Bradman, A., Gerona, R., Patton, S., Zakharevich, I., Gunier, R. B., & Klein, K. (2019). Organic diet intervention significantly reduces urinary pesticide levels in U.S. children and adults. *Environmental Research*, *171*, 568–575. https://doi.org/10.1016/j.envres.2019.01.024

Kim, K.-H., Kabir, E., & Jahan, S. A. (2017). Exposure to pesticides and the associated human health effects. *Science of The Total Environment*, *575*, 525–535. https://doi.org/10.1016/j.scitotenv.2016.09.009

Li, A. J., Martinez-Moral, M.-P., & Kannan, K. (2019). Temporal variability in urinary pesticide concentrations in repeated-spot and first-morning-void samples and its association with oxidative stress in healthy individuals. *Environment International*, *130*, 104904. https://doi.org/10.1016/j.envint.2019.104904

Li, S., Cao, Y., Pan, Q., Xiao, Y., Wang, Y., Wang, X., … Ran, B. (2021). Neonicotinoid insecticides triggers mitochondrial bioenergetic dysfunction via manipulating ROS-calcium influx pathway in the liver. *Ecotoxicology and Environmental Safety*, *224*, 112690. https://doi.org/10.1016/j.ecoenv.2021.112690

Norén, E., Lindh, C., Rylander, L., Glynn, A., Axelsson, J., Littorin, M., … Nielsen, C. (2020). Concentrations and temporal trends in pesticide biomarkers in urine of Swedish adolescents, 2000–2017. *Journal of Exposure Science & Environmental Epidemiology*, *30*(4), 756–767. https://doi.org/10.1038/s41370-020-0212-8

Ntzani, E. E., Ntritsos G, C. M., Evangelou, E., & Tzoulaki, I. (2013). Literature review on epidemiological studies linking exposure to pesticides and health effects. *EFSA Supporting Publications*, *10*(10). https://doi.org/10.2903/sp.efsa.2013.EN-497

Oerlemans, A., Figueiredo, D. M., Mol, J. G. J., Nijssen, R., Anzion, R. B. M., van Dael, M. F. P., … Scheepers, P. T. J. (2021). Personal exposure assessment of pesticides in residents: The association between hand wipes and urinary biomarkers. *Environmental Research*, *199*, 111282. https://doi.org/10.1016/j.envres.2021.111282

Pellizzari, E. D., Woodruff, T. J., Boyles, R. R., Kannan, K., Beamer, P. I., Buckley, J. P., … Bennett, D. H. (2019). Identifying and Prioritizing Chemicals with Uncertain Burden of Exposure: Opportunities for Biomonitoring and Health-Related Research. *Environmental Health Perspectives*, *127*(12), EHP5133. https://doi.org/10.1289/EHP5133

Plassmann, M. M., Brack, W., & Krauss, M. (2015). Extending analysis of environmental pollutants in human urine towards screening for suspected compounds. *Journal of Chromatography A*, *1394*, 18–25. https://doi.org/10.1016/j.chroma.2015.03.040

Pons, P., & Latapy, M. (2005). Computing Communities in Large Networks Using Random Walks (pp. 284–293). https://doi.org/10.1007/11569596_31

Pourchet, M., Debrauwer, L., Klanova, J., Price, E. J., Covaci, A., Caballero-Casero, N., … Antignac, J.-P. (2020). Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: From promises to challenges and harmonisation issues. *Environment International*, *139*, 105545. https://doi.org/10.1016/j.envint.2020.105545

Rokushima, M., Omi, K., Imura, K., Araki, A., Furukawa, N., Itoh, F., … Ishizaki, J. (2007). Toxicogenomics of Drug-Induced Hemolytic Anemia by Analyzing Gene Expression Profiles in the Spleen. *Toxicological Sciences*, *100*(1), 290–302. https://doi.org/10.1093/toxsci/kfm216

Rotter, S., Beronius, A., Boobis, A. R., Hanberg, A., van Klaveren, J., Luijten, M., … Solecki, R. (2018). Overview on legislation and scientific approaches for risk assessment of combined exposure to multiple chemicals: the potential EuroMix contribution. *Critical Reviews in Toxicology*, *48*(9), 796–814. https://doi.org/10.1080/10408444.2018.1541964

Saad-Hussein, A., Shahy, E. M., Ibrahim, K. S., Mahdy-Abdallah, H., Taha, M. M., Abdel-Shafy, E. A., & Shaban, E. E. (2022). Influence of GSTM1, T1 genes polymorphisms on oxidative stress and liver enzymes in rural and urban pesticides-exposed workers. *Archives of Environmental & Occupational Health*, *77*(10), 800–808. https://doi.org/10.1080/19338244.2021.2025024

Sapbamrer, R., & Hongsibsong, S. (2019). Effects of prenatal and postnatal exposure to organophosphate pesticides on child neurodevelopment in different age groups: a systematic review. *Environmental Science and Pollution Research*, *26*(18), 18267–18290. https://doi.org/10.1007/s11356-019-05126-w

Scher, D. P., Alexander, B. H., Adgate, J. L., Eberly, L. E., Mandel, J. S., Acquavella, J. F., … Brzak, K. A. (2007). Agreement of pesticide biomarkers between morning void and 24-h urine samples from farmers and their children. *Journal of Exposure Science & Environmental Epidemiology*, *17*(4), 350–357. https://doi.org/10.1038/sj.jes.7500505

Schymanski, E. L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H. P., & Hollender, J. (2014). Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environmental Science & Technology*, *48*(4), 2097–2098. https://doi.org/10.1021/es5002105

Sivilotti, M. L. A. (2004). Oxidant stress and haemolysis of the human erythrocyte. *Toxicological Reviews*, *23*(3), 169–188. https://doi.org/10.2165/00139709-200423030-00004

Socianu, S., Bopp, S. K., Govarts, E., Gilles, L., Buekers, J., Kolossa-Gehring, M., … Franco, A. (2022). Chemical Mixtures in the EU Population: Composition and Potential Risks. *International Journal of Environmental Research and Public Health*, *19*(10), 6121. https://doi.org/10.3390/ijerph19106121

Taxvig, C., Hass, U., Axelstad, M., Dalgaard, M., Boberg, J., Andeasen, H. R., & Vinggaard, A. M. (2007). Endocrine-disrupting activities in vivo of the fungicides tebuconazole and epoxiconazole. *Toxicological Sciences : An Official Journal of the Society of Toxicology*, *100*(2), 464–473. https://doi.org/10.1093/toxsci/kfm227

Teysseire, R., Manangama, G., Baldi, I., Carles, C., Brochard, P., Bedos, C., & Delva, F. (2020). Assessment of residential exposures to agricultural pesticides: A scoping review. *PLOS ONE*, *15*(4), e0232258. https://doi.org/10.1371/journal.pone.0232258

Teysseire, R., Manangama, G., Baldi, I., Carles, C., Brochard, P., Bedos, C., & Delva, F. (2021). Determinants of non-dietary exposure to agricultural pesticides in populations living close to fields: A systematic review. *Science of The Total Environment*, *761*, 143294. https://doi.org/10.1016/j.scitotenv.2020.143294

van Borkulo, C. D., Borsboom, D., Epskamp, S., Blanken, T. F., Boschloo, L., Schoevers, R. A., & Waldorp, L. J. (2015). A new method for constructing networks from binary data. *Scientific Reports*, *4*(1), 5918. https://doi.org/10.1038/srep05918

Vitale, C. M., Lommen, A., Huber, C., Wagner, K., Garlito Molina, B., Nijssen, R., … Antignac, J.-P. (2022). Harmonized Quality Assurance/Quality Control Provisions for Nontargeted Measurement of Urinary Pesticide Biomarkers in the HBM4EU Multisite SPECIMEn Study. *Analytical Chemistry*. https://doi.org/10.1021/acs.analchem.2c00061

Wang, A., Gerona, R. R., Schwartz, J. M., Lin, T., Sirota, M., Morello-Frosch, R., & Woodruff, T. J. (2018). A Suspect Screening Method for Characterizing Multiple Chemical Exposures among a Demographically Diverse Population of Pregnant Women in San Francisco. *Environmental Health Perspectives*, *126*(7), 077009. https://doi.org/10.1289/EHP2920

Willenbockel, C. T., Prinz, J., Dietrich, S., Marx-Stoelting, P., Weikert, C., Tralau, T., & Niemann, L. (2022). A Critical Scoping Review of Pesticide Exposure Biomonitoring Studies in Overhead Cultures. *Toxics*, *10*(4), 170. https://doi.org/10.3390/toxics10040170

Yang, J.-D., Liu, S.-H., Liao, M.-H., Chen, R.-M., Liu, P.-Y., & Ueng, T.-H. (2018). Effects of tebuconazole on cytochrome P450 enzymes, oxidative stress, and endocrine disruption in male rats. *Environmental Toxicology*. https://doi.org/10.1002/tox.22575

Yusà, V., F. Fernández, S., Dualde, P., López, A., Lacomba, I., & Coscollà, C. (2022). Exposure to non-persistent pesticides in the Spanish population using biomonitoring: A review. *Environmental Research*, *205*, 112437. https://doi.org/10.1016/j.envres.2021.112437

SUPPLEMENTARY MATERIAL

Table A. *Specific sampling dates, per study site and season*

		Season 1		***Season 2***	
Area		**Start**	**End**	**Start**	**End**
Spain					
	Agricultural	07/11/2019	20/12/2019	01/09/2020	02/10/2020
	Non-agricultural	05/11/2019	19/12/2019	01/09/2020	05/10/2020
Latvia					
	Agricultural	18/02/2020	31/03/2020	02/06/2020	18/06/2020
	Non-agricultural	18/02/2020	31/03/2020	02/06/2020	18/06/2020
Hungary					
	Agricultural	29/01/2020	10/02/2020	07/09/2020	16/09/2020
	Non-agricultural	11/02/2020	18/02/2020	16/09/2020	17/09/2020
Czech Republic					
	Agricultural	14/1/2020	13/3/2020	26/5/2020	30/7/2020
	Non-agricultural	14/1/2020	13/3/2020	26/5/2020	30/7/2020
The Netherlands					
	Agricultural	22/01/2020	06/03/2020	02/06/2020	24/06/2020
	Non-agricultural	22/01/2020	06/03/2020	02/06/2020	24/06/2020

Table B. *Descriptive characteristics of the SPECIMEn study participants based on the questionnaire, by study site and location.*

Study Site	*Spain*		*Latvia*		*Hungary*		*Czech Republic*		*Netherlands*	
Area	Agricultural	Non-agricultural	Agricultural	Non-agricultural	Agricultural	Non-agricultural	Agricultural	Non-agricultural	Agricultural	Non-agricultural
Smoking status adult1, %										
No-current smoker	65.4	73.6	88.0	82.4	54.9	78.8	84.3	91.7	94.5	100.0
Household income, % of country average1										
< 25%	7.7	0	14.0	9.8	27.4	17.3	15.7	16.7	1.8	0
25-50%	5.8	0	0	0	39.2	19.2	39.2	26.7	5.5	6.0
50-75%	17.3	3.8	0	0	13.7	7.7	35.3	33.3	49.1	44.0
>75%	57.7	75.5	74.0	70.6	5.9	44.4	9.8	21.7	20.0	44.0
Don't Know/NA	11.5	20.8	12.0	19.6	13.7	9.3	0	1.7	23.6	6.0
Professional contact with pesticides in the past month, n adults										
Season 1	0	1	2	0	0	0	1	0	0	0
Season 2	2	0	4	1	2	1	3	1	0	0
Having other adult household member(s) who had professional contact with pesticides, n adults	1	0	16	2	9	1	0	0	1	0
Usage of any type of products for treating the plants in the garden up to 3 days prior to sample collection, n adults										
Season 1	0	2	1	1	0	2	1	0	1	0
Season 2	4	2	4	2	1	2	6	4	4	2
Usage of any type of products for treating the plants inside the house up to 3 days prior to sample collection, n adults										
Season 1	2	0	2	4	0	2	0	1	1	3
Season 2	2	0	3	3	17	4	2	0	2	2
Usage of external antiparasitic treatments for pets in the 3 days prior to sample collection, n adults										
Season 1	2	2	0	1	2	11	1	1	1	0
Season 2	1	2	4	1	4	5	0	1	6	1
Usage of insect repellent or antiparasitic human products in the 3 days prior to sample collection, n adults										
Season 1	6	1	0	3	2	2	2	1	0	2
Season 2	25	6	5	4	4	2	6	8	5	0

1. 50% is country mean average income

Table C. *Annotated pesticide-related compounds (parent pesticides and/or metabolites) of confidence levels 1 - 5 (p = 95) and their overall detection frequency (%) per study site.*

ID	Pesticide type[1]	Parent pesticide	Pesticide (metabolite) annotation[2]	precursor ion	exact m/z	RT[3] urine [min]	Conf. level[4]	Overall Detection Frequency (%) ES[5]	LV	HU	CZ	NL
P1	H	2,4-D	Parent compound	[M-H]-	218.9623	9.93	**1**	4.07	0	2.2	2.71	0
P2_a	I	Acetamiprid	-CH2	[M-H]-	207.0443	8.71	**1**	98.56	32.84	94.13	98.19	93.29
P2_b			-CH2	[M+H]+	209.0589	8.55	4	81.82	10.95	45.23	41.18	47.00
P2_c			Parent compound	[M+H]+	223.0745	8.67	4	1.44	0	0.49	0	0.72
P3_a	F	Ametoctradin	-C2H6 +2O	[M+H]+	278.1612	9.47	**1**	5.02	2.74	1.22	4.75	2.88
P3_b			-C2H6 +2O	[M-H]-	276.1466	8.17	5	0.72	0.5	0.49	0.45	2.16
P4	I, Ac	Bifenthrin/Cyhalothrin	F3CCA + C6H8O6	[M-H]-	417.0570	11.95	4	40.43	3.23	7.09	3.62	13.91
P5_a	F	Boscalid	+O +SO3	[M-H]-	436.9771	10.26	2b	35.65	18.41	3.91	22.85	32.85
P5_b			+O +SO3	[M+H]+	438.9917	10.49	2b	7.18	0	0	0.45	0.24
P5_c			+O (M510F01)	[M-H]-	357.0203	11.89	4	0.48	0	0	0	0
P5_d			+O (M510F01)	[M+H]+	359.0349	11.69	4	0.48	0	0	0	0
P6	I	Chlorantraniliprole	+O	[M-H]-	497.9564	12.67	2b	3.83	0.25	0.24	0	0.24
P7_a	Ac	Chloropropylate	-C3H6	[M-H]-	294.9934	12.93	4	0	0	0	0.23	0
P7_b			-C3H6 -CO2	[M-H]-	251.0036	12.93	4	0.24	0	0	0	0
P8_a	H, GR	Chlorpropham	+O +SO3 (4-HSA)	[M-H]-	308.0003	9.5	**1**	55.74	31.59	31.05	34.16	75.06
P8_b			-C4H6O +SO3	[M-H]-	221.9633	6.19	3	29.19	32.09	21.03	28.05	63.07
P8_c			+2O +SO3	[M-H]-	323.9950	7.5	3	7.66	6.97	9.78	9.28	26.86
P8_d			+O +C6H8O6	[M-H]-	404.0757	8.55	4	15.55	15.42	12.96	14.03	44.84
P8_e			+O	[M-H]-	228.0433	10.97	4	1.2	0	0.73	1.36	9.35
P9_a	I	Chlorpyrifos (/methyl)	TCPy	[M-H]-	195.9129	10.1	**1**	1.67	0	0.24	0.23	0.24
P9_b			-CH2	[M-H]-	305.8723	10.72	**1**	36.12	0	6.85	21.72	6.47
P9_c			TCPy+C6H8O6	[M-H]-	371.9450	8.38	4	50.00	0	2.69	13.35	7.19
P10	H	Clopyralid	Parent compound	[M-H]-	189.9465	3.5	**1**	0.96	0	0	1.36	0.72
P11_a	I	Clothianidin (can come from thiamethoxam)	Parent compound	[M-H]-	248.0015	8.09	**1**	34.45	1.74	21.52	24.66	19.42
P11_b			-NO2 +H	[M+H]+	205.0309	5.77	**1**	0.48	0	0.24	0	0.24
P11_c			-CH2	[M-H]-	233.9858	7.51	2b	21.05	0.75	9.78	6.56	3.12
P11_d			Parent compound	[M+H]+	250.0160	8.08	4	1.67	0	3.18	0	2.16
P12_a	I	Cypermethrin, cyfluthrin, permethrin, transfluthrin	DCCA	[M-H]-	206.9985	10.73	**1**	0.48	0	0	0	0
P12_b			DCCA+C6H8O6	[M-H]-	383.0306	10.95	4	84.93	9.2	14.67	25.34	48.20
P13_a	F	Cyprodinil	+O +SO3	[M-H]-	320.0710	11.87	2b	14.11	7.71	2.69	10.18	26.38
P13_b			+2O +SO3	[M-H]-	336.0660	9.22	3	9.09	4.98	1.71	7.47	22.78
P14	I	Deltamethrin	DBCA+C6H8O6	[M-H]-	470.9296	11.43	4	76.32	0.75	7.33	9.5	21.82
P15_a	H, Al	Diuron	-CH2 -CH2	[M-H]-	202.9786	12	4	5.5	1	0.24	1.13	0.48
P15_b			-CH2	[M-H]-	216.9942	12.45	4	1.2	0.25	0	0	0
P15_c			-CH2	[M+H]+	219.0084	12.14	4	0.24	0	0	0	0

ID	Pesticide type[1]	Parent pesticide	Pesticide (metabolite) annotation[2]	precursor ion	exact m/z	RT[3] urine [min]	Conf. level[4]	Overall Detection Frequency (%)				
								ES[5]	LV	HU	CZ	NL
P16	F	Fenhexamid	+O +C6H8O6	[M+NH3]+	511.1244	9.34	3	0.96	1	1.22	2.49	6.71
P17_a	I, Ac	Fipronil	Parent compound	[M-H]-	434.9310	15.02	4	0.96	0	0	0	0
P17_b			+O	[M-H]-	450.9260	15.43	4	3.59	0.5	0	0	0
P18_a	I	Flonicamid	Parent compound	[M-H]-	228.0397	6.9	**1**	1.67	0.75	1.96	2.71	5.76
P18_b			-C2HN	[M+H]+	191.0427	6.1	2b	15.07	0.25	27.38	0.23	57.31
P18_c			Parent compound	[M+H]+	230.054	6.8	4	1.44	0	0.98	1.36	3.6
P19_a	H	Fluazifop	Parent compound	[M-H]-	326.0647	11.74	**1**	19.86	2.49	11.00	18.33	21.10
P19_b			Parent compound	[M+H]+	328.079	13.57	**1**	8.13	1.49	4.89	5.20	8.15
P20	F	Fludioxonil	+O +C6H8O6	[M-H]-	439.0609	11.81	2b	16.27	14.68	1.96	14.48	26.86
P21_a	F	Fluopyram	+O +SO3	[M-H]-	490.9908	12.68	2b	3.59	0.5	0.24	1.13	0.96
P21_b			+O +C6H8O6	[M+H]+	589.0807	13.08	2b	2.39	0.75	0.49	3.17	4.8
P21_c			-2H	[M+H]+	395.0385	13.07	2b	10.77	6.72	0.49	3.39	3.12
P22_a	I	Flupyradifurone	Parent compound	[M+H]+	289.0557	8.79	**1**	2.63	0.25	0.24	0.68	2.16
P22_b			-C2H2F2	[M+H]+	225.0425	7.54	4	1.67	0	0.24	0.23	3.12
P23	H	Fluroxypyr	Parent compound	[M+H]+	254.973	10.47	4	0.24	0	0	0	0
P24	F	Flutolanil	-C3H6 +O +SO3	[M-H]-	376.0108	8.18	3	14.11	0	4.65	0	0.24
P25_a	I, Ac	Fluvalinate	-C14H9NO	[M-H]-	294.0514	13.94	2b	0.96	0	0.73	0.23	0
P25_b			-C14H9NO +O	[M-H]-	310.0463	12.78	3	0.72	0	0.49	0	0
P25_c			-C14H9NO	[M+H]+	296.066	14.35	4	0.96	0	0.49	0	0
P26	H	Haloxyfop	-CH2	[M-H]-	360.026	13.39	4	60.53	3.23	2.69	34.39	21.34
P27_a	F	Imazalil	+C6H8O6	[M+H]+	473.0869	11.52	2b	19.38	10.70	8.31	4.52	4.56
P27_b			+H2O2 +C6H8O6	[M+H]+	507.0946	9.15	3	14.35	8.21	4.16	1.81	3.6
P28_a	I	Imidacloprid	-NO2 +H	[M+H]+	211.0739	6.01	**1**	17.46	1.74	4.16	0.68	9.35
P28_b			Parent compound	[M+H]+	256.0596	8.04	4	5.02	0	2.44	3.85	3.84
P28_c			+O	[M+H]+	272.054	7.48	4	10.53	0.75	1.47	2.71	2.4
P28_d			-2H	[M+H]+	254.0439	7.3	4	8.37	0.25	0.98	2.94	2.88
P29	F	Iprodione	-C3H6 (RP32490)	[M-H]-	285.9786	12.93	4	5.02	0	0.24	1.58	2.64
P30_a	H	MCPA	+O	[M-H]-	215.0117	7.57	3	14.59	0.75	1.22	5.66	12.47
P30_b			Parent compound	[M-H]-	199.0167	9.95	4	0.48	0	0	0.45	0.96
P31	F	Myclobutanil	-H2 +2O	[M-H]-	317.0811	9	3	7.18	0.50	0.24	4.30	0.96
P32_a	F	Penconazole	+O +C6H8O6	[M+H]+	476.0982	11.45	2b	6.46	1.74	2.2	2.04	2.4
P32_b			-2H +2O	[M+H]+	314.0457	11.91	3	2.63	0.25	0.73	1.13	1.68
P33	F, H, I, M, GR	Pentachlorophenol	in source fragment of +SO3	[M-H]-	264.8368	13.19	4	3.11	0	2.44	3.85	0.24

Table C. *Continued*

ID	Pesticide type[1]	Parent pesticide	Pesticide (metabolite) annotation[2]	precursor ion	exact m/z	RT[3] urine [min]	Conf. level[4]	Overall Detection Frequency (%) ES[5]	LV	HU	CZ	NL
P34_a	I, Ac	Pirimiphos-methyl	-CH2	[M-H]-	290.0734	10.75	**1**	85.17	10.20	6.60	23.98	47.72
P34_b			-CH2 -C2H4	[M-H]-	262.0422	7.47	5	16.75	0	0	0.23	4.08
P34_c			-CH2 -C2H4	[M+H]+	264.0564	6.22	5	0	0.25	0	0	0.48
P35_a	F	Propamocarb	Parent compound	[M+H]+	189.1597	6.00	**1**	9.57	1	11.49	4.98	23.26
P35_b			+O	[M+H]+	205.1546	6.45	2b	20.81	5.47	18.34	12.67	42.69
P36_a	F	Propiconazole	-C5H10O +H2 +C6H8O6	[M-H]-	432.0371	9.00	3	2.39	0	0.98	0	1.2
P36_b			-C5H10O (CGA91304)	[M-H]-	253.9888	12.30	4	0	0	0	0	0.24
P37	H	Propyzamide	+H2O3	[M-H]-	304.0143	11.36	2b	8.61	0	0.49	0.9	0.96
P38_a	F	Pyrimethanil	+O +SO3	[M-H]-	294.0556	9.15	2b	26.79	14.43	4.89	21.95	31.89
P38_b			+O	[M+H]+	216.1133	11.69	2b	0.72	0	2.69	0	0.48
P39	G	Quinmerac	Parent compound	[M-H]-	220.0171	8.54	4	86.12	22.64	25.92	74.89	23.26
P40_a	F	Tebuconazole	-2H +2O	[M-H]-	336.1124	12.18	2b	71.29	5.47	25.18	52.26	35.97
P40_b			+O +C6H8O6	[M+H]+	500.1794	12.71	3	41.15	17.16	30.56	23.08	13.91
P41_a	F	Thiabendazole	+O +C6H8O6	[M-H]-	392.0551	5.96	2b	0	0.75	0.24	0	0.48
P41_b			+O (5-hydroxy)	[M+H]+	218.0381	6.80	5	2.15	1.49	1.47	0	3.36
P42_a	I	Thiacloprid	+O	[M-H]-	267.0107	9.19	2b	8.37	0.75	2.93	7.92	4.56
P42_b			+H2 +O	[M-H]-	269.0271	7.05	4	3.11	0.5	0.49	0.9	1.92
P43_a	I	Thiamethoxam	Parent compound	[M+H]+	292.0262	7.10	**1**	0.72	0	2.44	0	0.48
P43_b			-NO2 +H	[M+H]+	247.0413	6.20	**1**	23.44	0	15.16	0	0.24
P44	F	Tolclofos-methyl	-CH2	[M-H]-	284.9309	10.31	4	0	0.25	0	0.45	0.24
P45_a	Af, Ab	Triclosan	+C6H8O6	[M-H]-	462.9759	13.23	**1**	84.69	16.17	24.45	46.15	12.71
P45_b			+O +C6H8O6	[M-H]-	478.9709	9.40	3	4.78	0.75	0.73	1.13	0.48
P45_c			+SO3	[M-H]-	366.9007	13.89	4	3.83	0.25	0.73	0.68	0.48
P45_d			Parent compound	[M-H]-	286.9439	16.12	4	2.15	0.5	1.22	0	0.96
P45_e			+C6H8O6	[M+NH3]+	482.0171	14.02	4	28.95	2.24	12.71	14.25	6.24
P46_a	F	Trifloxystrobin	-CH2 -CH2	[M-H]-	379.0911	13.07	2b	0.72	0.5	0	3.62	3.84
P46_b			-CH2 (CGA 321113)	[M+H]+	395.1213	14.88	5	0.24	0	0.24	2.04	0.24

1. H: Herbicide, F: Fungicide, I: Insecticide, GR: Plant Growth Regulator, Ac: Acaricide, M: molluscide, Al: Algicide, Ab: antibacterial, Af: antifungal,
2. Metabolite annotation: "-CH2" means the molecular formula of the metabolite is that of the parent minus CH2 (corresponding to demethylation). Similarly, "+O" means the metabolite is the parent compound plus one oxygen atom (hydroxylation). "+SO3" and "+C6H8O6" indicate sulfation and glucuronidation, respectively.
3. RT: Retention Time
4. Schymanski confidence level, ranging from 1 to 5, (Schymanski et al., 2014)
5. ES: Spain, LV: Latvia, HU: Hungary, CZ: Czech Republic, NL: the Netherlands

Table D. *Frequency (number of urine samples, n=2,088) of co-occurrent parent pesticides; the most frequent (in 2 or more urine samples) co-occurrences are shown. Different study sites are indicated by colors (CZ=Czech Republic, ES=Spain, HU=Hungary, LV=Latvia, NL=Netherlands), the detection frequency (%) of the listed parent pesticides is given on the right. Pesticides are co-occurring in the same sample when both have a black connected dot. Multiple metabolites and/or parent compounds related to the same parent pesticide were considered as one.*

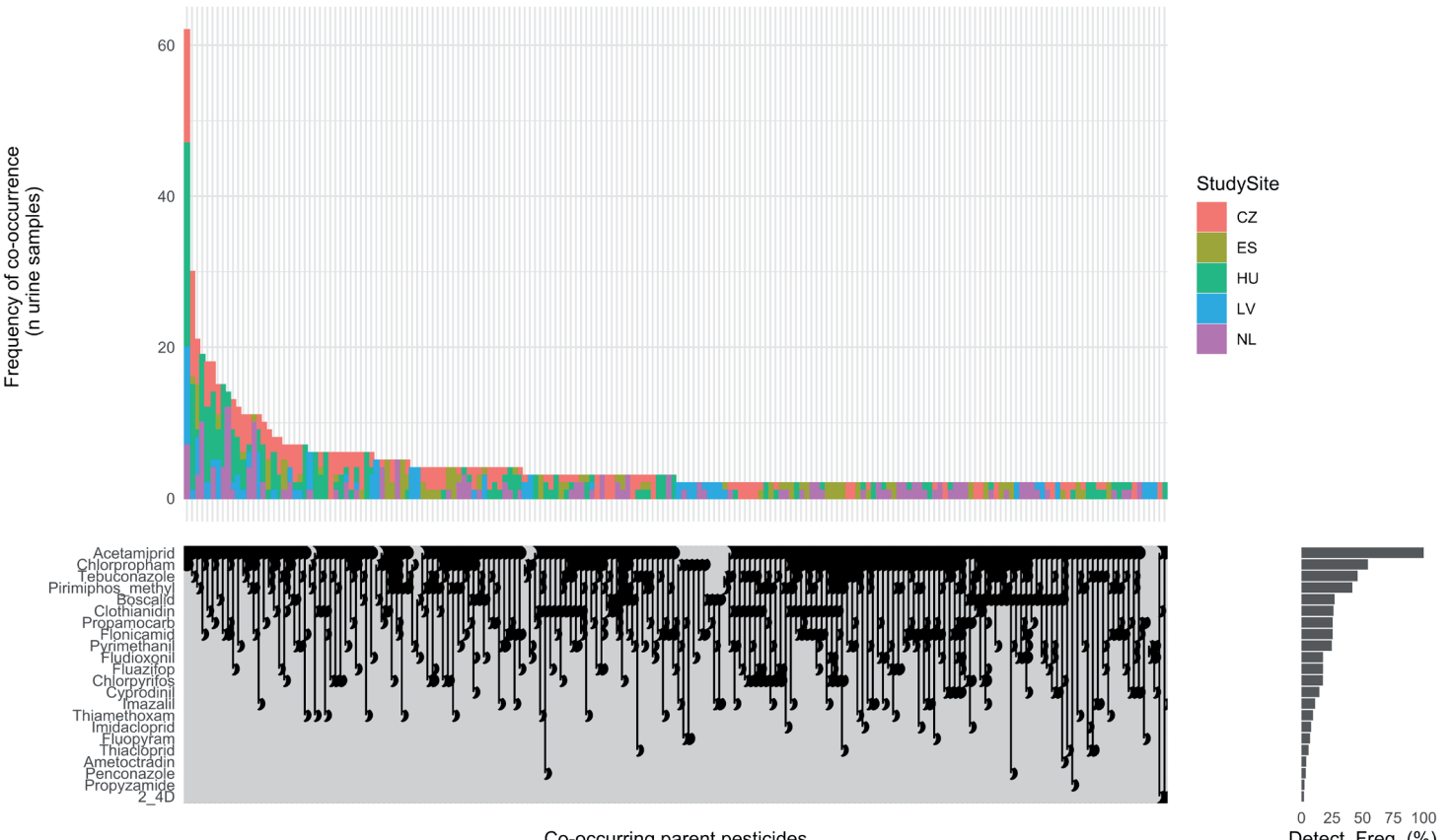


Table E. *Results of logistic mixed effects models, main and extended. Results are presented as Odds Ratios (OR) with 95% confidence intervals (CI). Significance levels based on p-value: '***' <0.001, '**' <0.01, '*' <0.05. Random effects are household and participant ID. Main model includes the predictors: location, season, and age category. Extended model includes additional predictors for pesticide usage, BMI, level of education and homegrown food consumption.*

ID	Parent pesticide	Category	ES		LV	
			Main OR (95% CI)	Extended OR (95% CI)	Main OR (95% CI)	Extended OR (95% CI)
P1	2,4-D	Season 2 vs 1[1] Parent vs Child Agricultural vs Non-agricultural	1.7 (0.5; 5.9) 2.5 (0.7; 8.8) 0.7 (0.1; 5.5)	1.6 (0.4; 5.9) 3.4 (0.5; 21) 0.8 (0.1; 7.3)	NA	NA
P2_a	Acetamiprid	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	0.5 (0.1; 2.7) 0.2 (0.0; 1.7) (100% detect in Agricultural area, no estimate possible)	0.5 (0.1; 2.7) 0.2 (0.0; 2.8) (Not possible)	0.6 (0.4; 1.0) . 0.8 (0.5; 1.3) 1.0 (0.6; 1.6)	0.6 (0.4; 1.0) * 0.7 (0.4; 1.5) 1.1 (0.7; 1.9)
P3_a	Ametoctradin	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	0.4 (0.1; 1.5) 2.1 (0.3; 17) 0.3 (0.0; 2.8)	0.6 (0.2; 1.4) 3.0 (0.9; 10) 0.4 (0.1; 1.0)	0.6 (0.2; 2.0) 0.8 (0.2; 2.8) 1.8 (0.5; 6.3)	0.5 (0.2; 1.9) 2.3 (0.3; 15) 1.3 (0.3; 5.3)
P5_a	Boscalid	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	0.7 (0.4; 1.0) . 2.9 (1.8; 4.6) *** 1.0 (0.6; 1.9)	0.6 (0.4; 1.0) . 2.5 (1.3; 4.9) ** 1.0 (0.5; 1.8)	0.9 (0.5; 1.6) 1.3 (0.7; 2.4) 1.2 (0.5; 2.6)	0.9 (0.5; 1.6) 1.0 (0.4; 2.6) 1.4 (0.6; 3.1)
P5_b			0.5 (0.2; 1.7) 1.2 (0.2; 6.8) 1.1 (0.2; 7.5)	0.6 (0.2; 1.7) 4.6 (0.9; 23) 1.1 (0.2; 8.0)	NA	NA
P6	Chlorantraniliprole	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	7.6 (1.7; 34) ** 1.0 (0.4; 2.7) 1.0 (0.4; 2.9) No random effects, this resulted in unreliable model	6.8 (1.5; 31) * 0.8 (0.2; 3.2) 0.7 (0.2; 2.4) No random effects	NA	NA
P8_a	Chlorpropham	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	0.4 (0.3; 0.7) *** 0.4 (0.2; 0.6) *** 0.7 (0.4; 1.3)	0.4 (0.3; 0.7) *** 0.3 (0.2; 0.6) *** 0.7 (0.4; 1.3)	1.6 (1.0; 2.6) . 0.3 (0.2; 0.6) *** 1.3 (0.7; 2.7)	1.5 (0.9; 2.4) 0.4 (0.2; 1.0) * 1.2 (0.6; 2.5)
P9_a	Chlorpyrifos (/methyl)	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	0.4 (0.1; 2.1) 6.2 (0.7; 52) . 1.4 (0.3; 6.3)	0.4 (0.1; 2.4) 3.9 (0.3; 52) 1.5 (0.3; 7.2) Not correct Educ	NA	NA
P9_b			0.2 (0.1; 0.4) *** 0.5 (0.3; 0.7) *** 0.8 (0.5; 1.3)	0.2 (0.1; 0.4) *** 0.4 (0.2; 0.7) ** 0.8 (0.5; 1.3)	NA	NA
P10	Clopyralid	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	NA	NA	NA	NA
P11_a	Clothianidin (can come from thiamethoxam)	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	0.6 (0.4; 0.9) ** 0.6 (0.5; 0.9) * 0.5 (0.3; 0.8) **	0.5 (0.3; 0.8) ** 0.6 (0.3; 1.0) . 0.4 (0.3; 0.7) ***	6.3 (0.7; 53) . 0.2 (0.0; 1.4) . 1.4 (0.3; 6.3)	5.7 (0.7; 50) 0.3 (0.0; 5.7) 0.9 (0.2; 4.6)
P11_b			NA	NA	NA	NA
P11_c			0.9 (0.5; 1.5) 0.7 (0.4; 1.3) 1.4 (0.8; 2.8)	0.8 (0.5; 1.4) 0.8 (0.4; 1.8) 1.4 (0.7; 2.8)	NA	NA

HU		CZ		NL	
Main OR (95% CI)	Extended OR (95% CI)	Main OR (95% CI)	Extended OR (95% CI)	Main OR (95% CI)	Extended OR (95% CI)
0.5 (0.1; 2.0) 0.8 (0.2; 3.0) 0.8 (0.2; 3.1)	0.4 (0.1; 1.7) 0.6 (0.1; 2.9) 0.6 (0.1; 2.8)	0.5 (0.1; 1.7) 1.0 (0.3; 3.2) 1.7 (0.5; 5.3)	0.4 (0.1; 1.4) 1.7 (0.2; 12) 1.6 (0.5; 5.4)	NA	NA
1.2 (0.5; 2.7) 0.4 (0.2; 1.0) * 1.2 (0.1; 2.7)	1.3 (0.5; 3.1) 0.5 (0.2; 1.5) 1.4 (0.5; 3.5)	0.6 (0.1; 2.5) 1.0 (0.2; 4.1) 0.5 (0.1; 2.2)	0.8 (0.2; 3.9) 12 (1.0; 149) . 0.6 (0.1; 3.0)	0.2 (0.0; 0.8) * 1.3 (0.2; 9.0) 2.0 (0.3; 14)	0.5 (0.2; 1.2) 1.7 (0.3; 8.0) 2.4 (0.9; 6.2) .
0.7 (0.1; 4.1) 4.0 (0.4; 37) 1.6 (0.3; 9.5)	Not reliable, 1.2% detected	0.4 (0.1; 1.0) * 0.6 (0.2; 1.5) 3.2 (1.1; 9.5) *	0.4 (0.1; 1.1) . 0.2 (0.0; 1.3) . 3.0 (1.0; 9.4) .	0.3 (0.1; 1.6) 0.1 (0.02; 0.9) * 0.8 (0.03; 20)	Not reliable, 2.9% detected
0.6 (0.2; 1.7) 2.1 (1.0; 9.9) . 0.8 (0.2; 2.2)	0.6 (0.2; 1.9) 2.4 (0.6; 9.3) 0.5 (0.1; 1.5)	0.7 (0.4; 1.2) 2.1 (1.3; 3.5) ** 1.4 (0.8; 2.5)	0.8 (0.5; 1.3) 2.5 (1.0; 6.1) . 1.4 (0.8; 2.7)	1.4 (0.9; 2.2) 1.4 (0.9; 2.1) 0.6 (0.3; 1.0) .	1.4 (0.9; 2.2) 1.1 (0.4; 2.6) 0.5 (0.3; 1.0) .
NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA
0.5 (0.3; 0.8) ** 0.5 (0.3; 0.7) ** 1.3 (0.7; 2.5)	0.5 (0.3; 0.9) * 0.4 (0.2; 0.7) ** 1.5 (0.7; 3.2)	2.1 (1.3; 3.3) ** 0.4 (0.2; 0.6) *** 1.0 (0.6; 2.0)	1.9 (1.2; 3.1) * 0.3 (0.1; 0.8) * 1.0 (0.5; 2.0) Not correct Edu	2.8 (1.7; 4.7) *** 0.6 (0.4; 1.1) . 2.1 (1.1; 3.9) *	2.7 (1.6; 4.6) *** 0.5 (0.2; 1.2) 2.1 (1.1; 4.1) *
NA	NA	NA	NA	NA	NA
2.5 (1.0; 6.1) * 0.5 (0.2; 1.1) . 0.2 (0.1; 0.7) *	2.7 (1.1; 6.5) * 0.2 (0.1; 0.8) * 0.3 (0.1; 1.0) .	0.6 (0.4; 1.0) . 0.5 (0.3; 0.7) ** 1.3 (0.7; 2.4)	0.6 (0.4; 1.0) . 0.7 (0.3; 1.7) 1.3 (0.7; 2.4)	0.5 (0.2; 1.1) . 0.8 (0.3; 1.8) 1.2 (0.4; 3.2)	0.4 (0.1; 1.0) * 0.9 (0.2; 5.2) 1.2 (0.4; 3.3)
NA	NA	1.0 (0.2; 5.1) 2.0 (0.4; 11) 0.2 (0.0; 2.0)	0.6 (0.4; 1.0) . 0.7 (0.3; 1.7) 1.3 (0.7; 2.4) not correct PestUse	NA	NA
3.1 (1.8; 5.2) *** 0.6 (0.4; 1.0) . 2.8 (1.6; 4.7) ***	3.5 (2.0; 6.1) *** 0.4 (0.2; 0.7) ** 2.8 (1.5; 5.1) **	0.6 (0.4; 1.0) . 1.0 (0.6; 1.5) 1.3 (0.8; 2.2)	0.7 (0.4; 1.1) 1.1 (0.5; 2.7) 1.4 (0.8; 2.5)	0.6 (0.4; 1.0) . 0.9 (0.5; 1.5) 1.3 (0.7; 2.3)	0.6 (0.4; 1.0) . 1.7 (0.6; 4.6) 1.4 (0.8; 2.7)
NA	NA	NA	NA	NA	NA
1.9 (0.9; 3.9) . 0.8 (0.4; 1.5) 4.4 (1.8; 11) ***	2.5 (1.2; 5.5) * 0.3 (0.1; 0.8) * 5.5 (2.1; 14) ***	0.5 (0.2; 1.1) . 0.9 (0.4; 2.0) 1.3 (0.6; 2.7)	0.5 (0.2; 1.1) . 1.2 (0.3; 4.6) 1.3 (0.6; 2.8)	0.02 (0.0; 0.3) ** 1.6 (0.0; 42) 1.5 (0.0; 38)	0.6 (0.4; 1.0) . 1.8 (0.7; 4.9) 1.4 (0.7; 2.6) Not corrected for PestUse &Educ, 3% detect

Table E. *Continued.*

ID	Parent pesticide	Category	ES		LV	
			Main OR (95% CI)	**Extended OR (95% CI)**	**Main OR (95% CI)**	**Extended OR (95% CI)**
P12_a	Cypermethrin, cyfluthrin, permethrin, transfluthrin	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	NA	NA	NA	NA
P13_a	Cyprodinil	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	1.2 (0.7; 2.1) 2.0 (1.6; 5.6) *** 1.4 (0.7; 2.7)	1.2 (0.6; 2.2) 2.3 (1.0; 5.4) . 1.4 (0.7; 2.7)	0.7 (0.3; 1.5) 0.9 (0.4; 2.1) 1.1 (0.4; 2.8)	0.7 (0.3; 1.5) 1.5 (0.5; 5.1) 1.1 (0.4; 2.9)
P18_a	Flonicamid	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	1.4 (0.3; 6.3) 6.2 (0.7; 52) 2.6 (0.5; 14)	1.4 (0.3; 7.1) 7.8 (0.6; 96) 2.6 (0.4; 15) Not correct Educ	NA	NA
P18_b			0.3 (0.2; 0.6) *** 0.6 (0.4; 1.1) 0.8 (0.5; 1.5)	0.3 (0.2; 0.6) *** 0.6 (0.3; 1.4) 0.9 (0.5; 1.6)	NA	NA
P19_a	Fluazifop	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	0.5 (0.3; 0.8) ** 1.0 (0.6; 1.7) 1.0 (0.6; 1.9)	0.5 (0.3; 0.9) * 0.7 (0.3; 1.5) 1.1 (0.6; 2.1)	0.7 (0.2; 2.4) 1.0 (0.3; 3.5) 4.2 (0.9; 20) .	0.6 (0.1; 2.3) 1.2 (0.2; 7.8) 3.1 (0.6; 16)
P19_b			0.8 (0.4; 1.6) 1.0 (0.5; 2.1) 1.6 (0.7; 3.6)	0.7 (0.3; 1.6) 0.4 (0.1; 1.2) 1.6 (0.6; 4.1)	0.5 (0.1; 2.7) 0.5 (0.1; 2.7) 5.2 (0.6; 45)	0.4 (0.1; 2.6) 1.9 (0.1; 29) 4.6 (0.5; 44)
P20	Fludioxonil	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	0.5 (0.3; 0.9) * 1.5 (0.8; 2.8) 1.1 (0.6; 2.1)	0.5 (0.3; 0.9) * 0.9 (0.4; 2.2) 1.0 (0.5; 2.0)	0.8 (0.4; 1.4) 1.1 (0.6; 2.1) 0.8 (0.4; 1.7)	0.8 (0.4; 1.4) 0.8 (0.3; 2.1) 0.9 (0.4; 1.9)
P21_a	Fluopyram	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	3.1 (0.8; 12) 0.8 (0.2; 2.8) 1.1 (0.1; 9.5)	3.6 (0.8; 15) . 1.0 (0.1; 8.1) 2.0 (0.2; 19)	NA	NA
P21_b			2.1 (0.4; 11) 1.0 (0.2; 4.5) 1.2 (0.1; 20)	2.2 (0.4; 13) 0.3 (0.0; 5.6) 1.7 (0.1; 41)	NA	NA
P21_c			1.0 (0.5; 1.9) 1.5 (0.8; 3.0) 1.4 (0.6; 3.4)	1.1 (0.5; 2.2) 0.9 (0.3; 2.6) 1.5 (0.6; 4.0)	0.2 (0.0; 0.8) * 0.9 (0.1; 6.1) 0.6 (0.1; 4.3)	
P22_a	Flupyradifurone	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	0.0 (0.0; 0.8) * 1.2 (0.0; 33) 0.6 (0.0; 18)	0.3 (0.1; 1.2) 0.7 (0.1; 4.2) 0.5 (0.1; 1.8)	NA	NA
P25_a	Fluvalinate	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	NA	NA	NA	NA
P27_a	Imazalil	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	0.2 (0.1; 0.3) *** 1.1 (0.7; 2.0) 1.0 (0.5; 2.0)	0.2 (0.1; 0.4) *** 0.8 (0.4; 1.9) 1.1 (0.6; 2.2)	0.4 (0.2; 0.8) * 2.4 (1.1; 5.2) * 1.7 (0.6; 4.9)	0.4 (0.2; 1.1) . 1.1 (0.3; 4.3) 0.7 (0.2; 2.2)
P28_a	Imidacloprid	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	0.6 (0.3; 1.0) * 1.4 (0.8; 2.5) 1.5 (0.7; 3.1)	0.5 (0.3; 1.0) * 1.8 (0.8; 4.2) 1.2 (0.6; 2.6)	0.2 (0.0; 1.4) 1.4 (0.3; 6.1) 1.4 (0.3; 6.2)	0.2 (0.0; 1.4) 0.7 (0.1; 6.2) 1.1 (0.2; 5.7)
P32_a	Penconazole	Season 2 vs 1 Parent vs Child Agricultural vs Non-agricultural	1.1 (0.5; 2.5) 2.2 (0.9; 5.0) . 0.7 (0.3; 1.7)	1.2 (0.5; 2.6) 2.9 (0.9; 9.0) . 0.8 (0.3; 1.9)	0.8 (0.2; 3.4) 0.8 (0.2; 3.4) 1.4 (0.3; 6.1)	0.8 (0.2; 3.4) 0.6 (0.1; 6.0) 1.6 (0.3; 7.7) Not correct PestUse

HU		CZ		NL	
Main OR (95% CI)	**Extended OR (95% CI)**	**Main OR (95% CI)**	**Extended OR (95% CI)**	**Main OR (95% CI)**	**Extended OR (95% CI)**
NA	NA	NA	NA	NA	NA
0.3 (0.1; 2.0)	0.4 (0.1; 2.9)	0.9 (0.5; 1.6)	0.9 (0.5; 1.7)	0.8 (0.5; 1.2)	0.8 (0.5; 1.2)
6.2 (0.9; 40)	2.9 (0.3; 26)	0.8 (0.4; 1.5)	1.0 (0.3; 3.0)	0.5 (0.3; 0.8) **	1.2 (0.5; 2.8)
1.9 (0.1; 44)	1.8 (0.1; 61)	0.3 (0.1; 0.6) **	0.3 (0.1; 0.6) **	0.6 (0.3; 0.9) *	0.5 (0.3; 0.9) *
3.1 (0.6; 16)	3.8 (0.7; 20)	1.0 (0.2; 5.3)	Not reliable, 2.7% detect	1.6 (0.6; 4.0)	1.5 (0.6; 3.8)
0.3 (0.1; 1.6)	0.2 (0.0; 1.5)	0.1 (0.0; 0.4) **		0.3 (0.1; 0.8) *	0.3 (0.0; 2.3)
1.0 (0.3; 4.3)	2.3 (0.5; 11)	2.7 (0.1; 84)		0.4 (0.1; 1.5)	0.4 (0.1; 1.3)
2.2 (1.4; 3.6) **	2.6 (1.6; 4.4) ***	NA	NA	0.9 (0.6; 1.4)	0.9 (0.6; 1.4)
0.9 (0.5; 1.5)	0.6 (0.3; 1.2)			0.7 (0.4; 1.0) .	1.2 (0.6; 2.8)
0.6 (0.4; 1.0) .	0.8 (0.4; 1.4)			1.0 (0.6; 1.6)	0.9 (0.5; 1.5)
1.1 (0.6; 2.0)	0.9 (0.4; 1.7)	0.9 (0.5; 1.4)	0.8 (0.5; 1.4)	1.0 (0.6; 1.6)	1.0 (0.6; 1.7)
0.6 (0.3; 1.2)	0.4 (0.2; 1.0) *	0.7 (0.4; 1.2)	0.5 (0.2; 1.3)	0.7 (0.4; 1.2)	0.5 (0.2; 1.5)
1.1 (0.5; 2.2)	0.9 (0.4; 2.1)	1.1 (0.6, 2.1)	1.3 (0.7; 2.4) not correct Educ	1.4 (0.7; 3.0)	1.5 (0.7; 3.1)
1.4 (0.1; 4.5)	1.0 (0.3; 2.6)	2.5 (1.0; 6.3) .	2.4 (0.9; 6.3) .	7.2 (1.6; 32) **	4.5 (1.1; 17) *
0.2 (0.0; 0.9) *	0.3 (0.1; 0.9) *	1.3 (0.6; 3.2)	1.0 (0.2; 4.6)	0.6 (0.1; 4.1)	0.0 (0.0; 2.9)
2.1 (0.3; 17)	2.4 (0.8; 7.2)	0.9 (0.3; 2.3)	0.9 (0.3; 2.5) not correct Educ	1.4 (0.2; 10)	3.1 (0.3; 29)
1.7 (0.4; 7.2)	1.7 (0.4; 7.5)	0.7 (0.4; 1.2)	0.7 (0.4; 1.2)	0.6 (0.3; 0.9) *	0.6 (0.4; 0.9) *
1.0 (0.2; 4.0)	0.9 (0.2; 5.5)	0.7 (0.4; 1.2)	1.1 (0.4; 3.1)	0.5 (0.3; 0.9) *	0.8 (0.3; 1.8)
0.6 (0.1; 2.6)	0.9 (0.2; 4.8)	0.5 (0.3; 0.9) *	0.6 (0.3; 1.1) .	0.9 (0.5; 1.6)	0.8 (0.5; 1.4)
NA	NA	NA	NA	NA	NA
NA	NA	3.9 (1.1; 14) *	4.0 (1.1; 15)	1.0 (0.2; 4.5)	1.1 (0.3; 3.5)
		1.0 (0.3; 2.9)	1.9 (0.3; 12)	0.5 (0.0; 6.2)	1.6 (0.1; 28)
		0.6 (0.2; 1.9)	0.7 (0.2; 2.3)	0.7 (0.1; 8.3)	0.7 (0.1; 5.4)
NA	NA	1.8 (0.5; 6.4)	1.8 (0.5; 7.2)	4.3 (0.5; 36)	Not reliable, 3% detected
		0.8 (0.2; 2.8)	1.0 (0.1; 12)	0.5 (0.2; 12)	
		1.3 (0.2; 10)	1.9 (0.2; 20)	1.0 (0.1; 22)	
NA	NA	NA	NA	0.1 (0.0; 1.9)	0.6 (0.1; 5.4)
				0.8 (0.0; 25)	0.1 (0.0; 3.0)
				0.8 (0.0; 27)	0.8 (0.0; 27)
NA	NA	NA	NA	NA	NA
0.2 (0.1; 0.5) ***	0.2 (0.1; 0.5) **	0.1 (0.0; 0.6) *	0.4 (0.2; 1.2)	0.3 (0.1; 1.0) *	0.4 (0.1; 1.0) .
0.5 (0.2; 1.2)	0.3 (0.1; 1.1) .	2.9 (0.2; 44)	3.1 (0.6; 15)	1.4 (0.6; 3.6)	2.2 (0.4; 11)
0.6 (0.2; 2.0)	0.7 (0.2; 2.3)	0.9 (0.1; 11)	1.0 (0.4; 2.6) Not correct Edu	1.6 (0.6; 4.2)	1.6 (0.6; 4.1) not corrected for Edu
0.9 (0.3; 2.5)	1.1 (0.4; 3.0)	NA	NA	1.1 (0.5; 2.1)	1.1 (0.5; 2.1)
0.9 (0.3; 2.4)	0.5 (0.1; 2.0)			0.8 (0.4; 1.7)	0.9 (0.2; 3.3)
0.9 (0.2; 3.3)	1.0 (0.3; 3.7)			1.2 (0.6; 2.5)	1.4 (0.6; 2.9)
3.6 (0.7; 18)	9.5 (1.0; 93)	0.5 (0.1; 2.0)	0.6 (0.1; 2.5)	1.0 (0.3; 3.6)	1.0 (0.3; 3.6)
1.3 (0.3; 4.8)	0.7 (0.0; 12)	2.0 (0.5; 8.3)	2.9 (0.3; 30)	1.5 (0.4; 5.5)	0.5 (0.1; 4.5)
0.5 (0.1; 2.1)	0.7 (0.0; 38)	2.4 (0.6; 9.7)	2.2 (0.5; 9.3) Not correct Educ. PestUse unreliable estim	0.9 (0.3; 3.2)	0.8 (0.2; 3.0) not corrected for PestUse

Table E. *Continued.*

ID	Parent pesticide	Category	ES		LV	
			Main OR (95% CI)	**Extended OR (95% CI)**	**Main OR (95% CI)**	**Extended OR (95% CI)**
P34_a	Pirimiphos-methyl	Season 2 vs 1	0.8 (0.4; 1.4)	1.0 (0.5; 2.0)	1.5 (0.8; 2.9)	0.8 (0.2; 3.8)
		Parent vs Child	0.4 (0.2; 0.9) *	0.2 (0.1; 0.7) *	0.4 (0.2; 0.9) *	0.6 (0.1; 5.4)
		Agricultural vs Non-agricultural	0.7 (0.3; 1.4)	1.1 (0.5; 2.5)	1.0 (0.5; (1.9)	1.5 (0.3; 7.4)
P35_a	Propamocarb	Season 2 vs 1	1.9 (0.9; 4.1) .	1.7 (0.8; 3.5)	NA	NA
		Parent vs Child	1.1 (0.5; 2.5)	1.2 (0.4; 3.4)		
		Agricultural vs Non-agricultural	1.3 (0.5; 3.1)	1.0 (0.4; 2.4)		
P35_b			1.2 (0.7; 1.9)	1.1 (0.7; 1.9)	0.3 (0.1; 0.9) *	0.3 (0.1; 0.8) *
			1.1 (0.7; 1.9)	1.1 (0.5; 2.3)	0.4 (0.1; 1.1) .	0.2 (0.0; 1.0) *
			1.3 (0.7; 2.4)	1.2 (0.7; 2.2)	4.2 (1.1; 16) *	4.0 (1.0; 17) .
P37	Propyzamide	Season 2 vs 1	1.2 (0.6; 2.4)	1.2 (0.6; 2.7)	NA	NA
		Parent vs Child	2.3 (1.1; 5.0) *	1.4 (0.5; 4.2)		
		Agricultural vs Non-agricultural	1.8 (0.7; 4.4)	1.8 (0.7; 4.7)		
P38_a	Pyrimethanil	Season 2 vs 1	0.6 (0.4; 1.0) *	0.7 (0.4; 1.0) .	2.1 (1.1; 3.8) *	2.1 (1.1; 3.8) *
		Parent vs Child	0.8 (0.5; 1.2)	0.7 (0.3; 1.3)	1.0 (0.6; 1.8)	1.3 (0.5; 3.2)
		Agricultural vs Non-agricultural	0.8 (0.5; 1.3)	0.8 (0.5; 1.3)	1.2 (0.6; 2.4)	1.1 (0.5; 2.3)
P38_b			NA	NA	NA	NA
P40_a	Tebuconazole	Season 2 vs 1	0.5 (0.3; 0.9) *	0.5 (0.3; 0.8) **	3.3 (1.1; 9.3) *	2.8 (1.1; 7.4)
		Parent vs Child	0.3 (0.2; 0.4) ***	0.2 (0.1; 0.4) ***	0.8 (0.3; 2.1)	2.9 (0.7; 12)
		Agricultural vs Non-agricultural	0.9 (0.5; 1.7)	0.7 (0.4; 1.4)	1.5 (0.4; 5.3)	1.5 (0.6; 4.0)
P41_a	Thiabendazole	Season 2 vs 1	NA	NA	NA	NA
		Parent vs Child				
		Agricultural vs Non-agricultural				
P42_a	Thiacloprid	Season 2 vs 1	0.9 (0.4; 2.0)	1.1 (0.5; 2.4)	NA	NA
		Parent vs Child	0.4 (0.2; 1.0) *	0.3 (0.1; 1.2) .		
		Agricultural vs Non-agricultural	1.6 (0.6; 4.4)	1.6 (0.6; 4.5)		
P43_a	Thiamethoxam	Season 2 vs 1	NA	NA	NA	NA
		Parent vs Child				
		Agricultural vs Non-agricultural				
P43_b			94 (22; 396) ***	93 (22; 390) ***	NA	NA
			1.1 (0.7; 1.8)	1.3 (0.6; 2.7)		
			0.6 (0.4; 1.1)	0.5 (0.3; 1.0) *		
P46_a	Trifloxystrobin	Season 2 vs 1	NA	NA	NA	NA
		Parent vs Child				
		Agricultural vs Non-agricultural				

1. Underlined is the reference category.

HU		CZ		NL	
Main OR (95% CI)	**Extended OR (95% CI)**	**Main OR (95% CI)**	**Extended OR (95% CI)**	**Main OR (95% CI)**	**Extended OR (95% CI)**
3.6 (1.4; 9.4) **	4.1 (1.5; 11) **	0.6 (0.4; 1.0) .	0.6 (0.4; 1.1) .	0.7 (0.4; 1.0) .	0.7 (0.4; 1.0) .
0.2 (0.1; 0.5) **	0.1 (0.0; 0.3) ***	0.5 (0.3; 0.8) **	0.6 (0.2; 1.5)	0.3 (0.2; 0.5) ***	0.3 (0.1; 0.8) *
0.1 (0.02; 0.4) ***	0.2 (0.1; 0.8) *	1.5 (0.8; 2.7)	1.3 (0.7; 2.5)	1.4 (0.8; 2.7)	1.4 (0.8; 2.7)
1.1 (0.6; 2.1)	1.1 (0.6; 2.2)	1.5 (0.4; 6.0)	1.4 (0.5; 3.5)	1.4 (0.9; 2.4)	1.5 (0.9; 2.5)
0.4 (0.2; 1.0) *	0.3 (0.1; 0.9) *	0.5 (0.1; 3.9)	1.5 (0.2; 10)	1.0 (0.6; 1.6)	3.1 (1.1; 8.5) *
0.5 (0.2; 1.0) .	0.5 (0.2; 1.2)	0.5 (0.1; 4.5)	0.5 (0.2; 1.9) Not correct educ	1.6 (0.8; 3.0)	1.9 (1.0; 3.6) .
2.1 (1.2; 3.6) *	2.3 (1.3; 4.1) **	1.4 (0.7; 2.6)	1.4 (0.7; 2.8)	1.0 (0.6; 1.5)	1.0 (0.7; 1.5)
0.5 (0.3; 0.9) *	0.4 (0.2; 0.8) *	0.5 (0.3; 1.0) .	0.9 (0.2; 3.3)	1.2 (0.8; 1.8)	3.2 (1.4; 7.3) **
0.5 (0.3; 0.9) *	0.5 (0.2; 1.0) *	0.7 (0.3; 1.9)	0.8 (0.3; 2.1)	1.2 (0.7; 1.9)	1.3 (0.8; 2.2)
NA	NA	NA	NA	NA	NA
1.3 (0.5; 3.8)	1.6 (0.5; 5.2)	0.4 (0.3; 0.7) **	0.5 (0.3; 0.8) **	0.6 (0.4; 1.0) *	0.6 (0.4; 1.0) *
2.3 (0.8; 7.2)	1.7 (0.8; 7.4)	1.4 (0.9; 2.3)	2.1 (0.9; 5.3)	0.8 (0.5; 1.2)	1.1 (0.5; 2.4)
0.4 (0.1; 2.0)	0.3 (0.0; 3.1)	0.9 (0.5; 1.7)	1.0 (0.5; 1.8)	0.6 (0.4; 1.0) *	0.6 (0.4; 0.9) *
0.6 (0.2; 2.0)	0.6 (0.2; 2.5)	NA	NA	NA	NA
1.2 (0.4; 4.0)	4.5 (0.9; 23) .				
2.8 (0.7; 11)	3.3 (0.7; 16)				
2.1 (1.2; 3.5) **	2.2 (1.3; 3.8) **	0.7 (0.5; 1.0) .	0.7 (0.4; 1.0) .	0.6 (0.4; 1.0) .	0.6 (0.4; 1.1) .
0.3 (0.2; 0.5) ***	0.5 (0.2; 0.9) *	0.2 (0.2; 0.4) ***	0.4 (0.2; 0.7) **	0.1 (0.1; 0.2) ***	0.2 (0.1; 0.5) ***
0.5 (0.3; 1.0) .	0.5 (0.2; 1.0) *	1.0 (0.6; 1.5)	1.1 (0.7; 1.7)	0.8 (0.4; 1.5)	0.58 (0.4; 1.6)
NA	NA	NA	NA	NA	NA
3.2 (0.8; 12) .	3.5 (0.9; 14) .	0.9 (0.4; 2.0)	1.2 (0.6; 2.7)	6.5 (1.5; 29) *	6.3 (1.3; 30) *
0.3 (0.1; 1.2) .	0.6 (0.1; 3.3)	0.3 (0.1; 0.7) **	0.2 (0.1; 1.1) .	0.6 (0.2; 2.1)	2.4 (0.2; 27)
0.1 (0.0; 0.7) *	0.1 (0.0; 0.5) * Not correct PestUse	1.6 (0.6; 4.7)	2.1 (0.8; 5.8)	0.4 (0.1; 3.5)	0.5 (0.1; 4.0) not corrected for Edu
4.3 (0.9; 21) .	3.7 (0.7; 20)	NA	NA	NA	NA
1.0 (0.3; 3.5)	0.3 (0.1; 1.7)				
10 (1.2; 80) *	7.2 (0.7; 79)				
NA	NA	NA	NA	NA	NA
1.4 (0.8; 2.3)	1.0 (0.5; 1.9)				
1.9 (1.1; 3.4) * (not detected in season 1..)	1.9 (1.0; 3.5) .				
NA	NA	2.3 (0.4; 15)	2.4 (0.3; 17)	2.3 (0.8; 6.8)	2.3 (0.8; 6.7)
		0.4 (0.0; 9.8)	0.1 (0.0; 25)	0.6 (0.2; 1.7)	2.6 (0.4; 17)
		1.0 (0.1; 19)	1.5 (0.1; 25)	1.5 (0.5; 4.3)	1.7 (0.6; 4.8)

Additional Information F. Information on area selection and recruitment of participants.

AREA SELECTION

This paragraph describes the selection of the agricultural and non-agricultural areas per country.

Spain

The agricultural area is located in Valencia, which is the second most important agricultural area in Spain and one of the regions with the highest pesticide use: 12.1% of the national total in 2009 (ECPA, 2010). The selected area in Valencia was the village of Godella, located in the "Valencian orchard", around 10km northwest of the capital, with a population of more than 10,000 inhabitants and in close vicinity to agricultural areas. The main crops in this municipality are orchards and citrus. In these kind of crops, pesticide application takes place regularly during the spraying season. Households located in the municipality of Godella, Rocafort, Masarojos, Moncada or Burjasot were included. Eligible households were located within 250 meters distance to an orchard or citrus field. Satellite images (Google maps) were used to confirm that the home location of each participant was within 250 m of an agricultural field. Active application of pesticides in these areas was confirmed according to data from the Municipal Tax Agency of Godella and the information of the warning bulletins of the Department of Plant Health of the "Conselleria de Agricultura, Medio Ambiente, Cambio Climático y Desarrollo Rural" of the Valencian Government of 2018.

The non-agricultural area is located in the peri-urban areas of Madrid (outside the ring road of M40 which defines central/urban Madrid), with low levels of agricultural activity. Eligible households were located at least 500 meters away from any agricultural area based on the information provided by participants and checked using Google Maps.

Latvia

Multiple agricultural areas were defined because of the low population density in countryside and also difficult recruitment of study participants. The agricultural areas were chosen from Kurzeme and Zemgale regions since historically these regions of Latvia are the most used for agriculture purposes. Area selection was based on the agricultural register from 2017, where farmers submit their land use (hectares and crops grown). The register indicated the largest total amount of agricultural land was located in Kurzeme and Zemgale regions.

Non-agricultural areas were defined as persons living at least 500m away from actively used agricultural lands – these were small villages, small cities and suburbs. We excluded possible study subjects that lived in the "big cities" that are known either because of their dense population (more than 10,000 inhabitants) or because of high economic activity – having many factories, a lot of traffic, etc.

Each study participant prior their acceptance in the study was asked how far from pesticide application sites do they live. This information was then evaluated using publicly available databases – one called kadastrs.lv was for checking the addresses to determine the cadastral

number of the property which was then submitted in a system for checking agricultural land usage (all crop types were considered, mostly cereals and potatoes are grown in Latvia) called https://karte.lad.gov.lv/. This system provides the opportunity to measure the distance from a specific area (one's address) to agricultural lands. In this way we determined whether our study subjects fitted as agricultural or non-agricultural addresses.

The system updates according to the season – this creates a situation where different cultures and crops can be grown in agricultural areas. The data was gathered taking into account the current situation – starting from March, 2020. The data of previous season was used to determine whether the person lives in an area with agricultural lands nearby where pesticides are used actively. Some study participants had only one type of crop/fruit/vegetable fields around their houses while most had several different types of fields.

Hungary

The selection of the agricultural and non-agricultural areas was based on the volume of apple growing. Szabolcs-Szatmár-Bereg Country has the largest area of apple orchards (17577 ha out of the 25044 ha). Based on the data provided by the Hungarian Central Statistical Office (KSH), the apple production amounted to around 0.3 million tons (approximately 60% of the total volume produced in Hungary) in 2016. Almost all settlements in the county have apple orchards where pesticides are used; however, we selected those settlements were several apple orchards are located. The selection of the household and participants was based on the predefined criteria and the distance between each household and the orchard was checked by Google Maps. Furthermore, the Division of Agriculture Plant Protection and Soil Conservation Department of the Government Office of Szabolcs-Szatmár-Bereg County provided information on the pesticide use at the exposed locations (e.g. date, name of pesticide product and active ingredient, dose).

We selected certain settlements in Nógrád Country as the non-agricultural area, since there is no significant fruit growing in this region. Most of the selected households were located in peri-urban area; however, some of them were in urban or rural areas. The distance from agricultural areas was checked by Google Maps.

Czech Republic

The area of interest was selected with the use of ArcGIS PRO. Two GIS layers containing information were used: the Land Parcel Identification System (LPIS CZ, Ministry of Agriculture) (https://eagri.cz/public/app/eagriapp/lpisdata/) and the Registry of territorial identification, addresses and real estate (RUIAN CZ, State Administration of Land Surveying and Cadaster) (https://www.cuzk.cz/ruian/RUIAN.aspx). LPIS CZ contains data on location, area, and general type of land parcel (e.g., field, orchard, vineyard, forest, pasture). RUIAN CZ contains information on addresses in the Czech Republic. The following procedure was then used:

1. Main focus was aimed at the South Moravian Region (SMR) in the Czech Republic (the Brno city is approximately in the center of SMR).
2. Only areas of fields, orchards, and vineyards were considered since we can expect the application of pesticides in these areas (LPIS CZ).

3. Street addresses in small cities (<5000 inhabitants) were extracted as layer (RUIAN CZ).
4. Buffer zone (250 m) around agricultural areas in SMR was created and intersected with the layer of street addresses. The street addresses within the buffer zone were considered potential agricultural areas.
5. Analogically, the buffer zone was expanded to 500 m and any street addresses not falling into this buffer zone were considered the non-agricultural area.

The address of those who expressed interest to join the study (and provided their home address) was then checked against the agricultural and non-agricultural area street addresses and potential participants were then categorized accordingly. The provisional check was also done via google.com/maps and mapy.cz. Finally, the surroundings were checked by field workers at the time of urine sample collection.

The Netherlands

Agricultural areas were areas with at least 100 inhabitants living within 250 meters from apple and pear orchards. A selection of addresses was made by combining two publicly available databases: the agricultural land-use database (2019), and the basic registration of buildings database (2019). All agricultural land use for apples and pears (orchards only) were selected, all buildings with a living function were selected. The focus was on the 'Betuwe' area, with the highest density of households fulfilling the criteria. This area is roughly located in the provinces Gelderland, Utrecht and part of North-Brabant between the rivers 'Nederrijn' and 'Waal'. Non-agricultural areas were defined as any address which was located at least 500 meters away from any agricultural land (including greenhouses). Households fulfilling these criteria from the Betuwe area and suburban Utrecht were included.

RECRUITMENT OF PARTICIPANTS

This paragraph describes the recruitment strategies implement in the different areas within each country.

Spain

For the agricultural area, recruitment started on October 15, 2019 and ended October 25, 2019. The recruitment was done in primary schools located in Godella (Valencia). This fact facilitated finding children with the age object of study (between 6 and 11 years old) and also their parents (or caretaker) living in households within 250 meters of agricultural area(s). The recruitment has been performed in public schools only, in which the number of volunteers was reached. After recruitment, 4 families withdrew, resulting in a total of 52 parent-child pairs participating.

In order to encourage participation, those in charge of recruiting followed a flexible policy with regards to dates and contact hours with children's parents. First of all, a first meeting with the school board of directors was organized at beginning of October-19 in order to inform school staff about the project, to request support from the centre and to organize the first

meeting with parents. The meetings with parents took place on 15th and 16th October 2019 in the following two primary schools of Godella: "CEIP Cervantes" and "CEIP El Barranquet". Copies of the information letter and invitation letter were provided to potential participants at this point. They also received the documents associated to the participation, such as a screening questionnaire, for further examination and consideration at home. Additionally, posters were displayed on schools to encourage the participation in the study.

For the non-agricultural area, recruitment started on June 14, 2019 and ended on September 30, 2019. Recruitment took place among co-workers. At the end of May 2019 a press note was released at the Spanish research institute webpage to inform workers about the project and about the 2 informative seminars that would take place in June. An email was sent at the beginning of June to all co-workers with basic information on the study. Additionally, posters were displayed to promote the seminars and to encourage participation in the HBM4EU study. At the seminars, the recruitment materials (information and invitation letters plus the screening questionnaire) were distributed to attendants. Recruitment started already at the seminars and followed by email among co-workers and co-workers' contacts willing to participate. A positive response was received from 60 families, however, 7 of them could not enter the study because they did not fulfil the selection criteria. This resulted in a total of 53 parent-child pairs participating. Those entering the study were given an envelope with the documentation associated to their participation i.e informed consent for parents, informed assent with an adapted language for child, FAQ sheet, information leaflet and the reply card as well as the urine sample collection kit in a portable coolbox with sampling instructions.

The study was approved by the medical ethical committee under number 20200109/10 for the agricultural area and by the Research Ethics Committee of the Instituto de Salud Carlos III under number CEI PI 34_2019-v2-Enmienda_2020 for the non-agricultural area

Latvia

Recruitment took place from February 18, 2020 until March 31, 2020.

There were many stages and ways of recruitment of study participants. First a list of contacts of all schools taking part in "eco-school" programme was made and the schools were contacted via e-mail (in total 70 schools). Only one responded via email and so the schools were contacted individually via phone and asked whether they are willing to participate by allowing to spread information on the project to children and their parents. Information envelopes containing a brief description of project activities, deadlines and contact persons were driven to schools for handing out. 33 respondents, mostly from agricultural areas, responded to this action.

A press release and a post on Facebook via Rīga Stradiņš University was made on October 21st, 2019, resulting in 400 shares. An email of general practitioners (family doctors) of Kurzeme and Zemgale regions were sent with information on this project as well.

Next banners and posters were made and sent out to Kurzeme and Zemgale local newspapers and the message was also put in "e-klase.lv" which is a system for all schools for organising the educational process – parents have access to the information on their child and checks the system regularly for grades, comments and information therefore a banner was made

visible to parents from Kurzeme and Zemgale for a week (10,000 views), with little result. At this point the research team concluded, that despite the effort to limit our study participants to be only from Kurzeme and Zemgale, the insufficient count of participants broadened the borders, and study participants are mostly from Kurzeme and Zemgale, but also from Vidzeme and Latgale. In total 50 parent-child pairs from agricultural areas, and 51 parent-child pairs from non-agricultural areas were included.

The study was approved by the medical ethical committee of Rīga Stradiņš University under number 6-3/3/48.

Hungary

The recruitment of the participants was performed between October and December 2019 and was coordinated by the staff of the Public Health Department of the Szabolcs-Szatmár-Bereg and Nógrád County Government Offices in close collaboration with the project team of the National Public Health Center. The recruitment of the volunteers was done through the primary schools in Nódrág County (non-agricultural area), while the health visitors being very familiar with the local circumstances were also involved in Szabolcs-Szatmár-Bereg County (agricultural area). In total, 54 YES reply cards from 11 settlements and 40 NO reply cards were collected at the agricultural area. Regarding the non-agricultural area, 68 YES reply cards from 8 settlements and 199 NO reply cards were received. It must be noted that the difference might be caused by the different recruitment strategies applied at the two areas; the health visitors at the agricultural locations selected families with whom they have already been in contact before the study. During the selection process, the volunteers were checked for the predefined selection criteria and the most suitable and the most committed adult – child pairs were included in the study. Before the sample collection, the signed informed consents were collected.In the case of agricultural areas, we have requested spray logs through the Szabolcs-Szatmár-Bereg County Government Office, so we know when and with what they were sprayed.

According to the spray logs, acetamiprid, an insecticide and acaricide, was also used on the apples (agricultural area), but chlorpropham was not used.

The study was approved by the Medical Research Council of Hungary under registration number 15521-3/2019/EKU.

Czech Republic

Recruitment started in mid-September 2019 and was finished at the end of February 2020 during at that time ongoing first sampling season.

Recruitment of all participants was done by post (letters, ~ 1000 sent, very low response rate <1%), promotional leaflets (1000 – 1500 delivered, very low response rate <1%), internet advertisement (e.g., posting on Social network web pages, short announcements on local news pages, announcements on internet pages of selected towns after communication with town mayor), announcements in a radio station and announcement in news relation on CZ - TV.

Approximately 200 people expressed interest to join the study. About 90 of them did not meet the criteria to join (children out of age range, occupation associated with pesticides,

etc.) or decided not to join (for whatever reason). Overall, 111 participants (adults) were eligible and willing to join. Of these 111 families, 16 were "double families" – families with both parents involved in the study with two children. This meant that samples were collected from 95 unique address points. The remaining 16 address points are associated with two sets of parent-child pairs.

We have encountered 3 cases where parents reported the incorrect age of their child in the initial screening questionnaire. This issue was discovered during the fieldwork of the first sampling season. We have ultimately decided to finish the collection of such samples in the second sampling season. The age of children out of the study target range (6 to 12 years old) was 4 years old (1 from an agricultural address and 1 non-agricultural) and 15 years (from agricultural address).

The SPECIMEn study in the Czech Republic received ethical approval under ref. no. ELSPAC/EK/3/2019.

The Netherlands

Recruitment of participants started in November 2019 and continued until February 2020.

In the Netherlands it turned out to be quite a challenge to combine the databases of land and building-use with the basic administration of municipalities (GBA) because of privacy regulations. Since we had no access to the age of subjects from a specific residential location, letters were sent out at random to addresses within the selected postal codes. Two batches of letters were sent out, the first of 1,000, the second of 10,000. Each letter contained an information letter, the screening questionnaire, a reply card, informed consent for both parent and child, and an information brochure about the study. The first batch consisted of 500 agricultural and 500 non-agricultural area addresses and was send in the first week of November 2019. The second batch of 4,000 agricultural and 6,000 non-agricultural area addresses was send half of December 2019. The numbers of letters are quite high, since most of letters would go to non-eligible households e.g. without children. The response was around 2%, of which about half was not eligible to participate. For example, when one of the household members was working with pesticides.

Because of time pressure and urge to start collecting the samples in January, we decided to combine recruitment strategies. A news item was placed in local newspapers (Figure 5.5.2) and on news-websites, including a QR-code directing to the website of the study. The study-website (https://www.rivm.nl/europees-onderzoek-naar-bestrijdingsmiddelen-in-urine) included an online sign-up form were potential participants could complete the screening questionnaire. It turned out that specifically the addresses within agricultural areas were interested in participating and some non-agricultural addresses were still missing. Therefore, additional recruitment was done among co-workers with children to participate. In total 55 parent-child pairs were recruited from agricultural areas, and 50 parent-child pairs from non-agricultural areas.

The medical research ethics committee confirmed that the Dutch Medical Research Involving Human Subjects Act (WMO) does not apply to the above mentioned study and that therefore an official approval of this study by the MREC Utrecht was not required under the WMO (reference number WAG/mb/19/027712).

Chapter 6

Urinary pesticide mixture patterns and exposure determinants in an adult population from the Netherlands and Switzerland: application of a suspect screening approach

I.B. Ottenbros[1,5], P. Ammann[2,3], M. Imboden[2,3], S. Fuhrimann[2,3], J-P. Zock[5], E. Lebret[1,5], R.C.H. Vermeulen[1], R. Nijssen[4], A. Lommen[4], H. Mol[4], J.J.Vlaanderen[1], N. Probst-Hensch[2,3]

[1] Institute for Risk Assessment Sciences, Utrecht University, Utrecht, Netherlands
[2] Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Kreuzstrasse 2, CH-4123 Allschwil, Switzerland
[3] University of Basel, P.O. Box, CH-4003 Basel, Switzerland
[4] Wageningen Food Safety Research – part of Wageningen University & Research, Wageningen, Netherlands
[5] Center for Sustainability, Environment and Health, National Institute for Public Health and the Environment (RIVM), Bilthoven, Netherlands

Submitted to Environmental Research (May 2023)

ABSTRACT

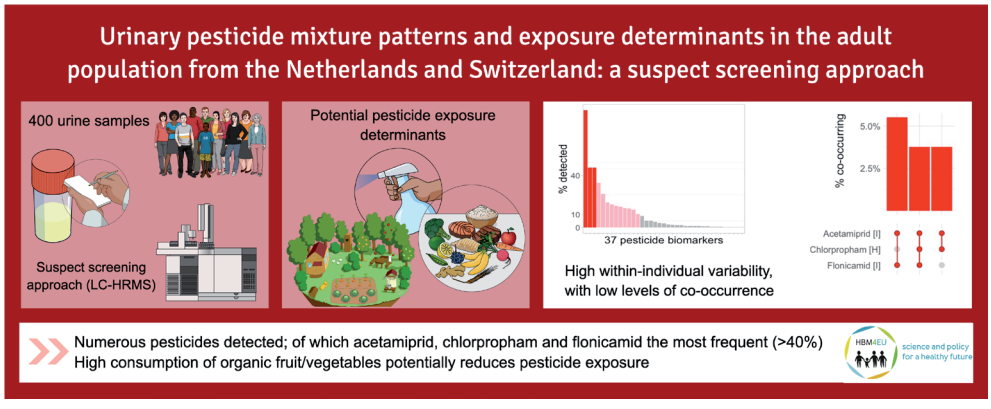
Introduction: Non-occupational sources of pesticide exposure may include domestic pesticide usage, diet, occupational exposure of household members, and agricultural activities in the residential area. We conducted a study with the ambition to characterize pesticide mixture patterns in a sample of the adult population of the Netherlands and Switzerland, using a suspect screening approach and to identify related exposure determinants.

Methods: A total of 105 and 295 adults participated in the Dutch and Swiss studies, respectively. First morning void urine samples were collected and analyzed in the same laboratory. Harmonized questionnaires about personal characteristics, pesticide-related activities, and diet were administered. Detection rates and co-occurrence patterns were calculated to explore internal pesticide exposure patterns. Censored linear and logistic regression models were constructed to investigate the association between exposure and domestic pesticide usage, consumption of homegrown and organic foods, household members' exposure, and distance to agricultural and forest areas.

Results: From the 37 detected biomarkers, 3 (acetamiprid (-CH2), chlorpropham (4-HSA), and flonicamid (-C2HN)) were detected in ≥40% of samples. The most frequent combination of biomarkers (acetamiprid-flonicamid) was detected in 22 (5.5%) samples. Regression models revealed an inverse association between high organic vegetable and fruit consumption and exposure to acetamiprid, chlorpropham, propamocarb (+O), and pyrimethanil (+O +SO3). Within-individual correlations in repeated samples (summer/winter) from the Netherlands were low (≤ 0.3), and no seasonal differences in average exposures were observed in Switzerland.

Conclusion: High consumption of organic fruit and vegetables was associated with lower pesticide exposure. In the two countries, detection rates and co-occurrence were typically low, and within-person variability was high. Our study results provide an indication for target biomarkers to include in future studies aimed at quantifying urinary exposure levels in European adult populations.

Graphical Abstract



INTRODUCTION

Pesticides are widely used in agriculture to protect crops. In Europe, more than 400 pesticide compounds are registered and marketed (European Commission, 2023). On a daily basis, the general population is exposed to a mixture of various pesticides by consumption of pesticide-containing food or drinks, domestic usage of pesticide-containing products, or living close to agricultural areas. The active ingredients of pesticides are intrinsically toxic and can adversely affect human health (Gilden et al., 2010; Cimino et al., 2017). Adverse health effects of single compounds, particularly reported in occupational settings, include cancer, neurological, mental, respiratory, reproductive, and developmental disorders as well as rheumatoid arthritis (Ohlander et al., 2020). The characterization of pesticide exposure patterns and exposure sources in the general population is an essential step toward understanding the full scope of health impacts of single compounds and mixtures. Exposure characterization by human biomonitoring (HBM) has the advantage, depending on the biomonitoring method chosen, to cover all possible exposure routes (dermal, oral, inhalation), thereby reflecting the internal exposure concentrations of a wide range of chemicals (Ganzleben et al., 2017). Since most pesticides are rapidly metabolized into more polar derivatives and excreted through urine, this matrix is typically used for pesticide exposure assessment (Angerer et al., 2007; Egeghy et al., 2011). Suspect Screening (SS) approaches based on liquid chromatography (LC) combined with high-resolution mass spectrometry (HRMS) make it possible to effectively measure large numbers of pesticides co-occurring in the same urine sample (Bonvallot et al., 2021; Huber et al., 2022; Vitale et al., 2022). This SS approach provides a list with annotations of pesticides and pesticide metabolites in a sample (Pourchet et al., 2020; Huber et al., 2022), and presents the measurements as semi-quantitative signal intensities. Although these signal intensities do not refer to absolute pesticide concentration levels, signal intensities for the same biomarker analyzed under identical laboratory settings can be compared.

Epidemiological studies describing exposure pathways to pesticides are mostly focusing on occupational populations or residents living in agricultural areas, where the exposure levels are typically higher compared to the general population (Deziel et al., 2015; Teysseire et al., 2021). Information on pesticide exposure in the general population is rather limited (Heffernan et al., 2016; Dahiri et al., 2021; Yusà et al., 2022), in particular with regard to exposure to pesticide mixtures at low concentrations (Hernández et al., 2017) and temporal variation of exposure (Attfield et al., 2014; Li et al., 2019). While for the general population, the exposure levels and the total number of exposure pathways might be lower, understanding the contribution of non-occupational exposure sources is crucial to study the link between pesticide exposure and adverse health effects, as well as to propose preventive measures to protect the general population, specifically those most vulnerable (pregnant or nursing women, infants or children, and the elderly) (European Commission, 2022).

The overall aim of this study was to explore pesticide mixture patterns of exposure in a sample of the adult population from the Netherlands and Switzerland using an HBM SS approach and to identify possible determinants.

MATERIAL AND METHODS

Study population and sample collection

The two studies presented here, the Dutch arm of the Survey on PEstiCIde Mixtures in Europe (SPECIMEn) and the Swiss pesticide suspect screening study, are part of the European Human Biomonitoring (HBM4EU) initiative and sought to generate new evidence on pesticide exposure in the general population. While the Dutch SPECIMEn study focused on exploring variations of pesticide exposure patterns in parent-child pairs by repeated sampling design in two seasons, the Swiss study provided exposure data of adults by taking single samples during three different seasons.

A concise summary of the Dutch sampling strategy is provided here, as the study population recruitment and sample collection procedure have already been described in detail elsewhere (Ottenbros et al., 2023). In short, first-morning void urine samples were collected from participants (parent-child pairs) from different locations (closer and further away from orchards) in two seasons (winter 2020: Jan-Mar 2020; and summer 2020: Jun-Jul 2020). For the current study, samples of 105 adults from the SPECIMEn study were included. In order to minimize the influence of pesticide applications in nearby agricultural areas on exposure intensity and to better assess baseline exposure of the adult population, only winter samples of adults were included in the main analyses. Farmers and other adults employed in the agricultural sector were not included. Participants mainly lived in the Betuwe area (between Arnhem-Gorinchem-Utrecht). The distance to the nearest agricultural plot and forest from each geo-coded address was calculated using QGIS software (v3.4.4) using publicly available data from the Dutch Central Bureau of Statistics (CBS). At the time of urine collection, a questionnaire was administered covering personal and household characteristics, activities prior to sampling, potential pesticide usage, and food consumption the day prior to sampling.

In the Swiss study, 300 adults from the canton of Basel-Stadt participated in the HBM4EU pesticide suspect screening study (Buekers et al., 2022). A total of five participants indicated that they used pesticides for occupational use, and were thus excluded from further analyses. Data and sample collection were performed between January 08, 2020, and October 10, 2020. Study participants were contacted in five sex-stratified recruitment waves via postal mail containing the study invitation letter, an information leaflet, and a response card. A total of 6,000 subjects, selected from the resident register based on age and long-term residency in Basel-Stadt, had been invited. Interested subjects were contacted after their successful electronic registration to the REDCap® (Research Electronic Data Capture) data collection tool (Harris et al., 2009), to identify a date for the urine sample collection. Instructions and a urine sample collection kit were sent by postal mail prior to the day of collection. The participants collected first-morning void urine samples at their homes and were asked to store the morning urine sample at 4°C using cooling pads and a cooling bag until the study team collected the sample. The urine samples were then transported to and processed in the study center at Swiss TPH, maintaining the cold chain throughout until the biobanking of urine aliquots at -80°C. An electronic self-administered pre-sample questionnaire (answered before the day of sample collection) and a post-sample questionnaire (answered on the day of sample collection) were

distributed. Participant recruitment, data collection, and laboratory workflow were performed and documented using REDCap® (Harris et al., 2009). The "minimal geo data model" (MGDM) for agricultural land use in the Basel-Stadt area was used to calculate the distance of the participants' geo-coded addresses to the nearest agricultural area and forest in QGIS 3.4.4 (*MMQGIS* and *NNJoin* plugin). An exact description of the definitions used for forest and agricultural areas for both countries is provided in the Appendix (Table A.1).

The harmonized questionnaires administered in the two countries were developed in the context of the HBM4EU project and mostly contained identical questions in both countries. Where necessary, they were additionally harmonized between the two study countries. Questions and variables of interest for the analysis were manually compared and, where necessary, re-coded by the authors (for details see Appendix, Table A.2). The medical research ethics committee confirmed that the Dutch Medical Research Involving Human Subjects Act (WMO) does not apply to the above-mentioned study and that therefore an official approval of this study by the Medical Research Ethics Committee (MREC) Utrecht was not required under the WMO (reference number WAG/mb/19/027712). The Swiss study acquired ethical approval from the local ethics committee (Ethikkommission Nordwest- und Zentralschweiz (EKNZ), 2019-02136). All participants provided their written informed consent.

Suspect screening approach

The urine samples from the Dutch and Swiss studies were both analyzed at the Wageningen Food Safety Research laboratory in the Netherlands under a harmonized and quality-controlled suspect screening (SS) analysis framework. Huber et al. (2022) and Vitale et al. (2022) provide a detailed description of the applied analytical workflow, which includes the following steps: 1) pH adjustment and solid phase extraction (SPE) cleanup, 2) LC coupled to full scan HRMS (LC-HRMS) to measure the extracts, 3) data processing and analysis, 4) prioritization of supposed detects, and 5) spectral comparison of suspected detects with reference standards for final confirmation.

Several metabolites of the same parent compound may be included in the final list of pesticides, but the confidence with which a compound can be determined may vary. Schymanski et al. (2014) developed a confidence score representing the (un)certainty about the identity of a compound, ranging from 1: 'Fully confirmed structure' to 5: 'Exact mass (m/z) of interest'. Only compounds identified by molecular structure, or confidence levels 1 (confirmed structure) and 2b (probable structure by diagnostic evidence), were considered in the current study.

The results of the SS analysis are presented as semi-quantitative signal intensities of the compound detects, i.e. as indicators of exposure rather than quantitative concentration levels. Higher signal intensity scores generally correspond to higher concentrations for the same compound, but may also depend on levels of ion suppression due to matrix effects. Equal signal intensity scores for different compounds may correspond to different concentration levels depending on their ionization efficiency.

Biomarkers are indicated with their parent pesticide name and the respective metabolite in parentheses the first time mentioned in the text. Upon the second mention, the name of

the biomarkers will be noted only by the parent pesticide for improved readability. To avoid confusion, metabolites in parentheses will remain stated if two or more biomarkers of the same pesticide were detected.

Statistical Analyses

For the analysis of the suspect screening data, the subset of 37 biomarkers confirmed with high confidence (Schymanski levels 1 and 2b) were considered. The detection frequency for each biomarker was calculated for the pooled dataset, as well as the Dutch and the Swiss population separately (see Appendix A, Figure A.1). The biomarkers were plotted with their log-transformed SS signal intensity score and their detection ratio. Co-occurrence of biomarkers (detected together in the same urine sample) was shown graphically using an UpSet plot (UpSetR, v1.4.0).

Temporal differences in both studies were assessed. In the Dutch study, two samples from the same individual were taken, one in each season, for which the intra-class correlation coefficient (ICC) was calculated. A linear mixed effects model with censored data (using the log-transformed intensity scores) was used, i.e. a multilevel Tobit model. This two-level (first: measurements, second: subjects) random intercept model was defined as follows:

$$\log(y_{ij}) = \beta + u_j^{(2)} + \varepsilon_{ij}^{(1)}$$

where β represents the intercept, u_j the between-subject error, and ε_{ij} the within-subject error. To assess the temporal differences in the Swiss study, average intensity scores for each season (winter, spring, summer) were displayed in boxplots for the 13 most detected biomarkers (see Appendix A, Figure A.2).

In order to explore determinants of exposure, censored regression models were constructed (Tobit, VGAM v1.1.7) (Henningsen, 2022). Given the explorative nature of the analyses, models were not adjusted for multiple comparisons. For biomarkers of only three pesticides (acetamiprid, chlorpropham and flonicamid), the detection rate was sufficiently high (at least 40% at each study side) to construct a censored linear regression model. For the remaining biomarkers, logistic censored regression models were constructed (based on detected yes/no). All models were *a priori* corrected for age (years), gender (male/female), BMI (normal: <25, overweight: 25-30, obese: >30), level of education (primary, secondary, tertiary, higher), income (<25%, 25-50%, 50-75%, >75% of country average), and country (not for country-specific models). The following exposure variables were mutually included in the models: having a household member who used pesticides occupationally (yes/no), pesticide usage in the garden (up to 3 days (3d) prior to sampling, yes/no), pesticide usage indoors (up to 3d prior sampling, yes/no), pesticide usage on pets (up to 3d prior sampling, yes/no), pesticide usage for hobby use (up to 3d prior sampling, yes/no), homegrown food consumption in summer (not-high (<50%), high (≥50%)), organic food consumption per food category (vegetables and fruit, bread, meat, rice, eggs, dairy; not-high (<50%), high (≥50%)), and distance (m) to the closest agricultural area or forest. Continuous variables were log-transformed (age, distance to agriculture, distance to forest). For sensitivity analyses, following the study of Baudry et al.

(2019), a 'low' category was created for less than 10% of homegrown/organic foods consumption (see Appendix A, Table A.3). Missing values in the independent variables were imputed using mice (v3.14.0), with normal distribution for the continuous variables, proportional odds model for the categorical (education, income, BMI), and logistic regression for the remaining variables. All regression models were built on the pooled dataset, including an adjustment for country. Additionally, models were constructed for each country separately (see Appendix A, Figures A.3 and A.4). Results of the censored linear and logistic models were shown in forest plots.

To identify whether exposure to the pesticides detected in our study was driven by the consumption of specific food items, we identified food items from the 2020 European Food Consumption Database by the European Food Safety Authority (EFSA) in which the 12 most frequently detected parent pesticides from our study were often detected (European Food Safety Authority et al., 2022). For each compound, the five most frequently contaminated food items from the EFSA database were selected, with contamination frequencies ranging from 2.5% to 75.9%. Food items irrelevant for our study population, i.e. infant formulas and ready-made meals for children, were excluded. A description of these food items and the percentage consumed in the Dutch and Swiss population is provided in Appendix A (Table A.4).

RESULTS

Characteristics of the study samples

An overview of the two study population characteristics is shown in Table 1. A total of 105 (70% female) and 295 (46% female) adults were included from the Dutch (NL) and Swiss (CH) HBM4EU study, respectively. Both the age range as well as the mean age was higher in the Dutch sample (range: 29-56 years, mean: 42 years) as compared to the Swiss participant population (range: 20-39 years, mean: 31 years), reflecting the differences in the target population. With regard to BMI and educational level, the two study populations were similar. Approximately 73% of participants were of normal weight/underweight, and 73% had a university degree. The majority of the Dutch participants had a household income between 50-75% of the country average (47%), while the majority of the Swiss participants had an income of 25-50% of the country average (37%). Distance to agricultural areas was similar in both populations, with an average of 976 (NL) and 979 (CH) meters to agricultural areas, but distance to forest areas was higher in Switzerland (NL: average of 271 m; CH: average of 566 m). The distribution of participants based on their distance to both areas is included in Appendix A (Figure A.5).

Table 1. *Participant characteristics of the Dutch (NL) and Swiss (CH) studies*

Country	**Netherlands (NL)**	**Switzerland (CH)**
Participants, n	105	295
Gender female, n (%)	73 (69.5)	136 (46.1)
Mean age, years [min-max]	42.1 [29-56]	30.8 [20-39]
BMI, n (%)		
Normal/Underweight (<25)	129 (72.3)	217 (73.6)
Overweight (25-30)	60 (22.9)	64 (21.7)
Obese (>30)	18 (4.8)	14 (4.7)
Education level, n (%)		
No or only primary education	1 (1.0)	2 (0.7)
Secondary education	5 (4.8)	1 (0.3)
Tertiary education (post-secondary)	19 (18.1)	67 (22.7)
University (BSc, MSc, PhD)	76 (72.3)	221 (74.9)
Don't Know/ NA	4 (3.8)	4 (1.4)
Household income[1], % of country average		
< 25%	1 (1.0)	61 (20.7)
25-50%	6 (5.7)	109 (36.9)
50-75%	49 (46.7)	47 (15.9)
>75%	33 (31.4)	53 (18.0)
Don't Know/NA	16 (15.2)	25 (8.5)
Mean distance to agricultural areas, m [min-max]	976 [21-2618]	979 [24-2221]
Mean distance to forest, m [min-max]	271 [0-1739]	566 [8-1294]

1. Income categories from the Swiss questionnaire were assigned to the <25th, 25th - 50th, 50th – 75th and >75th percentile categories based on the publication by the Swiss Federal Department of Finance (2014): https://biblio.parlament.ch/e-docs/377581.pdf

Pesticide distribution in the study samples

A total of 37 biomarkers were confirmed with high confidence (Schymanski levels 1 and 2b), relating to 27 different parent pesticides. An overview of all 37 biomarkers (including detected metabolites and types of pesticides) are presented in Appendix A (Table A.5). A graphical presentation of the detection rates and the distribution of the signal intensity scores (log-transformed) of the 13 most frequently detected biomarkers (related to 12 parent pesticides) are shown in Figure 1. Supplementary Figure A.1 (Appendix A) shows the distribution of the intensity scores and the detection rates of all 37 biomarkers. The magnitude of detection rates was comparable in both countries, but detection rates were generally higher in the Dutch population than in the Swiss population, regardless of the between-country differences in the biomarkers' intensity scores. Metabolites of the three pesticides acetamiprid (-CH2, insecticide), flonicamid (-C2HN, insecticide), and chlorpropham (4-HSA, herbicide) had detection rates above 40% in both the Dutch and Swiss study samples. In the Netherlands, also propamocarb (+O, fungicide) and pirimiphos-methyl (-CH2, insecticide/acaricide) were detected in at least 40% of the samples. Biomarkers of an additional eight pesticides, namely fludioxonil (+O +C6H8O6), boscalid (+O +SO3), pyrimethanil (+O +SO3), fluazifop (parent), clothianidin (parent), propamocarb (parent), cyprodinil (+O +SO3), and tebuconazole (-2H +2O), were detected in at least 20 urine samples in each country (total detection ratio of >10%). The intensity scores of each biomarker were comparable for both countries. The highest intensity scores were found for metabolites of propamocarb (+O and parent), yet intensity score differences between biomarkers cannot be directly translated into concentration differences.

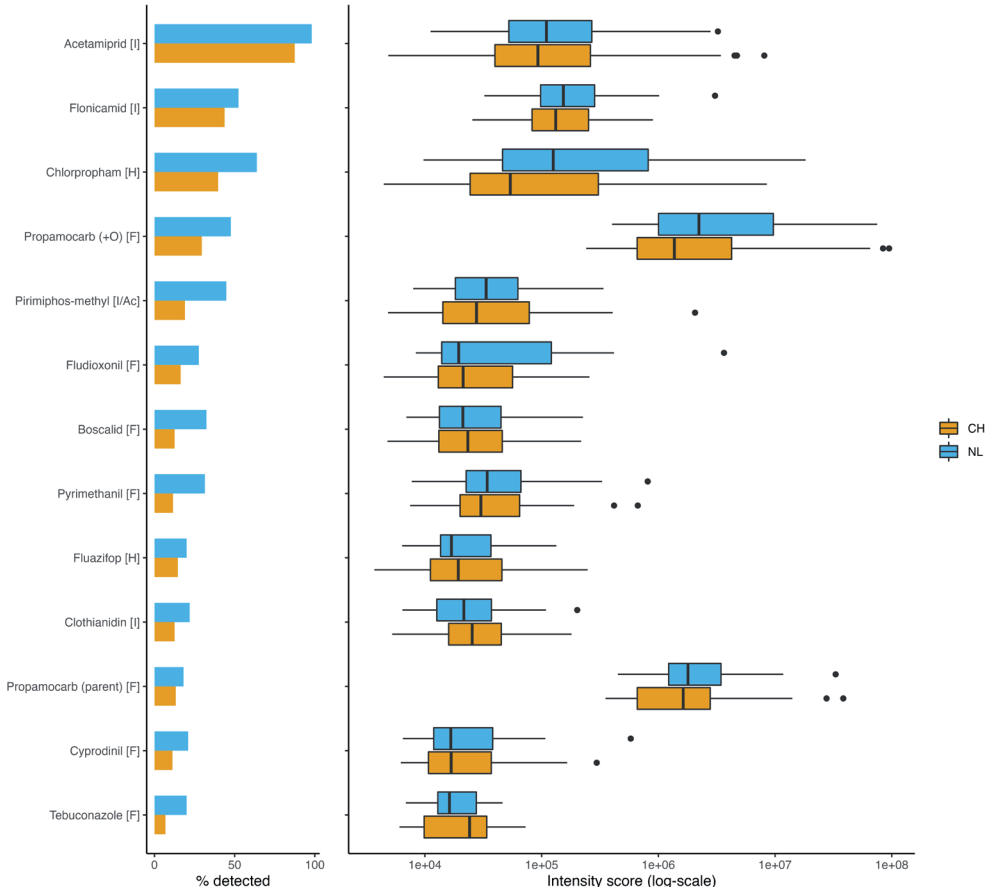


Figure 1. *The distribution of the percentage detected and the intensity scores of the 13 most frequently detected biomarkers for pesticides (noted as parent: F(ungicide), I(nsecticide), H(erbicide), Ac(aricide)), based on n=105 samples from the Netherlands (NL) and n=295 samples from Switzerland (CH). Note: Signal intensity scores reported here are semi-quantitative and can therefore not be directly translated into urine concentration levels.*

Based on the 13 most frequently detected biomarkers, the combinations that co-occurred the most in each country are shown in Figure 2. Only combinations occurring at least four times (arbitrary cut-off point) are shown, resulting in a co-occurrence pattern based on eight biomarkers and 17 different combinations. In 211 out of 400 samples (52.8%), at least two biomarkers were detected. The most frequent co-occurring pesticide biomarker patterns included different combinations of acetamiprid-flonicamid-chlorpropham, of which the acetamiprid-flonicamid combination occurred in 22 urine samples, or 5.5%. In general, the frequency of co-occurrence for a specific combination of biomarkers was low. In the Netherlands, fewer co-occurrences were detected (relatively), while in Switzerland more co-occurrences were identified, with the most common pair acetamiprid-flonicamid found in 7.1% of the samples (see Appendix A, Figure A.6).

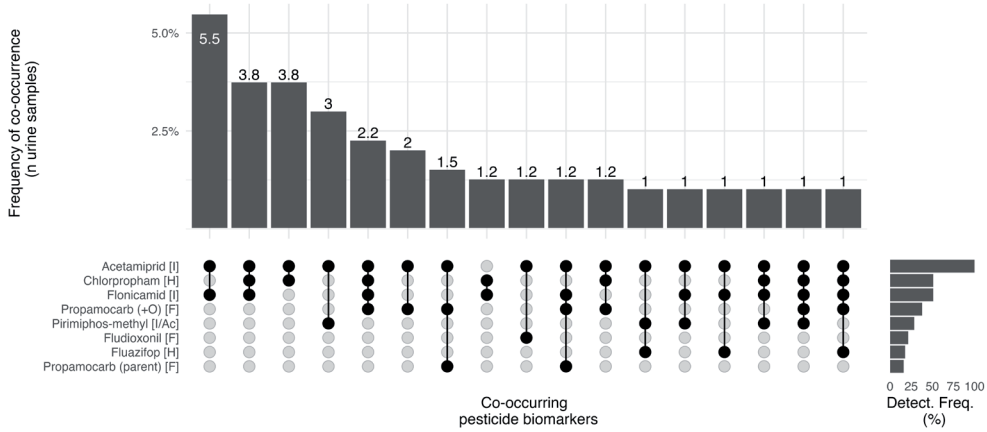


Figure 2. *Frequency of co-occurrences of the 13 most detected pesticide biomarkers in the pooled dataset, (noted as parent: F(ungicide), I(nsecticide), H(erbicide), Ac(aricide)). Only the most frequent combinations (in at least four urine samples) are presented, based on n=105 samples from NL and n=295 samples from CH.*

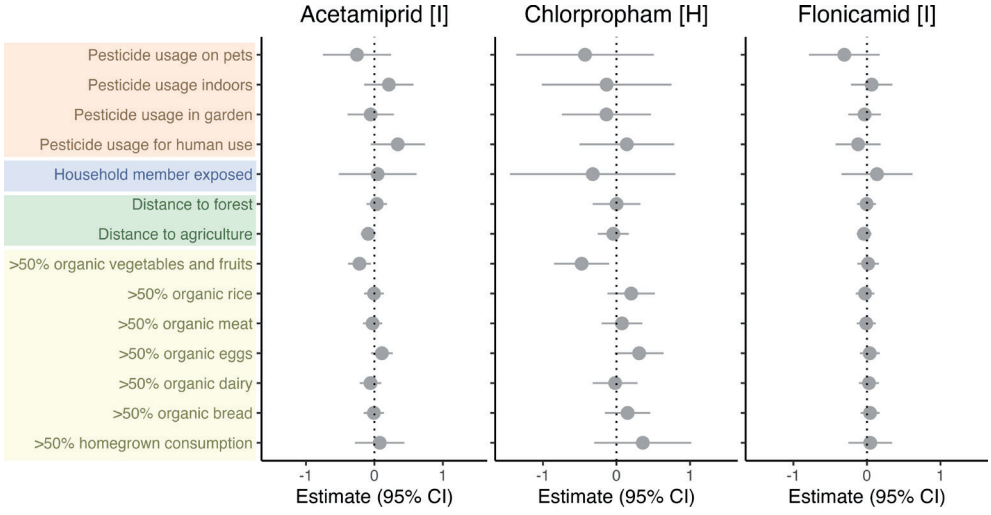


Figure 3. *Association between potential exposure determinants with the intensity scores of acetamiprid, chlorpropham, and flonicamid in urine; results of the Tobit regression models for the pooled dataset. All models were corrected for age, gender, BMI, level of education, income, and country. Factors related to pesticide usage (orange box), household member exposure (blue), distance to agriculture/forest (green), and diet (yellow) are shown. All variables are mutually adjusted.*

Temporal variation in pesticide distribution

Since the study design in the Netherlands included two samples per individual from two different seasons, we utilized this opportunity to calculate the ICC based on the intensity scores. The ICC values (considering within-individual and between-season variation) for all 13 biomarkers were low (<0.3), indicating high within-person variability.

The Swiss samples were collected across multiple seasons (ranging from January until October 2020). Seasonal averages of intensity scores (winter, spring and summer) for the Swiss samples showed no temporal differences. Results of the ICC calculations (Table A.6) and Swiss seasonal averages in boxplots (Figure A.2) are displayed in Appendix A.

Determinants of exposure to acetamiprid, chlorpropham and flonicamid

For acetamiprid, chlorpropham, and flonicamid, censored linear regression (Tobit) models were constructed to explore the potential role of exposure determinants. The covariate mutually adjusted associations of product usage (orange box around the variable names), occupational exposure of household members (blue), homegrown food consumption (yellow), organic diet (yellow), and distance to agricultural areas and forest (green) with the intensity score of the respective metabolite in the pooled study sample are shown in Figure 3 by forest plots. For the country-stratified analyses, see Appendix A (Figure A.3).

The only discernible association for the pooled data models was a lower urinary intensity score for acetamiprid and chlorpropham when organic vegetables and fruit were frequently (>50%) consumed. For flonicamid, no effect was detected in the pooled data model.

Determinants of exposure to biomarkers detected in <40% of samples

For the 10 biomarkers detected in between 10% and 40% of the samples, logistic regression models revealed no discernible association (log odds) with any potential determinant across biomarkers and countries. For the pooled dataset, forest plots for each biomarker are presented in Figure 4. Country-specific results of the logistic regression models can be found in Appendix A (Figure A.4). A high organic vegetable and fruit consumption was associated with a lower detection rate in the pooled data models for propamocarb (+O) and pyrimethanil (+O +SO3). For fluazifop, a greater distance to forest areas was associated with an increase in detection rate in the pooled data models. In the pooled data model for pyrimethanil, distance to agricultural areas was positively associated with the detection rate. In general, the log odds of the non-dietary determinants had large confidence intervals, mainly due to low numbers of occurrence in the study population (see Appendix A, Table A.3). In addition, some exposure variables dropped out of the regression models for certain biomarkers due to the same reason of low occurrence (see Appendix, Table A.3).

DISCUSSION

The present study examines and compares human biomonitoring data on pesticides collected in a sample of the adult population in the Netherlands and in Switzerland as part of the HBM4EU project. Overall, 37 biomarkers (relating to 27 parent pesticides) were detected in 400 urine samples by a suspect screening methodology conducted at the same laboratory. The pesticides present in the urine samples obtained in the two countries were comparable, despite some differences in population characteristics. Detection rates were typically low, co-occurrence of biomarkers not common, and temporal variation at the level of individuals high. Detection rates were highest for acetamiprid, chlorpropham and flonicamid. We observed that consumption of organic fruit and vegetables was an important determinant for exposure to several of the measured pesticide metabolites. In contrast, no clear association of other determinants, such as non-occupational pesticide usage, household member's exposure, distance to agricultural or forest areas, and other dietary habits, with signal intensity and exposure probability was found.

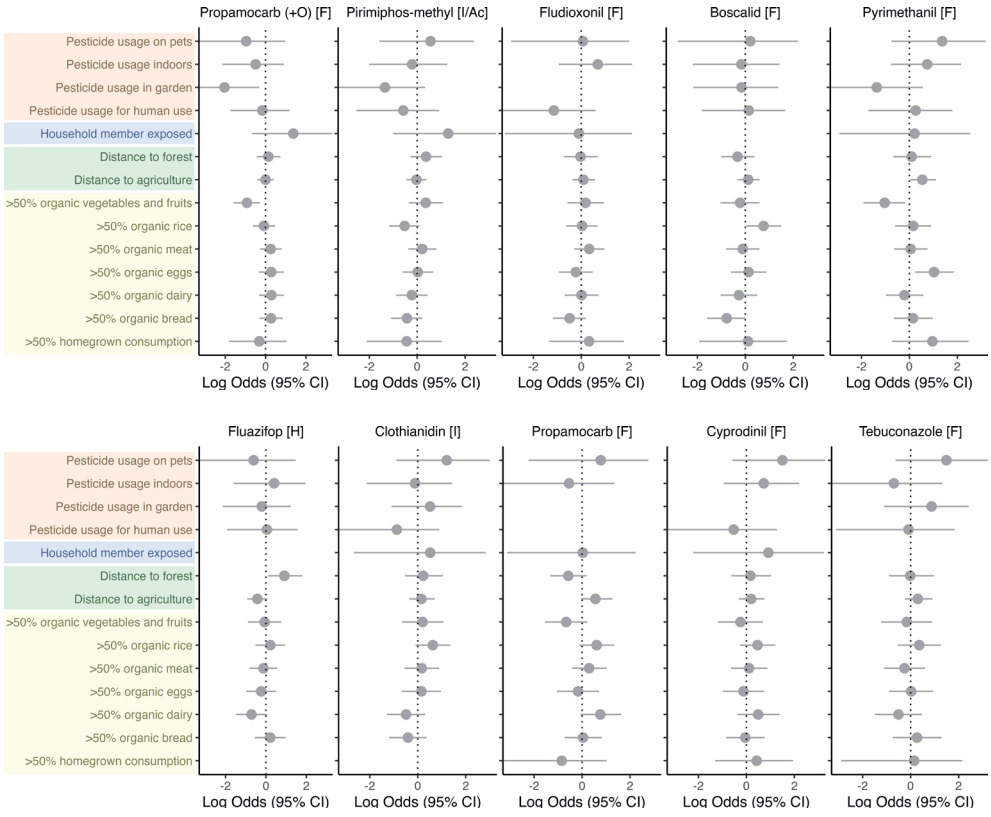


Figure 4. *Association between potential exposure determinants with the urinary presence of biomarkers detected 10-40%; results of the logistic regression models for the pooled dataset. All models were corrected for age, gender, BMI, level of education, income, and country. Factors related to pesticide usage (orange box), household member exposure (blue), distance to agriculture/forest (green), and diet (yellow) are shown. All variables are mutually adjusted. Note: Odds for certain exposure variables are missing due to low frequency of participants using pesticides, e.g. in the garden. Hence, variation in the detected biomarkers was too low for the variable to be included in the regression model.*

Detection rates and exposure pathways

Despite the considerable number of detected biomarkers (confirmed by molecular structure, n=37), only three biomarkers of the parent pesticides acetamiprid, chlorpropham and flonicamid were detected in at least 40% of all samples. In the Dutch data, two additional compounds (pirimiphos-methyl and propamocarb) had a detection rate of ≥40%. These 3 most frequently detected parent pesticides were also part of 8 selected pesticides for targeted analysis by the OBO (Research on exposure of residents to pesticides in the Netherlands) study, based on their usage frequency, monitoring data, analytical possibilities, and possible exposure of the residential population Figueiredo et al. (2021). Two of our frequently detected biomarkers (chlorpropham and flonicamid) have hardly been studied thus far, but our data suggests including these markers in future pesticide exposure studies.

The high detection frequency in the Dutch (98%) and the only slightly lower detection frequency in the Swiss data (87%) indicate ubiquitous exposure to acetamiprid, despite its relatively quick excretion time. The negative association between high organic vegetable and

fruit consumption and acetamiprid exposure in our study points to conventionally grown vegetables and fruits as potential determinants of exposure. Acetamiprid is a neonicotinoid (insecticide) and is approved in the EU as well as in Switzerland for professional use on mainly fruit trees and vegetables, as well as for non-occupational use (only certain acetamiprid-containing products). Due to neonicotinoids' systemic mechanism of action, i.e. their ability to enter and persist in plant tissue, residues of neonicotinoids in food cannot be removed by peeling or washing (Magalhaes et al., 2009; Simon-Delso et al., 2015). Hence, for the general population, fruit and vegetable intake is likely to be the main exposure pathway and target to reduce exposure to acetamiprid (Zhang et al., 2018; Zhang and Lu, 2022). Based on the EFSA database, residues of acetamiprid are mainly found in cherries (48%), chili peppers (38%), pomelos (30%), roman rocket (29%), and pears (26%) (see Appendix A, Table A.4). While data on cherries, chili peppers, pomelos, and roman rocket consumption is not available for the Dutch and Swiss study population, leafy greens were consumed by 32% of the Dutch, and 54% of the Swiss participants. Pears were consumed slightly less in Switzerland (NL: 11%, CH: 7%).

Detection rates for chlorpropham, an herbicide and plant growth regulator, were high in both the Dutch (63%) and Swiss (40%) samples, indicating frequent exposure in both populations. Our Tobit regression results for chlorpropham point to dietary exposure as an important exposure pathway, showing a negative association between organic vegetable and fruit consumption and chlorpropham exposure. The pesticide is mainly used to prevent sprouting of potatoes during storage, and the application is usually done using fogging or spraying equipment (European Food Safety Authority et al., 2017). EU and Swiss approval for chlorpropham was withdrawn in 2019, but periods of grace lasted until autumn 2020 (European Commission, 2019). Hence, exposure through diet in the two study populations was still possible and likely in the year 2020. EFSA data shows that residues of chlorpropham can be found in 15-29% of potatoes (see Appendix A, Table A.4). The relatively high consumption of potatoes in the Netherlands (72 kg/year), as compared to Swiss consumption (47 kg/year), might explain the difference in detection rates (Helgi Library, 2020). Results from the food frequency questionnaires (FFQ) additionally show frequent consumption of potatoes in both the Dutch and Swiss study population (see Appendix A, Table A.4), with 38.1% of Dutch and 28.7% of Swiss participants stating they consumed potatoes within 24 hours before urine collection.

The insecticide flonicamid is authorized for occupational use in both countries and is mainly applied on fruit, vegetables, wheat, and potatoes. Detection rates in the Netherlands were slightly higher (52%) than in Switzerland (43%), and point towards a frequent exposure to flonicamid in both populations. The regression model for flonicamid exposure in the Netherlands indicates that having a household member who is occupationally exposed to pesticides (self-assessment by the participant) is associated with higher exposure. The occupational exposure to flonicamid specifically in orchards was confirmed in another study as well (Zhao et al., 2015). Preferential consumption of organic bread was also associated with lower exposure to flonicamid in the Dutch model. In the Dutch study population, bread was consumed by 83.8% participants within 24 hours before urine collection, as compared to 49% in the Swiss population. The top-5 flonicamid-contaminated food items in the EFSA database, however, do not include bread or other wheat products (see Appendix A, Table A.4). Instead, residues

of flonicamid were mainly found in cucumbers, sweet peppers, peas, peaches, and brussel sprouts. In the Dutch study, data on the consumption of sweet peppers (30%) and peas (12%) is available, but the Swiss FFQ did not inquire about these food items.

The logistic models for propamocarb (+O), pirimiphos-methyl, fludioxonil, fluazifop, clothianidin, propamocarb (parent), boscalid, cyprodinil, pyrimethanil, and tebuconazole revealed no consistent direction of association between the determinants and exposure probability across biomarkers and countries. However, a high consumption of organic vegetables and fruit was associated with lower exposure probability for propamocarb (+O) and pyrimethanil. Based on the EFSA database, residues of propamocarb are found in about 24% of lettuces, and pyrimethanil is found in about 35-45% of citrus fruit (see Appendix A, Table A.4). Data on the consumption of lettuces and various citrus fruits within 24 hours before urine collection is not available for the Dutch and Swiss study population.

Our findings add evidence to previous studies indicating that food choices have an influence on pesticide exposure in the general population (Fortes et al., 2013; Ye et al., 2015; Rempelos et al., 2022). Especially for vegetable and fruit consumption, prior research consistently shows negative correlations between organic food consumption and urinary pesticide concentrations (Baudry et al., 2019; Hyland et al., 2019; Baudry et al., 2021). In our study, a high consumption of organic fruits and vegetables was related to a lower exposure to four biomarkers. However, for the other food groups, the direction of association varied, with high levels of uncertainty. Although not assessed in this study, consumption of imported foods might explain rather small differences in detection frequencies for compounds which are applied in much larger quantities in the Netherlands (see Appendix A, Table A.4) as compared to Switzerland, such as chlorpropham (NL: 39t, CH: 0.06t).

Within-individual variability and mixture exposure patterns

For all biomarkers, correlation between winter and summer season samples of the same individual in the Dutch population was low (≤0.3). This indicates a high within-individual variability of exposure and pesticide levels in urine. Potentially important sources of within-individual variability are changes in lifestyle, including dietary habits, and environmental influences. Longer-term exposure profiles are not well captured in the light of quick metabolization of most pesticides and short biological half-lives (Egeghy et al., 2011). This is in line with previous research on pesticide exposure levels over time, showing high within-individual variability of pesticide levels in urine (Morgan et al., 2016; Li et al., 2019). High within-individual variability was also detected in occupationally highly exposed groups, which can lead to challenges in capturing exposure windows (Fuhrimann et al., 2020). Similarly, a previous study analyzing the Dutch SPECIMEn and HBM4EU data of four other countries found no consistent effect of season on detection frequencies (Ottenbros et al., 2023). Considering also the absence of seasonal differences in average exposure in Switzerland, day-to-day variations of lifestyle and environmental exposures may be more important drivers of exposure than seasonal variations.

Based on the exploration of co-occurrence of biomarkers within the same sample, the most frequent combination (acetamiprid with flonicamid) was only detected in 5.5% (n = 22) of the samples. The 17 most frequent combinations (found in at least four urine samples)

were based on different variations of eight biomarkers, which also reflect the most frequently detected biomarkers in the studies. This points towards individualized and variable pesticide mixture exposure profiles among the general population, as found in previous HBM studies (Aerts et al., 2018; Ottenbros et al., 2023). This is also in agreement with the observed high within-person variability of exposure in our study. The fewer co-occurrence patterns in the Netherlands likely can be explained by the overall smaller number of biomarkers detected in comparison to Switzerland.

Strengths and limitations

The harmonized data collection with questionnaires filled in at the time of urine sampling and standardized urine sample analysis at the same laboratory within the HBM4EU project allowed for the joint analysis of the two datasets. The comparison of pesticide mixtures and exposure pathways in adult populations across the Netherlands and Switzerland was additionally justified by similar pesticide regulations in both countries. The employment of an innovative SS methodology offered the opportunity to semi-quantitatively measure exposure to a large number of compounds previously rarely examined within a single study. Despite the informative insights gained from our study, a few limitations have to be addressed. Regarding sample collection, it should be noted that many pesticides are metabolized and excreted quickly. Hence, the distribution of individual long-term exposures will not adequately be captured by the collection of one first morning void urine sample. Longitudinal and repeated study designs (or longer sampling times, such as 24h voids) will be necessary to adequately monitor temporal variations and estimate temporally integrated pesticide exposure in the general population. We should also point out that part of the urine samples from Switzerland and all samples from the summer season in the Netherlands were collected during the COVID-19 pandemic. This might have affected diet, daily activities, and habits (Bertrand et al., 2021).

Second, a myriad of labels for organic and biological foods exist within the EU and in Switzerland. Thus, it might not have been straightforward for participants to declare how much of their usual food intake is produced organically and participants' might have based their answers on different labels. Additionally, the employed FFQs did not query the consumption of specific food items frequently contaminated with pesticide residues, as reported by the EFSA database. Hence, we did not include single food items in the regression models, which could have diminished the ability to detect any effect. Future studies may profit from a more detailed FFQ that is better aligned with pesticide exposure databases. Additionally, there might be country-specific differences in the proportion of consumption of imported foods, which was not assessed in this study.

Third, we did not observe any effect of distance to agricultural or forest areas on the exposure estimates. Although it must be noted that definitions for agricultural and forest areas were different in the two geospatial datasets for the two study countries (see Table A.1), with the Dutch definitions being more precise. In addition, in contrast to the Swiss sample, the Dutch study design focused specifically on distance to orchards.

Lastly, although the SS approach is a useful analytical methodology to explore exposure to large numbers of pesticides, targeted methods have a higher specificity and sensitivity (Pour-

chet et al., 2020). Moreover, due to time and budget reasons, experts reviewed and prioritized the tentative annotations before starting the compound confirmation workflow, resulting in a suspect screening analysis biased toward halogenated and PO3-containing pesticides (Huber et al., 2022). The list of 37 identified biomarkers is additionally limited by technical possibilities and conventions. Therefore, technological advances might increase the number of identified biomarkers in the future. This might be one of the reasons why the three most commonly detected pesticides might not reflect the most commonly used pesticides in the Netherlands and Switzerland.

For the most commonly detected pesticides, targeted methods will need to be applied for a more precise estimation of determinants. In addition, future studies should carry out improvements regarding the compilation of the list of tentative annotations.

Nonetheless, taking these limitations into account, the results of this multi-country study contribute to the growing field of HBM of pesticides and offer first insights into pesticide mixture patterns and exposure sources and pathways in two countries in the European context. Future studies with a more detailed dietary and behavioral assessment, as well as targeted quantitative, ideally multi-biomarker, screenings of several HBM samples will be able to draw on these results for a more complete assessment of the general populations' exposure to pesticides and determinants thereof.

CONCLUSION

Using a semi-quantitative suspect screening approach, 37 well-annotated pesticide metabolites (relating to 27 parent pesticides) were present in urine collected from participants of the adult population in the Netherlands and Switzerland. Detection rates were typically low, yet three pesticides (acetamiprid, chlorpropham, flonicamid) were detected in at least 40% of the samples at both study sites. High consumption of organic fruits and vegetables was associated with decreased urinary levels for acetamiprid, chlorpropham, propamocarb and pyrimethanil. The suspect screening applied in this study provides an example of how a first-tier screening exercise for pesticide exposure can be conducted. Our study provides an indication for target biomarkers to include in follow-up studies dedicated to the quantification of urinary exposure levels. Also, it highlights the importance of repeated sampling in light of substantial within-individual variability, as well as food contamination reduction as a preventive target to lower pesticide exposures.

CRediT author contribution statement

JV, NPH, SF, RV - Conceptualization; **HM, RN, AL** - Data curation; **IO, PA** - Formal analysis; **NPH, RV, EL, JV** - Funding acquisition; **IO, MI, JPZ** - Investigation; **IO, PA, JV, NPH, SF** - Methodology; **JV, NPH** - Project administration; **IO, PA** - Visualization; **IO, PA** - Writing - original draft; **All Authors** - Writing - review & editing.

Funding statement

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 733032 HBM4EU, No 857560 and No 857340. HM, AL and RN acknowledge the Dutch Ministry of Agriculture, Nature and Food Quality for co-funding through projects 1277354201 and 1267333701. The Swiss participation in this European Program is funded by the Swiss State Secretary for Education Research and Innovation (SERI).

Acknowledgements

Max Oosterwegel for his help with the calculation of the ICCs.

Daan Leseman, Paul Schwillens, Paul Fokkens, and Karin Boer for their help with the Dutch sample collection.

Thomas Vermes, Tamara Wirth, Melissa Witzig, Susanna Nussbaumer, and Nicolas Bringolf for their help with the Swiss sample collection.

Abbreviations

CH	Switzerland
EKNZ	Ethikkommission Nordwest- und Zentralschweiz (local ethics committee in Switzerland)
HBM4EU	European Human Biomonitoring initiative
MREC	Medical Research Ethics Committee Utrecht;
NL	the Netherlands
SPECIMEn	Survey on PEstiCIde Mixtures in Europe
SS	Suspect Screening
WMO	Dutch Medical Research Involving Human Subjects Act

REFERENCES

Aerts R, Joly L, Szternfeld P, Tsilikas K, De Cremer K, Castelain P, Aerts J-M, Van Orshoven J, Somers B, Hendrickx M. Silicone wristband passive samplers yield highly individualized pesticide residue exposure profiles. *Environmental science & technology*, 2018; 52:1: 298-307. DOI: https://doi.org/10.1021/acs.est.7b05039.

Angerer J, Ewers U, Wilhelm M. Human biomonitoring: state of the art. *International journal of hygiene and environmental health*, 2007; 210:3-4: 201-228. DOI: https://doi.org/10.1016/j.ijheh.2007.01.024.

Attfield KR, Hughes MD, Spengler JD, Lu C. Within- and between-child variation in repeated urinary pesticide metabolite measurements over a 1-year period. *Environmental Health Perspectives*, 2014; 122:2: 201-206. DOI: https://doi.org/10.1289/ehp.1306737.

Baudry J, Debrauwer L, Durand G, Limon G, Delcambre A, Vidal R, Taupier-Letage B, Druesne-Pecollo N, Galan P, Hercberg S. Urinary pesticide concentrations in French adults with low and high organic food consumption: results from the general population-based NutriNet-Santé. *Journal of exposure science & environmental epidemiology*, 2019; 29:3: 366-378. DOI: https://doi.org/10.1038/s41370-018-0062-9.

Baudry J, Rebouillat P, Allès B, Cravedi J-P, Touvier M, Hercberg S, Lairon D, Vidal R, Kesse-Guyot E. Estimated dietary exposure to pesticide residues based on organic and conventional data in omnivores, pesco-vegetarians, vegetarians and vegans. *Food and chemical toxicology*, 2021; 153: 112179. DOI: https://doi.org/10.1016/j.fct.2021.112179.

Bertrand L, Shaw KA, Ko J, Deprez D, Chilibeck PD, Zello GA. The impact of the coronavirus disease 2019 (COVID-19) pandemic on university students' dietary intake, physical activity, and sedentary behaviour. *Applied Physiology, Nutrition, and Metabolism*, 2021; 46:3: 265-272. DOI: https://doi.org/10.1139/apnm-2020-0990.

Bonvallot N, Jamin EL, Regnaut L, Chevrier C, Martin J-F, Mercier F, Cordier S, Cravedi J-P, Debrauwer L, Le Bot B. Suspect screening and targeted analyses: Two complementary approaches to characterize human exposure to pesticides. *Science of The Total Environment*, 2021; 786: 147499. DOI: https://doi.org/10.1016/j.scitotenv.2021.147499.

Buekers J, Remy S, Bessems J, Govarts E, Rambaud L, Riou M, Halldorsson TI, Ólafsdóttir K, Probst-Hensch N, Ammann P, Weber T, Kolossa-Gehring M, Esteban-López M, Castaño A, Andersen HR, Schoeters G. Glyphosate and AMPA in Human Urine of HBM4EU-Aligned Studies: Part B Adults. *Toxics*, 2022; 10:10: 552. DOI: https://doi.org/10.3390/toxics10100552.

Cimino AM, Boyles AL, Thayer KA, Perry MJ. Effects of neonicotinoid pesticide exposure on human health: a systematic review. *Environmental health perspectives*, 2017; 125:2: 155-162. DOI: https://doi.org/10.1289/EHP515.

Dahiri B, Martín-Reina J, Carbonero-Aguilar P, Aguilera-Velázquez JR, Bautista J, Moreno I. Impact of pesticide exposure among rural and urban female population. An overview. *International Journal of Environmental Research and Public Health*, 2021; 18:18: 9907. DOI: https://doi.org/10.3390/ijerph18189907.

Deziel NC, Friesen MC, Hoppin JA, Hines CJ, Thomas K, Freeman LEB. A review of nonoccupational pathways for pesticide exposure in women living in agricultural areas. *Environmental health perspectives*, 2015; 123:6: 515-524. DOI: https://doi.org/10.1289/ehp.1408273.

Egeghy PP, Cohen Hubal EA, Tulve NS, Melnyk LJ, Morgan MK, Fortmann RC, Sheldon LS. Review of pesticide urinary biomarker measurements from selected US EPA children's observational exposure studies. *International journal of environmental research and public health*, 2011; 8:5: 1727-1754. DOI: https://doi.org/10.3390/ijerph8051727.

European Commission. *Commission Implementing Regulation (EU) 2019/989 of 17 June 2019 concerning the non-renewal of approval of the active substance chlorpropham, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market, and amending the Annex to Commission Implementing Regulation (EU) No 540/2011 (Text with EEA relevance.)*. 2019. [Accessed 2023]; Available from: http://data.europa.eu/eli/reg_impl/2019/989/oj.

European Commission. *Consolidated text: Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC*. 2022. [Accessed 2023]; Available from: http://data.europa.eu/eli/reg/2009/1107/2022-11-21.

European Commission. *EU pesticide database*. 2023. [Accessed 2023]; Available from: https://ec.europa.eu/food/plant/pesticides/eu-pesticides-database/start/screen/active-substances.

European Food Safety Authority, Arena M, Auteri D, Barmaz S, Bellisai G, Brancato A, Brocca D, Bura L, Byers H, Chiusolo A, Court Marques D. Peer review of the pesticide risk assessment of the active substance chlorpropham. *EFSA Journal*, 2017; 15:17: e04903. DOI: https://doi.org/10.2903/j.efsa.2017.4903.

European Food Safety Authority, Carrasco Cabrera L, Medina Pastor P. The 2020 European Union report on pesticide residues in food. *EFSA Journal*, 2022; 20:3: e07215. DOI: https://doi.org/10.2903/j.efsa.2022.7215.

Figueiredo DM, Krop EJM, Duyzer J, Gerritsen-Ebben RM, Gooijer YM, Holterman HJ, Huss A, Jacobs CMJ, Kivits CM, Kruijne R, Mol HJGJ, Oerlemans A, Sauer PJJ, Scheepers PTJ, van de Zande JC, van den Berg E, Wenneker M, Vermeulen RCH. Pesticide Exposure of Residents Living Close to Agricultural Fields in the Netherlands: Protocol for an Observational Study. *JMIR Research Protocols*, 2021; 10:4: e27883. DOI: 10.2196/27883.

Fortes C, Mastroeni S, Pilla MA, Antonelli G, Lunghini L, Aprea C. The relation between dietary habits and urinary levels of 3-phenoxybenzoic acid, a pyrethroid metabolite. *Food and Chemical Toxicology*, 2013; 52: 91-96. DOI: https://doi.org/10.1016/j.fct.2012.10.035.

Fuhrimann S, Staudacher P, Lindh C, De Joode BVW, Mora AM, Winkler MS, Kromhout H. Variability and predictors of weekly pesticide exposure in applicators from organic, sustainable and conventional smallholder farms in Costa Rica. *Occupational and environmental medicine*, 2020; 77:1: 40-47. DOI: http://dx.doi.org/10.1136/oemed-2019-105884.

Ganzleben C, Antignac J-P, Barouki R, Castaño A, Fiddicke U, Klánová J, Lebret E, Olea N, Sarigiannis D, Schoeters GR. Human biomonitoring as a tool to support chemicals regulation in the European Union. *International journal of hygiene and environmental health*, 2017; 220:2 Pt A: 94-97. DOI: https://doi.org/10.1016/j.ijheh.2017.01.007.

Gilden RC, Huffling K, Sattler B. Pesticides and health risks. *Journal of Obstetric, Gynecologic & Neonatal Nursing*, 2010; 39:1: 103-110. DOI: https://doi.org/10.1111/j.1552-6909.2009.01092.x.

Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, 2009; 42:2: 377-381. DOI: https://doi.org/10.1016/j.jbi.2008.08.010.

Heffernan A, English K, Toms L, Calafat A, Valentin-Blasini L, Hobson P, Broomhall S, Ware R, Jagals P, Sly P. Cross-sectional biomonitoring study of pesticide exposures in Queensland, Australia, using pooled urine samples. *Environmental Science and Pollution Research*, 2016; 23: 23436-23448. DOI: https://doi.org/10.1007/s11356-016-7571-7.

Helgi Library. *Potato Consumption Per Capita*. 2020. [Accessed 2023]; Available from: https://www.helgilibrary.com/indicators/potato-consumption-per-capita/.

Henningsen A. *censReg: Censored Regresion (Tobit) Models. R package version 0.5*. 2022. [Accessed 2023]; Available from: http://CRAN.R-Project.org/package=censReg.

Hernández AF, Gil F, Lacasaña M. Toxicological interactions of pesticide mixtures: an update. *Archives of toxicology*, 2017; 91: 3211-3223. DOI: https://doi.org/10.1007/s00204-017-2043-5.

Huber C, Nijssen R, Mol H, Antignac JP, Krauss M, Brack W, Wagner K, Debrauwer L, Vitale CM, Price EJ. A large scale multi-laboratory suspect screening of pesticide metabolites in human biomonitoring: From tentative annotations to verified occurrences. *Environment International*, 2022; 168: 107452. DOI: https://doi.org/10.1016/j.envint.2022.107452.

Hyland C, Bradman A, Gerona R, Patton S, Zakharevich I, Gunier RB, Klein K. Organic diet intervention significantly reduces urinary pesticide levels in US children and adults. *Environmental research*, 2019; 171: 568-575. DOI: https://doi.org/10.1016/j.envres.2019.01.024.

Li AJ, Martinez-Moral M-P, Kannan K. Temporal variability in urinary pesticide concentrations in repeated-spot and first-morning-void samples and its association with oxidative stress in healthy individuals. *Environment International*, 2019; 130: 104904. DOI: https://doi.org/10.1016/j.envint.2019.104904.

Magalhaes LC, Hunt TE, Siegfried BD. Efficacy of neonicotinoid seed treatments to reduce soybean aphid populations under field and controlled conditions in Nebraska. *Journal of economic entomology*, 2009; 102:1: 187-195. DOI: https://doi.org/10.1603/029.102.0127.

Morgan MK, Sobus JR, Barr DB, Croghan CW, Chen F-L, Walker R, Alston L, Andersen E, Clifton MS. Temporal variability of pyrethroid metabolite levels in bedtime, morning, and 24-h urine samples for 50 adults in North Carolina. *Environmental research*, 2016; 144: 81-91. DOI: https://doi.org/10.1016/j.envres.2015.11.003.

Ohlander J, Fuhrimann S, Basinas I, Cherrie JW, Galea KS, Povey AC, van Tongeren M, Harding A-H, Jones K, Vermeulen R, Kromhout H. Systematic review of methods used to assess exposure to pesticides in occupational epidemiology studies, 1993–2017. *Occupational and Environmental Medicine*, 2020; 77:6: 357. DOI: http://dx.doi.org/10.1136/oemed-2019-105880.

Ottenbros I, Lebret E, Huber C, Lommen A, Antignac JP, Čupr P, Šulc L, Mikeš O, Szigeti T, Középesy S, Martinsone I, Martinsone Z, Akulova L, Pardo O, Fernández SF, Coscollá C, Pedraza-Diaz S, Krauss M, Debrauwer L, Wagner K, Nijssen R, Mol H, Vitale CM, Klanova J, Molina BG, León N, Vermeulen R, Luijten M, Vlaanderen J. Assessment of exposure to pesticide mixtures in five European countries by a harmonized urinary suspect screening approach. *International Journal of Hygiene and Environmental Health*, 2023; 248: 114105. DOI: https://doi.org/10.1016/j.ijheh.2022.114105.

Pourchet M, Debrauwer L, Klanova J, Price EJ, Covaci A, Caballero-Casero N, Oberacher H, Lamoree M, Damont A, Fenaille F. Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: From promises to challenges and harmonisation issues. *Environment international*, 2020; 139: 105545. DOI: https://doi.org/10.1016/j.envint.2020.105545.

Rempelos L, Wang J, Barański M, Watson A, Volakakis N, Hoppe H-W, Kühn-Velten WN, Hadall C, Hasanaliyeva G, Chatzidimitriou E, Magistrali A, Davis H, Vigar V, Średnicka-Tober D, Rushton S, Iversen PO, Seal CJ, Leifert C. Diet and food type affect urinary pesticide residue excretion profiles in healthy individuals: results of a randomized controlled dietary intervention trial. *The American Journal of Clinical Nutrition*, 2022; 115:2: 364-377. DOI: https://doi.org/10.1093/ajcn/nqab308.

Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, Hollender J. Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environmental Science & Technology*, 2014; 48:4: 2097-2098. DOI: https://doi.org/10.1021/es5002105.

Simon-Delso N, Amaral-Rogers V, Belzunces LP, Bonmatin J-M, Chagnon M, Downs C, Furlan L, Gibbons DW, Giorio C, Girolami V. Systemic insecticides (neonicotinoids and fipronil): trends, uses, mode of action and metabolites. *Environmental Science and Pollution Research*, 2015; 22: 5-34. DOI: https://doi.org/10.1007/s11356-014-3470-y.

Swiss Federal Department of Finance. *Wohlstandsbericht: Einkommen und Vermögen steigen, Einkommensverteilung bleibt weitgehend stabil*. Finance F.D.o. 2014, Bern: Swiss Confederation. https://biblio.parlament.ch/e-docs/377581.pdf.

Teysseire R, Manangama G, Baldi I, Carles C, Brochard P, Bedos C, Delva F. Determinants of non-dietary exposure to agricultural pesticides in populations living close to fields: A systematic review. *Science of The Total Environment*, 2021; 761: 143294. DOI: 10.1016/j.scitotenv.2020.143294.

Vitale CM, Lommen A, Huber C, Wagner K, Garlito Molina B, Nijssen R, Price EJ, Blokland M, van Tricht F, Mol HG. Harmonized quality assurance/quality control provisions for nontargeted measurement of urinary pesticide biomarkers in the HBM4EU multisite SPECIMEn study. *Analytical Chemistry*, 2022; 94:22: 7833-7843. DOI: https://doi.org/10.1021/acs.analchem.2c00061.

Ye M, Beach J, Martin JW, Senthilselvan A. Associations between dietary factors and urinary concentrations of organophosphate and pyrethroid metabolites in a Canadian general population. *International Journal of Hygiene and Environmental Health*, 2015; 218:7: 616-626. DOI: https://doi.org/10.1016/j.ijheh.2015.06.006.

Yusà V, Fernández SF, Dualde P, López A, Lacomba I, Coscollà C. Exposure to non-persistent pesticides in the Spanish population using biomonitoring: A review. *Environmental Research*, 2022; 205: 112437. DOI: https://doi.org/10.1016/j.envres.2021.112437.

Zhang D, Lu S. Human exposure to neonicotinoids and the associated health risks: A review. *Environment International*, 2022; 163: 107201. DOI: https://doi.org/10.1016/j.envint.2022.107201.

Zhang Q, Li Z, Chang CH, Lou JL, Zhao MR, Lu C. Potential human exposures to neonicotinoid insecticides: A review. *Environmental Pollution*, 2018; 236: 71-81. DOI: https://doi.org/10.1016/j.envpol.2017.12.101.

Zhao M-A, Yu A, Zhu Y-Z, Kim J-H. Potential dermal exposure to flonicamid and risk assessment of applicators during treatment in apple orchards. *Journal of Occupational and Environmental Hygiene*, 2015; 12:8: D147-D152. DOI: https://doi.org/10.1080/15459624.2015.1009984.

SUPPLEMENTARY MATERIAL

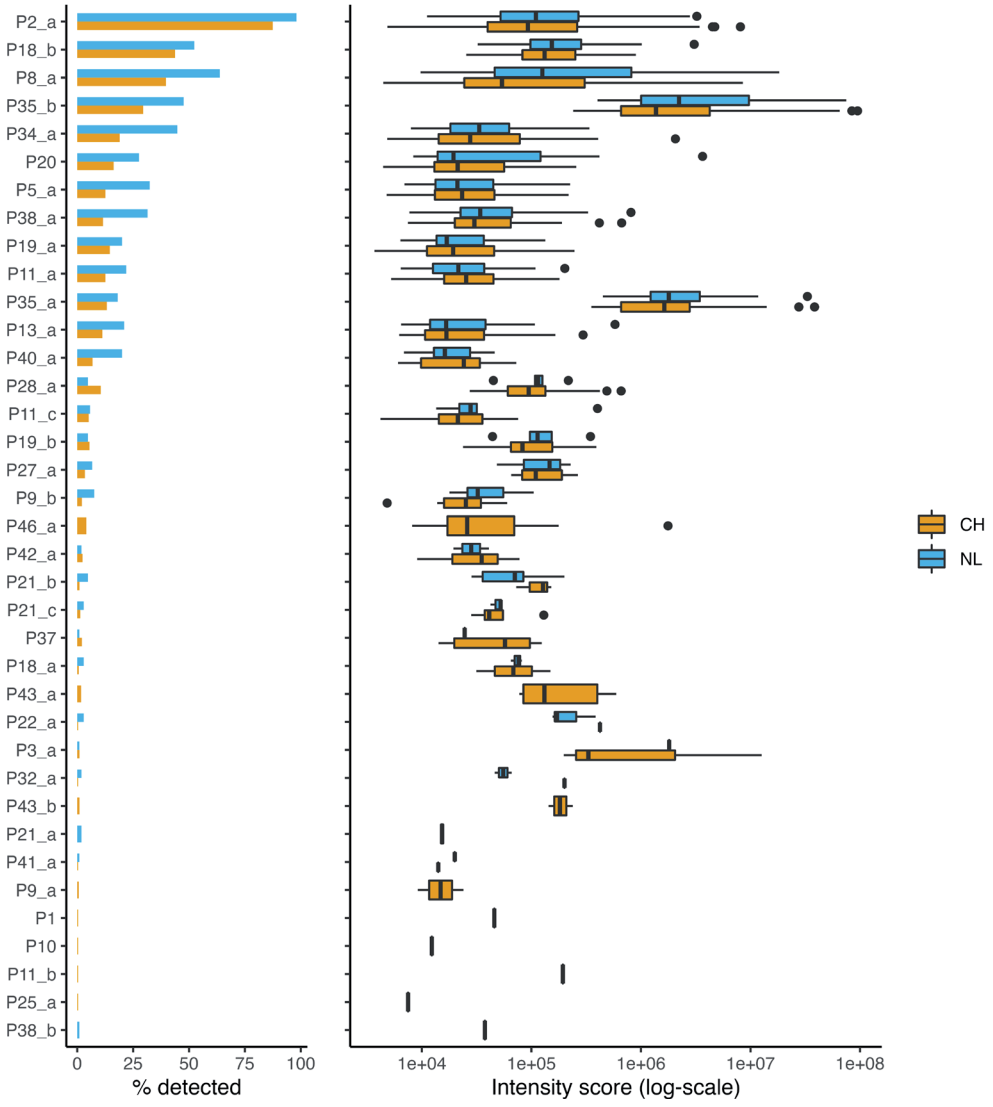


Figure A.1. *The distribution of intensity scores and percentage detected of the 37 most frequently detected biomarkers, based on n=105 samples from the Netherlands (NL) and n=295 samples from Switzerland (CH). For the full name of each biomarker, see Table A.5.*

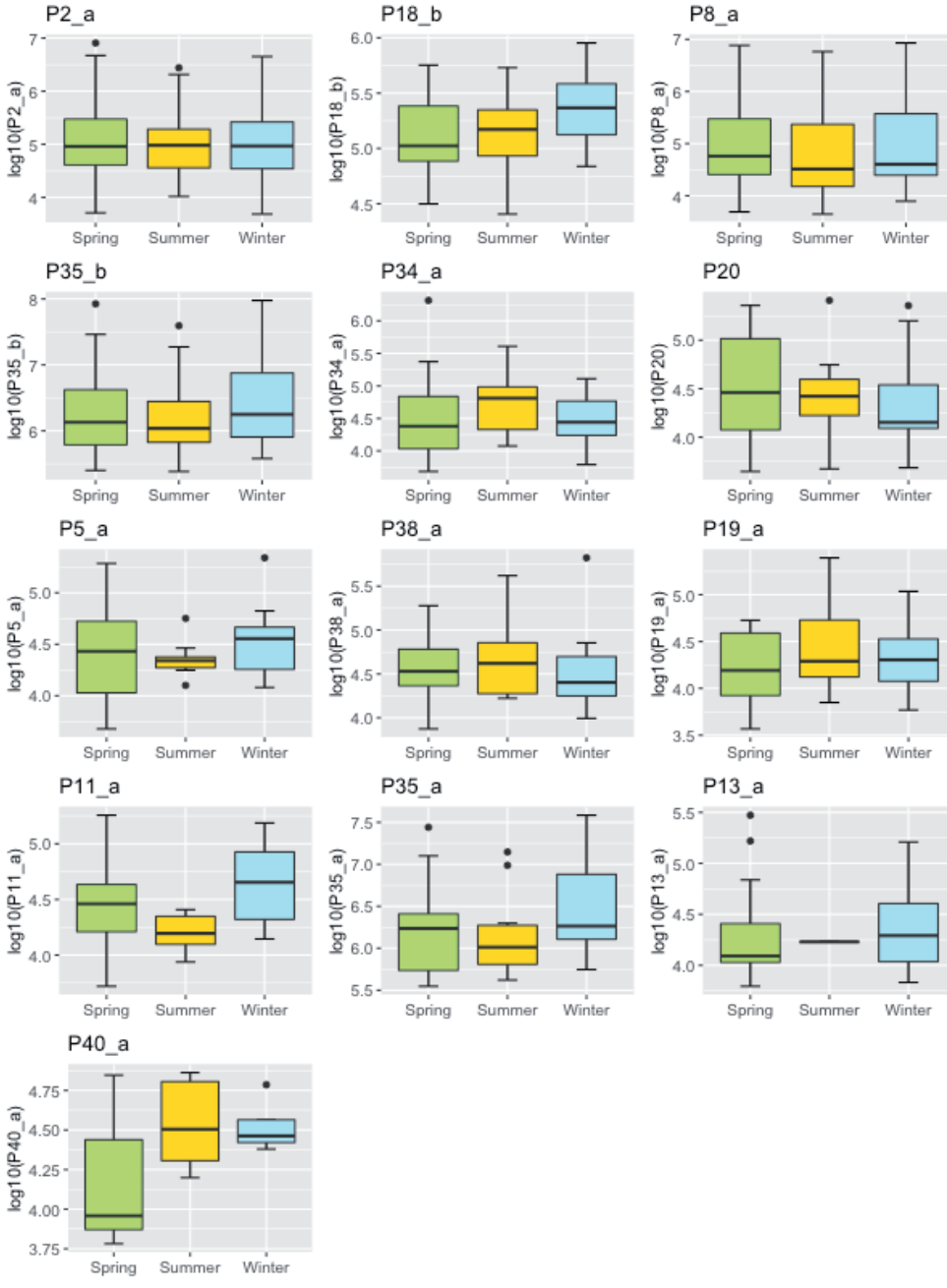


Figure A.2. *Distributions of signal intensity scores (log-transformed) across seasons in the Swiss population.*

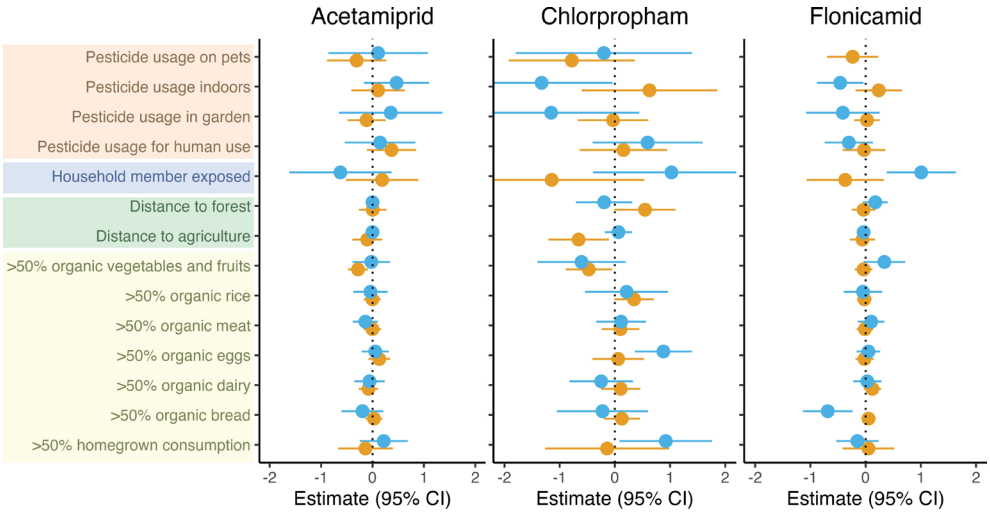


Figure A.3. *Associations between potential exposure determinants and the intensity scores of acetamiprid, chlorpropham, and flonicamid in urine; results of the Tobit regression models for the country-specific dataset. All models were corrected for age, gender, BMI, level of education, and income. Factors related to pesticide usage (orange box), household member exposure (blue), distance to agriculture/forest (green), and diet (yellow) are shown. All variables are mutually adjusted.*

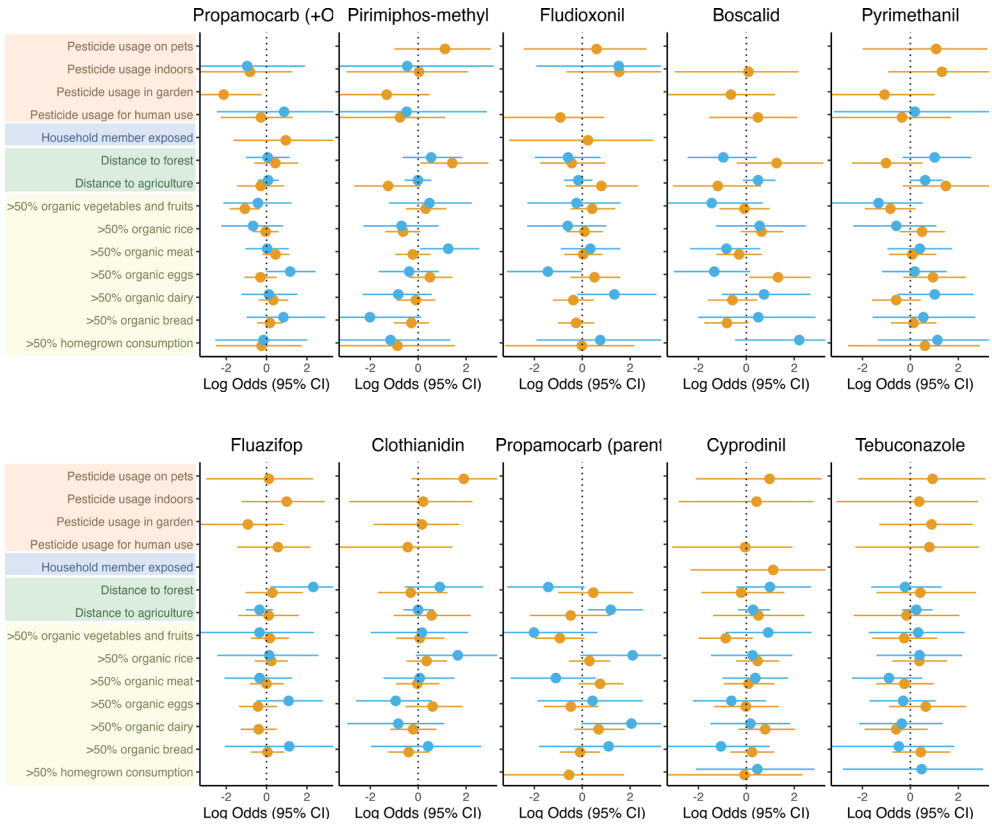


Figure A.4. *Association of exposures with the urinary presence of biomarkers detected 10%-40%; results of the logistic regression models for the country-specific dataset. All models were corrected for age, gender, BMI, level of education, and income. Factors related to pesticide usage (orange box), household member exposure (blue), distance to agriculture/ forest (green), and diet (yellow) are shown. All variables are mutually adjusted.*

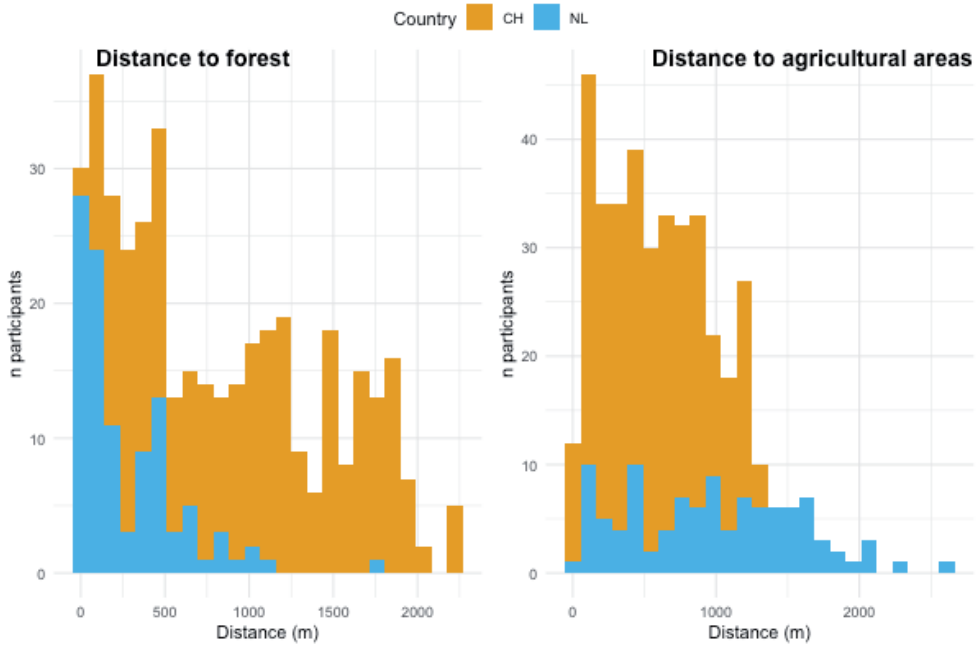


Figure A.5. *Distribution of distance to agriculture and forest for the Netherlands (NL) and Switzerland (CH).*

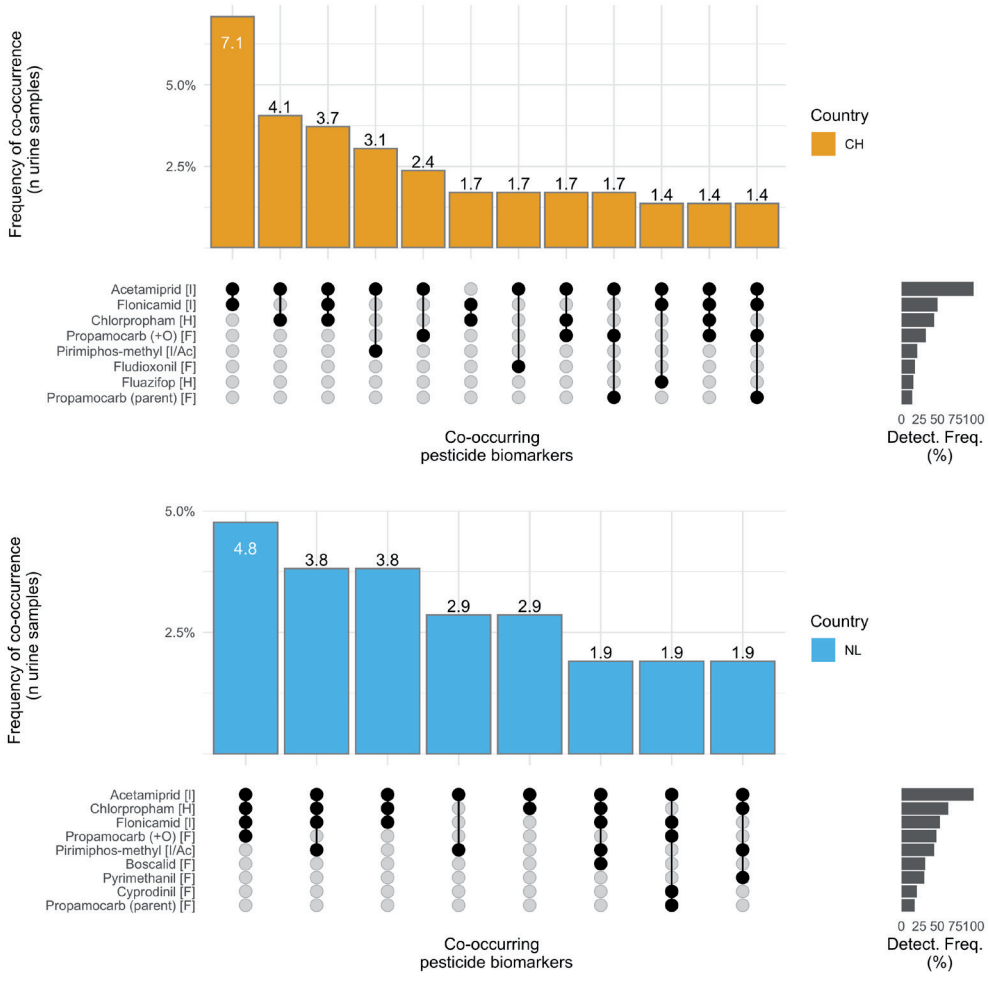


***Figure A.6**. Frequency of co-occurrences of the 13 most detected pesticide biomarkers, for both the Netherlands (NL, blue, left) and Switzerland (CH, orange, right). Based on n=105 samples from the Netherlands (NL) and n=295 samples from Switzerland (CH).*

Table A.1. *Country definitions of agricultural and forest areas used for distance calculations.*

	Netherlands (NL)	**Switzerland (CH)**
Agricultural land	Arable land, farm (excluding detached dwelling house if it can be included as a residential area), grassland, orchard, nursery, cultivated land, grassy dike, high and low stem orchard, apiary, farmland, poplar pasture, horticulture, meadow, pasture; land for greenhouse horticulture includes land used for growing crops under standing glass, in- and adjacent water basins, associated gas tanks;	Arable land, permanent green space, area with permanent crops, greenhouses, area with hedges, riparian and field woodland that is not forest
Forest	Area grown with trees destined for wood production and or nature management. Forest = area with trees where the crowns are about to be closed (or will be closed soon), cut down trees area, burned down trees area, newly planted trees, forest path, tree nursery, wood collection area, spread out houses inside the forest (such as single houses, but not multiple streets), wet forest, poplars field, small waters inside the forest (smaller than 1ha and or smaller than 6m). Not forest = forest areas of parks, not in forest located tree nurseries, residential areas inside the forest area (with multiple streets), open strips for high voltage masts Minimal area = 1ha, minimal 25m in width	Area planted with forest trees or shrubs and capable of performing forest functions; pasture forests, stocked pastures; unstocked or unproductive areas of a forest property, such as forest blocks, forest roads and other forestry structures and facilities

Table A.2. *Comparison and re-coding of questionnaire variables of interest for the analysis from the Dutch and Swiss questionnaires applied on the day of urine sampling.*

Question NL [Unit/Categories]	**Question CH [Unit/Categories]**	**Analysis variable [Unit/Categories]**	**Recoded categories**
Date of birth [DD/MM/YYYY]	Date of birth [DD/MM/YYYY]	Age [YY]	-
Sex [0=M; 1=F; 2=Other]	Sex [1=F; 2=M]	Sex [0=M; 1=F; 2=Other]	CH: M=0
Height [cm] and weight [kg]	BMI [kg/cm^2]	BMI [kg/cm^2]	NL: calculate BMI (weight*10000/ (height*height))
Highest education [1=No education or only primary education ;2=Secondary education; 3=Tertiary education (post-secondary); 4=University (BSc, MSc, PhD); 99=Don't know; 999=NA]	Highest education [1=Lower level than primary school; 2=Primary school; 3=Secondary school; 4=High school/tertiary school; 5=Apprenticeship; 6=Vocational school; 7=University: Bachelor; 8=University: Master; 9=University: Doctorate/ PhD; 999=NA]	Highest education [1=No education or only primary education ;2=Secondary education; 3=Tertiary education (post-secondary); 4=University (BSc, MSc, PhD); 999=Don't know/NA]	CH: combine categories into 4 (Lower level than primary school OR primary school=1; Secondary school=2; High school/tertiary school OR Apprenticeship OR Vocational school=3; University: Bachelor OR University: Master OR University: Doctorate/PhD=4)
Total household income (% of country average) [1=<25 %; 2=25-50 %; 3=50-75 %; 4=>75 %; 99=Don't know; 999=NA]	Total average monthly household income [1=Less than CHF 3'000; 2=Between CHF 3'000.- and CHF 4'500; 3=Between CHF 4'500 and CHF 6'000; 4=Between CHF 6'000 and CHF 9'000; 5=Between CHF 9'000 and CHF 11'000; 6= More than CHF 11'000; 7=I prefer not to answer this question]	Total household income (% of country average) [1=<25 %; 2=25-50 %; 3=50-75 %; 4=>75 %; 999=Don't know/Prefer not to answer/NA]	CH: combine categories into 4 (Less than CHF3'000 OR Between CHF 3'000 and 4'500=<25%=1; Between CHF 4'500 and CHF 6'000 OR Between CHF 6'000 and CHF 9'000=25-50%=2; Between CHF 9'000 and CHF 11'000=50-75%=3; More than CHF 11'000=75%;=4; I prefer not to answer this question=999) NL: combine Don't know and NA (Don't know OR NA=999)
Family members' professional contact with pesticides [1=Yes; 999=NA]	Number of household members coming in contact with pesticides in their profession [1= Nobody; 2=1 Person; 3=2 People; 4=More than 2 people]	Household members' professional contact with pesticides [0= No; 1=Yes; 999=NA]	NL: make 3 categories (not 1 or 999=0) CH: make 3 categories (Nobody=0; 1 Person OR 2 People OR More than 2 people=1; not answered=999]
Pesticide use in garden up to 3 days prior to urine sampling [0=No; 1=Yes; 99=Don't know; 999=NA]	Pesticide use in garden up to 3 days prior to urine sampling [1=No, not on these days; 2=Yes, in the 12 hours before the urine collection; 3=Yes, the day (in the 12 to 24 hours) before the urine collection; 4=Yes, on the 2. day (24 to 48 hours) before the urine collection; 5= Yes, on the 3. day (48 bis 72 Stunden) before the urine collection; 6=I don't know]	Pesticide use in garden up to 3 days prior to urine sampling [0=No; 1=Yes; 999=Don't know/NA]	CH: make 4 categories (No, not on these days=0; 12 hours before OR 12-24 hours before OR 24-48 hours before OR 48-72 hours before=1; I don't know=999; NA=999)

Table A.2. *Continued.*

Question NL [Unit/Categories]	Question CH [Unit/Categories]	Analysis variable [Unit/Categories]	Recoded categories
Pesticide use indoors at home up to 3 days prior to urine sampling [0=No; 1=Yes; 99=Don't know; 999=NA]	Pesticide use indoors at home or at work up to 3 days prior to urine sampling [1=No, not on these days; 2=Yes, in the 12 hours before the urine collection; 3=Yes, the day (in the 12 to 24 hours) before the urine collection; 4=Yes, on the 2. day (24 to 48 hours) before the urine collection; 5= Yes, on the 3. day (48 bis 72 Stunden) before the urine collection; 6=I don't know]	Private pesticide use indoors up to 3 days prior to urine sampling [0=No; 1=Yes; 999=Don't know/NA]	CH: make 4 categories (No, not on these days=0; 12 hours before OR 12-24 hours before OR 24-48 hours before OR 48-72 hours before=1; I don't know=999; NA=999)
Pesticide use on pets up to 3 days prior to urine sampling [0=No; 1=Yes; 99=Don't know; 999=NA]	[1=No, not on these days; 2=Yes, in the 12 hours before the urine collection; 3=Yes, the day (in the 12 to 24 hours) before the urine collection; 4=Yes, on the 2. day (24 to 48 hours) before the urine collection; 5= Yes, on the 3. day (48 bis 72 Stunden) before the urine collection]	Pesticide use on pets up to 3 days prior to urine sampling [0=No; 1=Yes; 999=Don't know/NA]	CH: make 4 categories (No, not on these days=0; 12 hours before OR 12-24 hours before OR 24-48 hours before OR 48-72 hours before=1; NA=999)
Insect repellent use on self-up to 3 days prior to urine sampling [0=No; 1=Yes; 99=Don't know; 999=NA]	Insect repellent use on self-up to 3 days prior to urine sampling [1=No, not on these days; 2=Yes, in the 12 hours before the urine collection; 3=Yes, the day (in the 12 to 24 hours) before the urine collection; 4=Yes, on the 2. day (24 to 48 hours) before the urine collection; 5= Yes, on the 3. day (48 bis 72 Stunden) before the urine collection; 6=I don't know]	Insect repellent use on self-up to 3 days prior to urine sampling [0=No; 1=Yes; 999=Don't know/NA]	CH: make 4 categories (No, not on these days=0; 12 hours before OR 12-24 hours before OR 24-48 hours before OR 48-72 hours before=1; I don't know=999; NA=999)
Pesticide use during DIY activities or hobbies 3 days prior to urine sampling [0=No; 1=Yes; 99=Don't know; 999=NA]	Pesticide use during activities 3 days prior to urine sampling [1=No, not on these days; 2=Yes, in the 12 hours before the urine collection; 3=Yes, the day (in the 12 to 24 hours) before the urine collection; 4=Yes, on the 2. day (24 to 48 hours) before the urine collection; 5= Yes, on the 3. day (48 bis 72 Stunden) before the urine collection]	Pesticide use during DIY activities or hobbies 3 days prior to urine sampling [0=No; 1=Yes; 999=Don't know/NA]	CH: make 4 categories (No, not on these days=0; 12 hours before OR 12-24 hours before OR 24-48 hours before OR 48-72 hours before=1; NA=999)
Professional contact with pesticides in past month [0=No; 1=Yes; 99=Don't know; 999=NA]	Professional contact with pesticides in past month [1=Yes; 2=No; 3=I don't know; 999=NA]	Professional contact with pesticides in past month [0=No; 1=Yes; 999=I don't know/NA]	CH: No=0; I don't know=999

Question NL [Unit/Categories]	**Question CH [Unit/Categories]**	**Analysis variable [Unit/Categories]**	**Recoded categories**
Percentage of diet that was organic in last 6 months [0=No consumption; 1=Consumption; 99=Don't know; 999=NA; % organic]	Percentage of diet that is organic in general [% organic, 999=NA	Percentage of diet that is organic [% organic, 999=I don't know/NA]	NL and CH: make 2 categories (<50% organic; ≥50% organic); for sensitivity analysis: <10% organic, 10-50% organic, >50% organic
Percentage of homegrown food in last 6 months, per season [1=Known; 99=Don't know; 999=NA; % homegrown]	Percentage of homegrown food in last 6 months, per season [1= I don't know; %; NA=999	Percentage of homegrown food per season [% homegrown, 999=I don't know/NA]	CH: I don't know=999 NL and CH: limit to summer season
Food items eaten during 24 hours before urine collection [0=Not eaten; 1=Eaten during the past 24h; 999=NA]	Food items eaten up to 3 days before urine collection [1=No; 2=Yes, in the 12 hours before the urine collection; 3=Yes, the day (in the 12 to 24 hours) before the urine collection; 4=Yes, on the 2. day (24 to 48 hours) before the urine collection; 5= Yes, on the 3. day (48 bis 72 Stunden) before the urine collection]	Food items eaten during 24 hours before urine collection [0=Not eaten; 1=Eaten during the past 24h; 999=NA]	CH: make 2 categories and limit to 24 hours before urine collection (No=0; 12 hours OR 12-24 hours before urine collection=1; NA=999) CH and NL: Combine single food items into food categories (vegetables and fruit, rice, bread, eggs, dairy, meat)

Variables for the planned analyses were chosen by the first authors, and listed and compared in an excel file. Recoding and categorization was done in R software.

Table A.3. *Exposure variables included in the Tobit and logistic regression models and their distribution. Numbers in brackets signify the total after imputation.*

Country		Netherlands	Switzerland
Samples, n		105	295
Having a household member exposed to pesticides, n		1	3 (4)
Usage of pesticides in the garden, n		1 (1)	13
Usage of pesticides indoors, n		4 (5)	6 (7)
Usage of pesticides on pets, n		1	5
Usage of pesticides for human use, n		2 (2)	9
Consumption of homegrown foods in summer, n			
	<10%	88	274
	10-50%	11	15
	>50%	6	6
Organic consumption of vegetables, n			
	<10 %	47	13
	10-50%	41	70
	>50%	17	212
Organic consumption of bread, n			
	<10%	88	73
	10-50%	8	94
	>50%	9	128
Organic consumption of meat, n			
	<10%	52	57
	10-50%	15	67
	>50%	38	171
Organic consumption of eggs, n			
	<10%	50	32
	10-50%	5	32
	>50%	50	231
Organic consumption of dairy, n			
	<10%	62	42
	10-50%	12	58
	>50%	31	195
Organic consumption of rice, n			
	<10%	68	68
	10-50%	16	109
	>50%	21	118

Table A.4. *Compound characteristics of the 12 parent compounds included in the censored linear (a) or logistic (b) regression.*

Compound	Biological half-life [h]	Regulation (NL) [14,15]	Regulation (CH)	Application (NL) [t/year (2020)][16]	Application (CH) [t/year (2020)][17]	Top 5 food items in which compound was most frequently detected (EFSA data 2020)	Consumption per food item (NL, n= 105) [%]	Consumption per food item (CH, n=295) [%]
Acetamiprid (a)	4-15 (rat model; measured in urine)[2]	approved for occupational use; certain acetamiprid-containing products are approved for non-occupational use	approved for occupational use; certain acetamiprid-containing products are approved for non-occupational use	4.486	1.182	Cherries (47.7)	-	-
						Chili peppers (37.7)	-	-
						Pomelos (30.1)	-	-
						Roman rocket (28.8)	32.4 (leafy greens)	54.0 (leafy greens)
						Pears (25.9)	11.4	7.3
Chlorpropham (a)	3-16 (rat model; measured in urine)[3]	no longer approved for occupational use (max. period of grace 8 Oct 2020); not approved for non-occupational use	no longer approved for occupational use (max. period of grace 30 Sept 2020); not approved for non-occupational use	39.292	0.063	Potatoes, general (14.8-27.9)	38.1 (incl. fries)	28.7
						Beetroots (10.0)	-	-
						Avocados (8.0)	-	-
						Radishes (6.9)	-	-
						Cucumbers (6.3)	-	-
Flonicamid (a)	5-12 (rat model, measured in urine)[4]	approved for occupational use	approved for occupational use	4.913	0.218	Cucumbers (21.7)	-	-
						Sweet peppers (17.8)	30.4	-
						Peas (12.7)	12.3	-
						Peaches (10.3)	0.0 (incl. apricots)	8.3 (incl. apricots)
						Brussel sprouts (10.2)	-	-
Boscalid (b)	< 24 (rat model; main excretion is fecal)[5]	approved for occupational use	approved for occupational use	9.306	1.950	Blackberries (45.3)	-	-
						Head lettuce (43.9)	32.4 (leafy greens)	54.0 (leafy greens)
						Dried vine fruits (38.9)	11.4 (dried fruits)	16.3 (dried fruits)
						Lamb's lettuce (37.6)	32.4 (leafy greens)	54.0 (leafy greens)
						Blueberries (36.3)	-	-

Table A.4. *Continued.*

Compound	Biological half-life [h]	Regulation (NL) [14,15]	Regulation (CH)	Application (NL) [t/year (2020)][16]	Application (CH) [t/year (2020)][17]	Top 5 food items in which compound was most frequently detected (EFSA data 2020)	Consumption per food item (NL, n= 105) [%]	Consumption per food item (CH, n=295) [%]
Clothianidin (b)	3-4 (rat model, measured in blood) 14 (human model, measured in urine)[6,7]	No longer approved for occupational use (except in greenhouses) since end of 2018	No longer approved for occupational use (except in greenhouses) since end of 2018	NA	0.000	Chili peppers (17.5) Beetroots (15.9) Head lettuce (10.0) Tea leaves (8.9) Radishes (8.7)	- - 32.4 (leafy greens) - -	- - 54.0 (leafy greens) - -
Cyprodinil (b)	<2 (rat model, measured in tissue)[8]	approved for occupational use	approved for occupational use; certain cyprodinil-containing products are approved for non-occupational use	9.129	4.571	Redcurrants (57.3) Blackberries (52.2) Raspberries (39.2) Strawberries (38.4) Dried vine fruits (34.5)	- - - 7.6 11.4 (dried fruits)	- - - - 16.3 (dried fruits)
Fluazifop (b)	9-37 (human model; measured in blood and urine)[9]	approved for occupational use	approved for occupational use	1.123	2.162	Beans (6.9) Linseeds (5.8) Potatoes (4.5) Aubergines (4.2) Carrots (2.5)	5.7 - 38.1 12.4 (incl. zucchini) 33.3	5.7 - 28.7 21.7 (incl. zucchini) 49.0
Fludioxonil (b)	2-5 (phase I), 30-60 (phase II) (rat model; main excretion is fecal)[10]	approved for occupational use	approved for occupational use; certain fludioxonil -containing products are approved for non-occupational use	8.082	3.389	Pineapples (75.9) Redcurrants (62.8) Blackberries (49.6) Nectarines (45.7) Strawberries (42.0)	6.7 - - 0 7.6	2.0 - - - -

Compound	Biological half-life [h]	Regulation (NL) [14,15]	Regulation (CH)	Application (NL) [t/year (2020)][16]	Application (CH) [t/year (2020)][17]	Top 5 food items in which compound was most frequently detected (EFSA data 2020)	Consumption per food item (NL, n= 105) [%]	Consumption per food item (CH, n=295) [%]
Pirimiphos-methyl (b)	1 (rat model; measured in blood)[11]	No longer approved for occupational use	No longer approved for occupational use	0.000	0.000	Wheat flour (27.0)	-	-
						Beans (26.2)	5.7	5.7
						Orange juice (15.0)	4.7	13.7
						Wheat flour white (14.1)	-	-
						Beetroots (12.9)	-	-
Propamocarb (b)	NA	approved for occupational use	approved for occupational use	92.911	1.147	Cucumbers (46.6)	-	-
						Lettuces (24.6)	32.4 (leafy greens)	54.0 (leafy greens)
						Head lettuce (23.8)	32.4 (leafy greens)	54.0 (leafy greens)
						Radishes (18.9)	-	-
						Crisp lettuce (14.6)	32.4 (leafy greens)	54.0 (leafy greens)
Pyrimethanil (b)	3-15 (phase I), 8-30 phase II) (human model; measured in urine)[12]	approved for occupational use	approved for occupational use	5.562	1.178	Clementines (44.2)	-	-
						Lemons (43.1)	-	-
						Mandarines (41.1)	-	-
						Oranges (37.2)	19.0	10.0
						Dried vine fruits (33.4)	11.4 (dried fruits)	16.3 (dried fruits)
Tebuconazole (b)	8 (oral administration)-16 (dermal administration) (human model; measures in urine)[13]	approved for occupational use; certain tebuconazole -containing products are approved for non-occupational use	approved for occupational use; certain tebuconazole -containing products are approved for non-occupational use	4.982	6.061	Nectarines (39.5)	0	-
						Cherries (38.8)	-	-
						Peaches (37.0)	0	8.3 (incl. apricots)
						Dried vine fruits (31.5)	11.4	16.3 (dried fruits)
						Passionfruit (29.9)	-	-

2 National Center for Biotechnology Information (2023e); 3 National Center for Biotechnology Information (2023a); 4 National Center for Biotechnology Information (2023f); 5 National Center for Biotechnology Information (2023d); 6 National Center for Biotechnology Information (2023g); 7 Harada et al. (2016); 8 United States Environmental Protection Agency (2003); 9 Woollen et al. (1991); 10 National Center for Biotechnology Information (2023c); 11 National Center for Biotechnology Information (2023b); 12 Faniband et al. (2019); 13 Oerlemans et al. (2019); 14 Dutch Board for the Authorisation of Plant Protection Products and Biocides (Ctgb) (2023); 15 Agriculture and Environment Research Unit (AERU) (2023); 16 Dutch Central Agency for Statistics (cbs) (2022); 17 based on sales data by the Swiss Federal Agency for Agriculture (2022)

Table A.5. *Detected biomarkers through the suspect screening approach (n=37). Only compounds of confidence levels (Schymanski score) 1 and 2b are presented. Biomarkers are ordered based on their overall detection frequency. Pesticides marked in bold were included in the linear regression models, pesticides underlined were included in the logistic regression model.*

ID	Pesticide type[1]	Parent pesticide	Metabolite found in urine	ID[2] level	Detection ratio, % NL	Detection ratio, % CH
P2_a	I	**Acetamiprid**	-CH2	1	98.1	87.5
P18_b	I	**Flonicamid**	-C2HN	2b	52.4	43.7
P8_a	H, GR	**Chlorpropham**	+O +SO3 (4-HSA)	1	63.8	39.7
P35_b	F	Propamocarb	+O	2b	47.6	29.5
P34_a	I, Ac	Pirimiphos-methyl	-CH2	1	44.8	19.0
P20	F	Fludioxonil	+O +C6H8O6	2b	27.6	16.3
P5_a	F	Boscalid	+O +SO3	2b	32.4	12.5
P38_a	F	Pyrimethanil	+O +SO3	2b	31.4	11.5
P19_a	H	Fluazifop	Parent	1	20.0	14.6
P11_a	I	Clothianidin (can come from thiamethoxam)	Parent	1	21.9	12.5
P35_a	F	Propamocarb	Parent	1	18.1	13.2
P13_a	F	Cyprodinil	+O +SO3	2b	21.0	11.2
P40_a	F	Tebuconazole	-2H +2O	2b	20.0	6.8
P28_a	I	Imidacloprid	-NO2 +H	1	4.8	10.5
P11_c	I	Clothianidin (can come from thiamethoxam)	-CH2	2b	5.7	5.1
P19_b	H	Fluazifop	Parent	1	4.8	5.4
P27_a	F	Imazalil	+C6H8O6	2b	6.7	3.4
P9_b	I	Chlorpyrifos (/methyl)	-CH2	1	7.6	2.0
P46_a	F	Trifloxystrobin	-CH2 -CH2	2b	0.0	4.1
P42_a	I	Thiacloprid	+O	2b	1.9	2.4
P21_b	F	Fluopyram	+O +C6H8O6	2b	4.8	1.0
P21_c	F	Fluopyram	-2H	2b	2.9	1.4
P37	H	Propyzamide	+H2O3	2b	1.0	2.0
P18_a	I	Flonicamid	Parent	1	2.9	0.7
P43_a	I	Thiamethoxam	Parent	1	0.0	1.7
P22_a	I	Flupyradifurone	Parent	1	2.9	0.3
P3_a	F	Ametoctradin	-C2H6 +2O	1	1.0	1.0
P32_a	F	Penconazole	+O +C6H8O6	2b	1.9	0.3
P43_b	I	Thiamethoxam	-NO2 +H	1	0.0	1.0
P21_a	F	Fluopyram	+O +SO3	2b	1.9	0.0
P41_a	F	Thiabendazole	+O +C6H8O6	2b	1.0	0.3
P9_a	I	Chlorpyrifos (/methyl)	TCPy	1	0.0	0.7
P1	H	2,4-D	Parent	1	0.0	0.3
P10	H	Clopyralid	Parent	1	0.0	0.3
P11_b	I	Clothianidin (can come from thiamethoxam)	-NO2 +H	1	0.0	0.3
P25_a	I, Ac	Fluvalinate	-C14H9NO	2b	0.0	0.3
P38_b	F	Pyrimethanil	+O	2b	1.0	0.0

1. H: Herbicide, F: Fungicide, I: Insecticide, GR: Plant Growth Regulator, Ac: Acaricide, M: molluscicide, Al: Algicide, Ab: antibacterial, Af: antifungal
2. Schymanski confirmation level

For the 3 pesticides acetamiprid, chlorpropham, and flonicamid, detection ratios of the corresponding biomarkers were above 40% in both the Dutch and Swiss study samples (marked in bold). Tobit regression was performed on these three compounds. An additional 10 biomarkers (underlined) were detected in at least 40 samples (overall detection ratio >10%). These 10 were included in logistic regression models.

Table A.6. *ICCs Netherlands. Correlation between seasons and within individuals for the 12 pesticide biomarkers.*

Pesticide	Parent pesticide	Mean ICC	SD
P2_a	Acetamiprid	0.05	0.05
P5_a	Boscalid	0.15	0.11
P8_a	Chlorpropham	0.28	0.10
P11_a	Clothianidin (can come from thiamethoxam)	0.14	0.13
P13_a	Cyprodinil	0.09	0.10
P18_b	Flonicamid	0.10	0.09
P19_a	Fluazifop	0.34	0.17
P20	Fludioxonil	0.25	0.15
P34_a	Pirimiphos-methyl	0.23	0.13
P35_b	Propamocarb	0.13	0.10
P38_a	Pyrimethanil	0.06	0.07
P40_a	Tebuconazole	0.31	0.16

Chapter 7

General Discussion

During daily life, humans are exposed to a plethora of manufactured chemical compounds originating from various sources. These exposures can potentially lead to an almost infinite number of co-occurring compounds. All these possible combinations simply cannot be covered by single chemical (or: chemical family) risk assessment approaches. The assessment of chemical mixtures is one of the major current challenges in environmental epidemiology (Price et al., 2022; Taylor et al., 2016; Vermeulen et al., 2020a). Part of this challenge is how to accurately measure real-life chemical mixtures. This thesis contributes to the existing knowledge by applying several approaches to identify chemical mixtures in the general population. These approaches can be distinguished by, i) visualization technique to describe co-exposure of chemicals in mixtures, ii) measuring real-life chemical mixtures at the individual level, and iii) applying a novel analytical suspect screening (SS) approach. The first part of the thesis is concentrated on a visualization technique, applied to existing human biomonitoring (HBM) data. By visualizing co-occurring compounds in correlation networks, clusters of exposure markers (chemical mixtures) were identified. The second part describes the application of the second and third approaches to pesticide mixtures, involving the measurement of pesticide mixtures at the individual level and employing a SS approach. Together, these three approaches aim to demonstrate the feasibility of assessing real-life chemical mixtures at the individual level in the general population.

The key messages of this thesis are:

- Graphical correlation network analysis enables the visualization of mixture patterns in HBM data, facilitating the identification of clusters of closely related biomarkers within these networks. Additionally, the role of factors influencing the clusters can be explored through covariate analysis.
- Personal exposure assessment, using HBM and personal samplers, such as silicone wristbands, effectively identifies co-occurrence patterns of pesticide mixtures at the individual level.
- A harmonized SS approach provides the opportunity to efficiently detect a wide range of pesticides within a pan-European HBM study.

In this chapter, the results of the three approaches to identify chemical mixtures are discussed, along with their applicability, limitations and alternative methods. Lessons learned from the research conducted in this thesis are presented, followed by a discussion on the interpretation and relevance of chemical mixtures, implications for mixture risk assessment, and future research opportunities. The author's perspective on the next steps in this challenging field will also be provided.

Visualization of chemical mixtures - *Main findings*

When describing exposure to chemical mixtures, one of the challenges lies in identifying co-occurrence patterns among the different chemicals. In **chapter 2** we presented a possible solution to address this issue: a correlation network approach that visualizes the chemical mixtures present in the data. This approach was applied on HBM samples obtained during three measurement campaigns conducted in Flanders, Belgium. Among these campaigns, the one with the most comprehensive data included measurements of 19 organic pollutants

and metals in cord blood samples collected from 281 children at birth (*FLEHS III*). Graphical correlation networks were used to describe dependence between multiple chemicals, providing a data-driven and intuitive way to identify correlated biomarkers. The identification of groups of highly connected biomarkers, "communities", within these networks highlighted which biomarkers should be considered collectively in the analysis and interpretation of epidemiological studies or the design of toxicological mixture studies. In **chapter 2,** we observed how biomarker networks and its communities changed across the three sampling campaigns, between smoking status during pregnancy, and between high and low maternal pre-pregnancy BMI. As expected, *FLEHS III* revealed communities consisting of similar chemical structures such as PCBs and PFASs. Additionally, a community emerged involving the metals Mn, Cu and Cd. Notably, p,p'-DDE (a metabolite of DDT) clustered together with the PCB community, both persistent organochlorines, highlighting the importance of considering them jointly when assessing potential risks (Longnecker et al., 1997; van den Berg et al., 2017). Comparisons of networks across covariates, such as stratification by smoking status or BMI, showed minor differences in network and community structure.

In **chapter 3,** the same network approach was applied to four HBM datasets from Belgium, Czech Republic, Germany and Spain. This chapter also addressed some methodological aspects such as a comparison of different approaches to correct for creatinine in urine samples and a comparison of unweighted and weighted networks. From the Belgium *3xG* study 125 mother-child pairs were included, with measurements of 36 different biomarkers in maternal blood, cord blood, cord blood plasma, or morning urine. The resulting network revealed nine communities, two of which consisted of chemicals from different families, while the metals formed three separate communities. The seven within chemical family communities mostly reflected samples from the same biological medium. From the Czech *FIREexpo* study 52 firefighters were included, in which 16 different biomarkers were measured in serum and urine. Four communities were identified in the resulting network, and one of them demonstrated a cross chemical family link between PFBS and 4-PHEN. From the German *GerES V* cohort, 515 participants aged 3 to 17 years were selected, with measurements of 51 different biomarkers in first morning void urine. In the resulting network, we identified ten communities, two of which did not exclusively represent the same chemical family. Lastly, from the Spanish *BIOAMBIENT* study 163 participants were selected with 31 different biomarkers measured in morning void urine or blood. The resulting network comprised six communities, of which half did have links across chemical families. An example was the connection between lead and DiBP (phthalate) metabolites. Overall, this was the first, to our knowledge, application of a harmonized and standardized (shared script) network approach on multiple datasets from different countries, incorporating a substantial number of biomarkers and individuals. Utilizing existing HBM datasets allowed us to conduct an initial exploration of co-occurrence of substances within the human body. The application of community detection was useful in identifying patterns within and between chemical families. It is important to note that the overlap between biomarkers measured in the different studies was limited, and networks were only applied on non-harmonized data (exposure markers and covariates) as each study was analyzed separately due to privacy regulations. Consequently, the observed differences

between study, such as the variations in phthalate communities between Germany and Spain (split into seven and four communities respectively), may be attributed to differences in study populations, design, chemical analytical procedures and actual exposure differences.

In **chapter 2 and 3** the communities identified within the networks provided an initial insight into the patterns of chemical mixtures present in the data. It is plausible that these communities reflect shared sources, routes of exposure, or physiochemical properties among the biomarkers. For instance, both PCBs and DDT can be ingested through diet. In **chapter 2,** the comparative network analysis conducted by stratifying the data based on covariates such as education level, age, BMI did show promise, with changing community structures for some of the stratifications. This comparative approach, holds potential for future investigations to identify specific sets of covariates that influences the community structures. Such findings could then be translated to action perspectives, such as making adaptations in diet or other lifestyle factors based on the identified differences.

In **chapter 3** HBM samples from different media were included, both blood and urine samples. Based on the Belgium data, we could see that metals detected in the urine of the mother were highly related to metals in the cord blood of their baby, which is in line with previous studies on migration of metals through the placenta to the fetus (Rísová, 2019; Vahter, 2008). For chemicals measured in urine, a standardization for creatine content was performed to account for the level of dilution and to allow for a better comparison with the biomarkers measured in blood. Comparing creatinine adjusted and non-adjusted networks for the German data, the non-adjusted network resulted in more spurious results (more false-positive edges), indicating the need to harmonize the creatinine adjustments across studies. Different creatinine adjustment procedures yielded similar results.

In **chapters 2 and 3,** networks were presented with both unweighted and weighted edges. The inclusion of weighted edges provided a more detailed understanding of the level of dependency (thickness of the edge) and the direction (positive or negative, by colors green or red) between biomarkers. While the same regularization method was used for both network types, some differences were observed. The weighted approach tended to capture more connections between biomarkers, as indicated by the presence of thin lines Moreover, the detected communities differed even more between both network types, as the clustering algorithm was applied to the final optimized network (first the network was estimated, then the communities were derived). It is important to acknowledge that the weighted approach enhanced interpretability for the datasets presented in **chapters 2 and 3.** In (future) high-dimensional HBM datasets, for example generated with SS, weighted networks are likely too computationally intensive and visually challenging to interpret. In such cases, unweighted networks would be preferable.

Compare our findings with existing literature is a challenge, as the application of correlation networks on datasets based on biomarkers of chemical exposure is still rather novel. In other scientific fields, network approaches are more commonly used, for example in cell biology (Schwikowski et al., 2000), metabolomics (Gauglitz et al., 2020) and psychology (Epskamp et al., 2018). An exception on HBM data is based on six birth cohort studies in Europe (the HELIX project), where network analysis was performed based on the within-cohort cor-

relation matrix (Tamayo-Uria et al., 2019). The actual correlations were used to determine the network structure and no regularization (as used in **chapters 2 and 3**) was applied. These networks of Tamayo-Uria et al. detected clusters of e.g. outdoor variables and PFASs with organochlorine compounds (Tamayo-Uria et al., 2019), while certain metals were not connected to other exposures, the latter being similar to the findings in **chapters 2 and 3**.

Visualization of chemical mixtures - *Applicability and limitations*

With the application of network analysis, I aimed to propose an easy to implement method to visualize prevalent chemical mixtures in HBM data. Visualizing chemical mixtures helps to better understand which chemicals are more closely related to each other, which can be explained by a commonality in source, exposure route or physiochemical properties in the human body. Detecting chemicals with a commonality, for example in source, could assist in performing more accurate risk assessments on those compounds that occur together in real-life.

Current methods for visualizing chemical mixtures in population-based HBM data include heatmaps, circos plots, and dendrograms. Heatmaps and circos plots display the correlation or partial correlation structure of compounds, with compounds ordered on the axes by the researcher's choice. The choice of axis order has a major impact on the ability to visually distinguish groups of more correlated chemicals. In circos plots, a manual threshold can be set to show only the most prominent links between chemicals, improving interpretability. However, this threshold must be manually determined and impacts the interpretation largely. Adding a dendrogram to heatmaps or circos plots helps to assess similarity between chemicals. The most commonly used methods to determine the optimal number of clusters are k-means or agglomerative clustering, which require prior decisions on the ideal number of clusters. The non-supervised network approach presented in **chapters 2 and 3** addresses these limitations by plotting chemicals in a network based on their dependency to others. By application of data-driven variable selection and a clustering algorithm the optimal number of clusters can be identified. In the data structure exploration phase, networks can complement heatmaps and circos plots. Networks are also valuable for detecting changes between datasets or within subsets of a dataset, such as different time points, smoking status or the use of specific products. The comparisons can potentially direct to previously unexplored links between chemical combinations (in certain subpopulations).

While this thesis has demonstrated the value of correlation networks in the field of HBM, it is important to note that networks are not a quick fix and there are some critical considerations and potential areas for improvement based on recent developments. The construction of networks relies on certain assumptions about the input data. Similar to other non-supervised statistical methods, the quality and construction of the input data impacts the correlation network structure. Presence of clear outliers, measurement errors, poor normalization, and batch effects can influence the network results (Santos et al., 2020). It is worth noting that networks become less unreliable or unstable with small sample sizes (<200) and a limited number of variables (<20, only when based on the partial correlation) (Epskamp et al., 2018). Furthermore, the dependency between two chemicals within a network is impacted by the

presence of other compounds included in the dataset, making it challenging to compare networks across studies, especially when different sets of chemicals are included.

During network estimation inducing sparsity is crucial to improve interpretability. Different methods to estimate sparsity can be chosen, such as the regularized nodewise regression of Meinshausen and Buhlmann (Meinshausen et al., 2006), or the graphical lasso (Glasso) by Friedman et al. (Friedman et al., 2008). In **chapters 2 and 3** the Glasso method was used, which simultaneously performs variable selection and estimation, resulting in a sparse graphical network. It should be noted that there is no one-size-fits-all approach, and choice of network estimation highly depends on the specific characteristics and structure of the data. The estimation of the true network can be challenging and there is a risk of both false-positive and false-negative edges. A proposed suggestion to control for false-positives as by Lafit et al. (Lafit et al., 2019), where correlations below a given threshold are set to zero, could be considered for future purposes, specifically when dealing with high-dimensional data. Also, recent advances in stability selection (based on resampling) as presented by Petrovic et al. (Petrovic et al., 2022) could be considered.

Exciting new developments in network modelling have taken place in the past couple years, offering potential improvements and alternative approaches for future research. Here, I discuss some of these considerations. First, the clustering algorithm *walktrap* used in this thesis detects biomarker sets with stronger connections within the set compared to outside (Fortunato, 2010; Pons et al., 2005). However, in the context of HBM mixtures, sometimes overlapping clusters may be more realistic. For instance when the research question relates to identifying different sources of exposure, such as pesticide exposure through both consumer use and diet. In such cases, fuzzy or overlapping clusters (such as the Clique Perculation Method (Palla et al., 2005)) would be more suitable to address this research question. In this thesis only non-overlapping clustering was applied, as the aim was to describe the co-occurrence, not related to source-identification. Second, the comparative network analysis in **chapters 2 and 3** was based on the presence or absence of edges. Alternative approaches, such as cluster-based comparisons or dynamic clustering (how clusters evolve over time periods) could be promising to address HBM mixture research questions (Palla et al., 2007). Third, it should be mentioned that many other methods are available to describe mixtures in HBM data, many complementary to network models. These methods can be categorized based on a description of the correlation structure (among which networks), of the dimensionality and of the variability (Santos et al., 2020). Unsupervised dimension reduction techniques, such as principal component analysis, factor analysis and sparse non-negative matrix underapproximation capture the variability of the dataset in a manageable number of variables (Gillis et al., 2013; Kalia et al., 2020). Furthermore, a recent study proposed a combination of networks with dimension reduction called graphical lasso-guided iterative principal component analysis (Harakawa et al., 2022), which removes trends with indirect correlations generated by other essential trends. This proposed methodology is especially interesting when looking into time trends of chemical mixtures.

Measuring pesticide mixtures at individual level - *Main findings*

When addressing exposure to chemical mixtures, individual measures are necessary to determine the frequency and magnitude of chemical mixtures, these measurements can either be external (active or passive) or internal. In large population-based studies passive sampling is often preferred, mostly due to practical constraints. Personal passive samplers, such as silicone wristbands as described in **chapter 4**, can effectively assess personal exposure to chemical mixtures over extended periods. **Chapter 4** focused on pesticide exposure among 19 residents living close to flower bulb fields in the Netherlands, as add-on to the existing pesticide exposure study OBO (Figueiredo et al., 2021). In the wristbands in **chapter 4**, out of 46 measured pesticides, 31 were detected. The overlap in detected pesticides between individuals was minimal, with only three pesticides detected in all wristbands (azoxystrobin, pymetrozine, carbendazim). Similar individual exposure profiles were detected in other wristband exposure studies (Dixon et al., 2019; Fuhrimann et al., 2022). Of the 31 measured pesticides, eight were not registered for use on flower bulb fields, and two were registered for use on flower bulb fields but were not recorded as being used during the study. Vacuumed floor dust samples were collected from participants' homes before and after wearing the wristbands, but no clear patterns emerged in the comparisons between the two measurement types.

In addition to external measures, internal measurements are informative in collecting the individual-level mixture exposure. Internal measures, also known as HBM, can be very time and labor intensive, invasive, and have typically higher costs, specifically when compared to passive external measures. Despite these challenges, we were able to collect HBM samples to measure pesticide mixtures for 1,345 participants in six countries across Europe as part of the European initiative HBM4EU. This study, SPECIMEn (Survey of Pesticide Mixtures in Europe), focused on the internal exposure assessment by pesticide biomarkers in urine, aimed to provide a comprehensive and integrated picture of human exposure to pesticide mixtures across various sources and pathways. In **chapters 5 and 6** urinary first morning void samples were collected from adults and children, in agricultural (< 250m from agricultural land) and non-agricultural (> 500m) residential areas, in six countries across Europe (Netherlands, Switzerland, Hungary, Czech Republic, Spain, Latvia), during two seasons. The hotspot-seasonal based design aimed to detect different levels of pesticide exposure related to e.g. spraying activities of a neighboring field during growing season. The application of a SS approach resulted in a qualitative assessment of a large number of pesticides detected across all different countries (**chapter 5**). Contrary to the hypothesis, the number of detects was not consistently higher among hotspot locations, and during spraying season (summer/autumn depending on the country). The number of detected pesticides was higher in samples from children compared to adults, likely due to their higher dietary pesticide intake per kg body mass. Analysis of co-occurrence showed that in 84% of the samples at least two different pesticides were detected, with a maximum of 13 pesticides detected in a single urine sample. Frequently detected specific combinations of pesticides consisted of acetamiprid-chlorpropham (62 samples) and acetamiprid-tebuconazole (30 samples). Acetamiprid with pirimiphos-methyl was the only combination found in all five study sites. Also, assessment of the co-occurrence patterns at country level (network analysis), did not result in strong relations and hardly any overlap across countries was seen.

Chapter 6 builds upon the data from the Netherlands and Switzerland, measured at the same laboratory, aiming to identify sources of exposure through construction of censored linear models. We found that the most frequently detected pesticides, namely acetamiprid and chlorpropham, were reduced in individuals with a high consumption of organic fruits and vegetables, which aligns to previous studies on these compounds (Hyland et al., 2019; Rempelos et al., 2022; Ye et al., 2015; Zhang et al., 2022). The indication of dietary exposure as the main source could explain the lack of differences observed in location and season in **chapter 5**. However, to precisely determine additional exposure through inhalation of outdoor air or contaminated house dust, a more in-depth analysis focused on pesticide-specific factors would be required.

The two applied measurement approaches (wristbands and HBM) were applied in different study designs, locations and time periods. However, when focusing on data from the Netherlands, there was some overlap in detected pesticides. A total of 14 pesticides were detected with both methods, of which 11 were confirmed with high levels of confidence in the urine samples. Chlorpropham was the most frequently detected pesticide, being found in 95% of the wristbands and 75% of the urine samples. The frequent detection of chlorpropham in the wristbands indicates not only dietary exposure but also exposure through inhalation. Acetamiprid was frequently detected in urine (93%) but in just 26% of the wristbands, likely indicating dietary exposure as the primary route. Boscalid and fludioxonil were detected in over 90% of the wristbands, but in just 33% and 27% of the urine samples, respectively. The short half-lives in soil of boscalid and fludioxonil (less than 2 weeks), may indicate infrequent exposure or exposure that did not occur around the time of urine collection for most participants. Although **chapters 4 and 5** had a different focus in study design (flower bulb fields versus apple and pear orchards) and data were collected in different years, the overlap in detected pesticides suggests airborne or direct physical contact exposure to those compounds in addition to dietary exposure.

Measuring pesticide mixtures at individual level - *Applicability and limitations*

With the application of personal measures as silicone wristbands and HBM the mixture of exposure at individual level was assessed, reflecting individual differences for example due to lifestyle. However, to fully capture the variability of mixtures present in the population, larger sample sizes and advanced modelling would be required. Moreover, to assess the determinants of individual pesticide mixture exposure, more detailed information on dietary intake and pesticide levels in and around the residences will be needed.

Wristbands are a relatively new (first described in 2014) measurement technique, and validation with HBM samples and other external exposure samplers is still ongoing (Hou et al., 2021; Samon et al., 2022; Wacławik et al., 2022a). These passive and low-impact samplers allow for longer timeframes of measurement and can detect a wide array of compounds using targeted or non-targeted high-resolution approaches (Bergmann et al., 2018; Travis et al., 2021). They provide insights into the chemical mixture to which an individual is exposed over time through air or direct physical contact (Samon et al., 2022; Wacławik et al., 2022b). Wristbands are easy to implement, have low burden, and are cost-effective. It's important to note that

wristbands represent time-weighted averages rather than episodic concentrations (Samon et al., 2022). In **chapter 4**, wristbands were worn for an average of 60 days, which is longer than the usual one to seven-day period (Hamzai et al., 2022a). This extended wearing period increases the potential of reaching equilibrium for compounds to which the wearer is repeatedly exposed (O'Connell et al., 2014; Samon et al., 2022). However, less frequent exposures may not remain stable in the wristbands for the entire 60-day duration (up to 7 days for some VOCs, and up to 30 days for SVOCs). Wristbands are able of sampling dermal exposure (both vapor phase compounds and compounds on the skin), as well as airborne exposure (capturing the portion of VOCs and SVOCs that can be inhaled, although particle exposure capture is still uncertain) (O'Connell et al., 2014; Samon et al., 2022). An alternative to silicone wristbands is for example Fresh Air wristbands, which focus on capturing NO_2, VOCs, and PAHs. Incorporated in the Fresh Air bands are an triethanolamine-coated pad besides a polydimethylsiloxane sorbent bar, together able to capture non-direct exposure from ambient air. These could be considered in research settings focused on for example longitudinal air pollutant exposure (Lin et al., 2020).

HBM is commonly applied to measure aggregated exposure, reflecting internal concentrations originating from different routes of exposure. In the case of pesticide mixtures, HBM reflects ingestion, dermal exposure and airborne exposure. Insights into the actual internal concentrations of chemicals is crucial for accurate risk assessment. The application of SS (see next paragraph) to HBM samples allows for the measurement of larger sets of pesticides in a single sample, although efforts to quantify these compounds still need further development. It should be noted that HBM samples are highly impacted by inter- and intra-individual variability which challenge the interpretation of differences across population, this variability can be caused by for example metabolism rate, toxicokinetics, timing of exposure, timing of sampling, and body mass index (Aylward et al., 2014a; Koch et al., 2014). Additionally, urinary flow is known to be variable and influenced by many short-term (e.g. hydration status) and long-term parameters such as age and BMI (Aylward et al., 2014b; Wacławik et al., 2022a). These aspects often result in a large day-to-day variation, particularly for short-lived compounds. Repeated samples would be preferred, although this is not always possible due to practical limitations. A possible solution would be to analyze pooled samples, for example from an entire week (Perrier et al., 2016; Vernet et al., 2018). Although from a mixture perspective, the analysis of individual samples would be preferred to not underestimate (for example by dilution when pooling) the mixture of exposure.

Silicone wristbands and HBM samples offer complementary advantages and disadvantages. Wristbands pose a lower participant burden and can be employed over longer periods, but lack health-based guidance values. HBM samples reflect internal concentrations and provide an accurate assessment of multiple compounds at a single timepoint, allowing for comparison with existing data.

The comparison of wristbands with HBM samples is limited by differences in exposure route reflection (wristbands only capture dermal and inhalation exposure). Correlation between wristbands and urine samples may be impacted by factors as timing of urine sample, use of pooled or spot urine samples, correction of urinary metabolites for creatinine, timing

of wristband application, and whether the chemical of concern reached equilibrium in the wristband (Samon et al., 2022). As a result, variable correlations between the two methods have been reported in the literature (Hamzai et al., 2022b; Samon et al., 2022; Wacławik et al., 2022a). It can be hypothesized that the correlation between wristbands and HBM samples would be highest in occupational settings where the ingestion route is minimal compared to dermal and inhalation exposure routes, this hypothesis was confirmed for phthalate esters exposure among nail salon workers (Craig et al., 2019).

Application of an analytical screening approach on pesticide mixtures - *Main findings*

When analyzing exposure to multiple chemicals in HBM samples, measurements are limited by known measurable exposure biomarkers with established (targeted) methods, the sample volume, and costs of analyses. Besides, these targeted measures need to be standardized to provide a reliable and reproducible quantitative exposure estimate. The characterization of the total real-life mixture exposure would require an increase of detection of biomarkers measured in a single sample in an efficient and effective manner. SS is based on liquid chromatography coupled with high-resolution mass spectrometry (LC-HRMS) and aims to generate a list of semi-quantitative annotations present in a sample set (Huber et al., 2022; Pourchet et al., 2020). SS enables detection of a large set of chemicals in a single sample, which is of great value for mixture exposure assessments. SS approaches have only been applied in a handful of HBM studies before, for example in the US by Pellizzari et al. (Pellizzari et al., 2019) and Wang et al. (Wang et al., 2018), gaining insight in chemicals not covered by targeted approaches. In **chapters 5 and 6** a pesticide SS approach was conducted on 2,383 collected urine samples from 1,345 participants across Europe (SPECIMEn study), performed in five laboratories. The main challenge for this SS approach is the harmonization across all laboratories of the i) sample preparation, ii) QA/QC provisions and criteria, and iii) data processing (Huber et al., 2022; Vitale et al., 2022). Novel in this study was the aggregation of previously curated proprietary suspect lists of pesticides and their metabolites among the five laboratories, which was made public with the publication (Huber et al., 2022). In total 498 tentative annotations were achieved (11% of the aggregated suspect list) (Huber et al., 2022), of which 40 pesticide biomarkers were confirmed with high levels of confidence (Schymanski scores 1 and 2b) and 54 with lower confidence levels (Schymanski levels 3-5) (Schymanski et al., 2014). The 40 high-confidence level biomarkers were used for analysis in **chapters 5 and 6**. These 40 related to a total of 29 parent pesticides and had generally low detection rates (<25% of the samples). The most frequently detected parent pesticides at all study sites were acetamiprid and chlorpropham.

Application of an analytical screening approach on pesticide mixtures - *Applicability and limitations*

The application of the SS approach on the HBM samples as described in **chapters 5 and 6,** provided a total of 40 detected pesticide biomarkers. This number is significantly larger compared to targeted studies, where typically a handful of biomarkers is measured. The advantages of SS over targeted methods are among others, a smaller sample volume needed (mg-mL range instead of g-mL range), quick simple non-selective extraction, and no or very limited purifica-

tion needed; this preserves sample integrity, limits sample preparation related variability and facilitates interlaboratory harmonization (Pourchet et al., 2020). Other advantages of SS are the higher number of compounds detected and the high-throughput capability, although at the cost of a lack of quantification, higher limits of detection, susceptibility to matrix effects and limited method repeatability.

In the application of SS on the SPECIMEn study (**chapter 5**), only biomarker identities (qualitative) and their detection rates could be compared across countries; across samples measured within the same laboratory, signal intensity could be used as a semi-quantitative measure as the concentration is proportional to the peak intensity (**chapter 6**). The five laboratories involved used harmonized methods in which the hardware was relatively comparable, but differences in sensitivity did occur (Huber et al., 2022; Vitale et al., 2022). Quantitative signal intensities (for only the 'unambiguously' identified biomarkers) were not standardized and harmonized across the laboratories, and could only be looked at from a within-laboratory perspective. Future efforts are needed to standardize between the laboratories. The lack of quantitative exposure data is a major restraint compared to targeted exposure studies, where patterns in exposure markers could be compared quantitatively. The data produced in the SPECIMEn study did provide a first glimpse of the potential overall exposure and can guide the prioritization of pesticide (metabolites) to be included as biomarkers in targeted methods.

Lessons learnt from the real-life example: pesticide mixtures

Based on the pesticide mixture characterization in the second part of my thesis, some lessons learnt specifically related to pesticide mixtures could be drawn.

First, setting up a harmonized pan-European pesticide mixture exposure study in six countries was challenging. Finding one overlapping crop in all countries was nearly impossible among others due to differences in climate. Additionally, spraying and non-spraying seasons for the selected set of crops (orchards, olive trees, citrus trees) varied between countries, particularly between Northern and Southern Europe. Moreover, active spraying periods varied per pesticide, and the spraying season of some pesticides, applied during e.g. planting instead of growing, could have been missed with the current study design. The use of SS eliminated the need for prior selection of pesticides, allowing for a harmonized exposure assessment across all countries without relying on specific sets of pesticides. At the time of collection, the SPECIMEn dataset was unique by its harmonized sampling collection and handling protocol, the harmonized laboratory analysis, and its scale, a total of 2,383 urine samples across Europe. Although a current European collaborative effort (the SPRINT project), within which large numbers of external and internal samples are collected, is exceeding this scale.

Second, out of the 40 pesticide biomarkers detected with high confidence, only a small number were detected within a substantial number of individuals. While it is known that SS is generally less sensitive compared to targeted approaches (Bonvallot et al., 2021; Pourchet et al., 2020), this likely only partially explains the observed low detection rates. Other factors that could contribute to low detection rates include limitations in the study design (seasonal sampling with only one sample per season), or simply low pesticide exposure in the study populations. A previous study by Bonvallot et al. showed less sensitivity of SS specifically for

highly polar low molecular-weight metabolites compared to targeted approaches; overall a target LOD below 1 µg/L was found and an estimated SS LOD of 200 µg/L (Bonvallot et al., 2021). When comparing the SS data with targeted pyrethroid measures in a subset of SPECIMEn participants (Dutch children, Tarazona et al., 2022), the metabolite DCCA (associated with cypermethrin, cyfluthrin, permethrin, transfluthrin) was the only overlapping pesticide. DCCA was not detected using SS in Dutch children, while it was found in 197 out of 207 samples using the targeted approach.

Third, the co-occurrence of pesticides detected by SS in the SPECIMEn study within individual urine samples was low, with most samples containing a small number of co-occurring parent pesticides (ranging from two to five). The intraclass correlation coefficients, a measure of within- and between-person variability, based on the Dutch data, were low (< 0.3) for the 12 most frequently detected pesticides (**chapter 6**). In 84% of the samples at least two different parent pesticides were detected, indicating the presence of combined exposures. The most frequent combination (acetamiprid with chlorpropham) occurred in only 3% of the total samples. Additionally, network analyses conducted at country level showed minimal overlap between countries. These findings suggest that pesticide mixtures are highly individual-specific, location-specific and time-specific. However, the used analytical approach (SS) and study design might have underestimated the detected pesticide mixture.

Fourth, in **chapter 6** we explored potential sources of pesticide exposure based on urine samples from the Netherlands and Switzerland. Considering the overall results of the SPECIMEn study, no seasonal differences were observed and only a small number of biomarkers was detected at high rates. For the two most frequently detected pesticides (acetamiprid and chlorpropham metabolites) it was found that high consumption of organic fruit and vegetables reduced the levels of exposure. Due to limitations in collected exposure source data and the low detection rates of biomarkers, further exploration of this finding was not possible.

Interpretation and relevance of mixtures

Mixture exposure is gaining increasing attention on scientific, societal and political agendas, due to a growing awareness of complex interrelationships between compounds. Mainly, there are growing concerns about the health effects of exposure to multiple compounds, even at low doses, as the additive dose of all compounds together could potentially result in adverse effects. Policy-makers are particularly interested in addressing the regulatory challenges associated with assessing risks linked to exposure to mixtures (Savitz. et al., 2023). The scientific interest in mixtures lies in a more accurate description of total exposure and their associated risk assessments (Lesliam et al., 2023). The research presented in this thesis also adds to the expanding field of exposome science, which encompasses all environmental exposures, including diet, consumer products and lifestyle factors (Vermeulen, 2022; Vermeulen et al., 2020b; Wild, 2005, 2012).

Exposure assessment of real-life mixtures

Based on current practice, a chemical mixture is defined by the specific focus and width of exposure measurements. Different measurement approaches may yield varying representa-

tions of the mixture. To accurately describe a chemical mixture, it is preferable to measure a wide range of chemicals uniformly. Besides, the sharing of individual-level data is preferred for detecting mixture patterns in the datasets. Currently, the overlap in HBM datasets across Europe is limited (both in terms of measured biomarkers, as in study characteristics), and factors such as privacy regulations hinder easy data sharing. The HBM4EU initiative aimed to establish a harmonized HBM database, but due to practical constraints this was limited to summary statistics of a preselected set of priority chemicals (Ougier et al., 2021). Moving forward, establishing a clear agreement on data collection and sharing at the European level would be highly beneficial. In terms of data collection, adopting a mixture perspective that encompasses multiple chemical families would be preferable. This agreement should outline the minimal level of information to be shared regarding the performed chemical analysis and co-occurrence of chemicals in the data. Publishing at least a correlation matrix, including potential variations in subpopulations, would advance the field of mixture exposure assessment.

Besides harmonized collection and sharing of exposure data, the use of SS approaches can efficiently expand the coverage of biomarkers per sample, accommodating the increasing number of chemicals on the market, including rapidly produced homologous compounds. A current major drawback of SS is the lack of quantitative exposure concentrations for detected biomarkers between laboratories, which hampers their integration in risk assessment efforts. Currently, quantitative interpretations of SS data mainly rely on relative comparisons, assuming linear relationships between instrument response and concentration (McCord et al., 2022a). Recently, advances in quantitative non-target analysis have shown promise in supporting risk-based decisions, such as quantification by a surrogate standard or response modeling based on chemical structure (McCord et al., 2022b). Efforts to standardize and quantify the SPECIMEn data will be the next step forward. To accurately assess the mixture risks to human health, insights in co-occurring compounds only will not be sufficient and actual exposure concentrations plus their level of toxicity should be considered as well. For example, if a strong correlation is observed between two biomarkers in a network, but their individual toxicity estimates are low, their contribution to the overall mixture risk would be minimal.

The timing of sampling largely impacts the detected mixture, especially for short-lived and rapidly metabolized compounds. In this thesis, the focus was on the combination of compounds measured in a single sample. For the HBM samples this also meant exposure characterization at a single point in time, possibly overlooking exposures of short-lived compounds with non-frequent exposures. Short-lived compounds can only be detected in the short to medium term (few days), e.g. after dietary consumption, or pesticide application. On the other hand, some long-lived compounds can accumulate in the body and be detected even after a drop in exposure. Understanding the differences between protracted and sequential exposures and their behavior in the overall mixture requires more knowledge on the stability of biomarkers. The use of different complementary measurement methods, such as HBM with silicone wristbands, can assist in detecting (pesticide) mixtures over a certain period of time without a substantial increase in participant burden.

Furthermore, it is worth noting that modelling exposure scenarios can help address the current gaps in available monitoring data and improve the focus of new monitoring campaigns in

terms of time and space. Examples of these exposure models are spatial models, dietary exposure models, and consumer use models.

Mixture Risk Assessment

To interpret chemical mixtures on their potential adverse effects to human health, various efforts have been published. The toxicological effect of mixtures is a debated topic in literature and of great interest to policy makers. Mixture risk assessment is specifically not the focus of this thesis, in fact it could result to another complementary thesis. Nevertheless, I will present here some MRA approaches that I think are relevant in the scope of this work.

Once a mixture of exposure has been characterized, the question arises what the functional or relevant real-life mixture is (Vermeulen et al., 2020b). A relevant mixture can include compounds that act together and share similar chemical structures, indicating a similar mode of action. In the context of HBM data, a common approach for risk assessment is to compare biomarker levels with existing health-based guidance or limit values (when available), resulting in a hazard quotient. For mixtures, the hazard quotients are summed to derive a composite hazard index.

The study by Loh et al. applied a toxicity weighing method on the communities identified in the GerES networks, as presented in **chapter 3** (Loh et al., 2023)**.** For this a biomonitoring hazard index was developed by summing hazard quotients, where each biomarker concentration was weighted (divided) by its associated HBM health-based guidance value (HBM-HBGV). However, this approach was limited by the availability of HBGVs for only 17 out of 51 substances, which greatly restricts its applicability. Also, this approach did not consider commonality in mode of action or health outcomes, leading to a relatively conservative assessment by simply summing risks per community. Furthermore, as the complexity of mixtures increase, for example with higher dimensional data, the HBGV approach becomes less applicable due to the lack of HBGVs and the fact that more compounds will always result in a higher sum and therefore higher risk.

Other available risk assessment approaches are taking a shared mechanism or target organ as starting point, however most are also limited by a lack of available guidance values. A recent example is the assessment of PFAS mixtures by Bil et al. which employed the Relative potency factor (RPF) approach, the Hazard Index approach (HI) and the sum value approach of EFSA (Bil et al., 2023). The HI approach resulted in the highest risk estimates (most conservative), followed by the RPF approach that considered differential potencies, and the sum value approach based on a limited set of PFAS with available HBGVs. Another approach suggested grouping based on toxic effects in six target organs or systems (CAGs), also allowing a tiered risk assessment (Boberg et al., 2019, 2021).

In the context of legislative settings, the European Regulation REACH is considering the application of generic factors (Mustafa et al., 2023). For instance, the Mixture Allocation Factor (MAF) of 5 is proposed by ECHA to account for potential mixture risks for chemicals produced above a certain tonnage limit. This generic factor assumes higher and more widespread exposure, resulting in a potentially higher risk for the entire population, although it lacks data-driven, evidence-based specificity. An alternative could be the use of a data-driven mixture

driver factor (MDF) for each chemical class/family separately, within which chemicals that are the main driver of the mixture effect for a specific adverse effect are assigned with an additional factor (Mustafa et al., 2023). It should be noted that the use of generic factors like MAF or MDF may be overly conservative, as not all compounds exhibit commonalities in mode of action or adverse outcome pathways, and may have different co-occurring exposure patterns.

Future research opportunities and author perspective

Exposure characterization of real-life chemical mixtures remains a scientific and regulatory challenge. Despite the growing attention for chemical mixtures in literature, integrating the results of mixture related studies into legislation is challenging. One of the challenges is that legislation per chemical class is regulated in separate regulatory silos, whereas chemical mixtures in real-life exceed the borders of these silos.

Based on the work presented in my thesis, I will highlight several opportunities in this challenging but interesting field of expertise. Including correlation networks in HBM studies would be a good addition to improve understanding of biomarker interdependencies. Networks are especially valuable for high-dimensional datasets, for example generated by SS. Networks can assist in the detection of clusters and how these change across covariates, providing direction for future actions for example related to source-identification. An extension of the methods with temporal network approaches, where nodes and links disappear and (re)emerge depending on factors e.g. related to variability in exposures across a life-course, would be valuable in the interpretation of differences over time (Vermeulen et al., 2020b). The application of dynamic clustering (changing clusters, e.g. also over time), and fuzzy clustering (overlapping clusters) can be informative for specific research questions. The SPECIMEn field survey provided valuable new insights into pesticide mixture patterns in Europe. An improved survey design would involve repeated urine samples, possibly in a subpopulation due to practical limitations, alongside silicone wristbands to capture internal concentrations and long-term exposure. High-resolution screening methods should be integrated in HBM studies to build a comprehensive database. Further development of the SS approach with quantitative analysis (McCord et al., 2022b) would benefit exposure science, epidemiology, mixture risk assessment, and legislative efforts. Lastly, to enhance the collection and sharing of uniform mixture exposure data, clear agreements at European level would be necessary regarding the minimal level of information to be shared when publishing HBM datasets which could contribute to mixture risk characterization. This minimal level of information would for example include a detailed description of the analytical method used and an overview of the observed correlations within the dataset.

In my thesis mostly positive and low to moderate correlations were observed between chemical families. For each of the detected clusters a hazard characterization would be necessary to conduct a chemical mixture risk assessment. The approach as performed by Loh et al. 2023 based on the results of **chapter 3**, demonstrated that the HI for certain clusters exceeded one, despite the fact that only a limited number of compounds was included. With an increasing number of chemicals in real-life exposure scenarios, the likelihood of higher risks becomes even more significant. Co-occurring chemicals that share a common mode of action

should be collectively considered in risk assessments. Based on the SPECIMEn study, three or more co-occurring pesticides were not frequently detected. These findings indicate that the mixture profile is likely different between individuals and a high-level of granularity in data collection is necessary for mixture risk assessment purposes. Also, it is important to note that for specific highly exposed groups or subpopulations, such as workers, the correlation pattern may differ, potentially indicating an increased link between chemical families. The use of health-based generic factors such as MDF could be a pragmatic starting point in cases where information is lacking. However, scientific knowledge should be carefully assessed for each detected mixture, and as more information on common exposures and risks becomes available this should be included into mixture risk assessment.

Applying the methods presented in this thesis, such as analyzing community structures in networks, could provide valuable insights into the most relevant mixtures of exposure for risk assessment purposes. While the focus of my work was on characterizing exposure to chemicals, it is important to prioritize the first tier of action, which should always be focused on prevention of exposure in the first place.

REFERENCES

Aylward, L. L., Hays, S. M., Smolders, R., Koch, H. M., Cocker, J., Jones, K., … Bevan, R. (2014a). Sources of Variability in Biomarker Concentrations. *Journal of Toxicology and Environmental Health, Part B*, *17*(1), 45–61. https://doi.org/10.1080/10937404.2013.864250

Aylward, L. L., Hays, S. M., Smolders, R., Koch, H. M., Cocker, J., Jones, K., … Bevan, R. (2014b). Sources of Variability in Biomarker Concentrations. *Journal of Toxicology and Environmental Health, Part B*, *17*(1), 45–61. https://doi.org/10.1080/10937404.2013.864250

Bergmann, A. J., Points, G. L., Scott, R. P., Wilson, G., & Anderson, K. A. (2018). Development of quantitative screen for 1550 chemicals with GC-MS. *Analytical and Bioanalytical Chemistry*, *410*(13), 3101–3110. https://doi.org/10.1007/s00216-018-0997-7

Bil, W., Govarts, E., Zeilmaker, M. J., Woutersen, M., Bessems, J., Ma, Y., … Vinggaard, A. M. (2023). Approaches to mixture risk assessment of PFASs in the European population based on human hazard and biomonitoring data. *International Journal of Hygiene and Environmental Health*, *247*, 114071. https://doi.org/https://doi.org/10.1016/j.ijheh.2022.114071

Boberg, J., Bredsdorff, L., Petersen, A., Löbl, N., Jensen, B. H., Vinggaard, A. M., & Nielsen, E. (2021). Chemical Mixture Calculator - A novel tool for mixture risk assessment. *Food and Chemical Toxicology*, *152*, 112167. https://doi.org/https://doi.org/10.1016/j.fct.2021.112167

Boberg, J., Dybdahl, M., Petersen, A., Hass, U., Svingen, T., & Vinggaard, A. M. (2019). A pragmatic approach for human risk assessment of chemical mixtures. *Current Opinion in Toxicology*, *15*, 1–7. https://doi.org/10.1016/j.cotox.2018.11.004

Bonvallot, N., Jamin, E. L., Regnaut, L., Chevrier, C., Martin, J.-F., Mercier, F., … Le Bot, B. (2021). Suspect screening and targeted analyses: Two complementary approaches to characterize human exposure to pesticides. *Science of The Total Environment*, *786*, 147499. https://doi.org/10.1016/j.scitotenv.2021.147499

Craig, J. A., Ceballos, D. M., Fruh, V., Petropoulos, Z. E., Allen, J. G., Calafat, A. M., … Webster, T. F. (2019). Exposure of Nail Salon Workers to Phthalates, Di(2-ethylhexyl) Terephthalate, and Organophosphate Esters: A Pilot Study. *Environmental Science & Technology*, *53*(24), 14630–14637. https://doi.org/10.1021/acs.est.9b02474

Dixon, H. M., Armstrong, G., Barton, M., Bergmann, A. J., Bondy, M., Halbleib, M. L., … Anderson, K. A. (2019). Discovery of common chemical exposures across three continents using silicone wristbands. *Royal Society Open Science*, *6*(2), 181836. https://doi.org/10.1098/rsos.181836

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Figueiredo, D. M., Krop, E. J. M., Duyzer, J., Gerritsen-Ebben, R. M., Gooijer, Y. M., Holterman, H. J., … Vermeulen, R. C. H. (2021). Pesticide Exposure of Residents Living Close to Agricultural Fields in the Netherlands: Protocol for an Observational Study. *JMIR Research Protocols*, *10*(4), e27883. https://doi.org/10.2196/27883

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, *486*(3–5), 75–174. https://doi.org/10.1016/j.physrep.2009.11.002

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432–441. https://doi.org/10.1093/biostatistics/kxm045

Fuhrimann, S., Mol, H. G. J., Dias, J., Dalvie, M. A., Röösli, M., Degrendele, C., … Vermeulen, R. (2022). Quantitative assessment of multiple pesticides in silicone wristbands of children/guardian pairs living in agricultural areas in South Africa. *Science of The Total Environment*, *812*, 152330. https://doi.org/10.1016/j.scitotenv.2021.152330

Gauglitz, J. M., Aceves, C. M., Aksenov, A. A., Aleti, G., Almaliti, J., Bouslimani, A., … Dorrestein, P. C. (2020). Untargeted mass spectrometry-based metabolomics approach unveils molecular changes in raw and processed foods and beverages. *Food Chemistry*, *302*, 125290. https://doi.org/10.1016/j.foodchem.2019.125290

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

Hamzai, L., Lopez Galvez, N., Hoh, E., Dodder, N. G., Matt, G. E., & Quintana, P. J. (2022a). A systematic review of the use of silicone wristbands for environmental exposure assessment, with a focus on polycyclic aromatic hydrocarbons (PAHs). *Journal of Exposure Science & Environmental Epidemiology*, *32*(2), 244–258. https://doi.org/10.1038/s41370-021-00359-9

Hamzai, L., Lopez Galvez, N., Hoh, E., Dodder, N. G., Matt, G. E., & Quintana, P. J. (2022b). A systematic review of the use of silicone wristbands for environmental exposure assessment, with a focus on polycyclic aromatic hydrocarbons (PAHs). *Journal of Exposure Science & Environmental Epidemiology*, *32*(2), 244–258. https://doi.org/10.1038/s41370-021-00359-9

Harakawa, R., Ito, T., & Iwahashi, M. (2022). Trend clustering from COVID-19 tweets using graphical lasso-guided iterative principal component analysis. *Scientific Reports*, *12*(1), 5709. https://doi.org/10.1038/s41598-022-09651-6

Hou, M., Shi, Y., Na, G., & Cai, Y. (2021). A review of organophosphate esters in indoor dust, air, hand wipes and silicone wristbands: Implications for human exposure. *Environment International*, *146*, 106261. https://doi.org/https://doi.org/10.1016/j.envint.2020.106261

Huber, C., Nijssen, R., Mol, H., Philippe Antignac, J., Krauss, M., Brack, W., … Lommen, A. (2022). A large scale multi-laboratory suspect screening of pesticide metabolites in human biomonitoring: From tentative annotations to verified occurrences. *Environment International*, *168*, 107452. https://doi.org/10.1016/j.envint.2022.107452

Hyland, C., Bradman, A., Gerona, R., Patton, S., Zakharevich, I., Gunier, R. B., & Klein, K. (2019). Organic diet intervention significantly reduces urinary pesticide levels in U.S. children and adults. *Environmental Research*, *171*, 568–575. https://doi.org/10.1016/j.envres.2019.01.024

Kalia, V., Walker, D. I., Krasnodemski, K. M., Jones, D. P., Miller, G. W., & Kioumourtzoglou, M.-A. (2020). Unsupervised dimensionality reduction for exposome research. *Current Opinion in Environmental Science & Health*, *15*, 32–38. https://doi.org/https://doi.org/10.1016/j.coesh.2020.05.001

Koch, H. M., Aylward, L. L., Hays, S. M., Smolders, R., Moos, R. K., Cocker, J., … Bevan, R. (2014). Inter- and intra-individual variation in urinary biomarker concentrations over a 6-day sampling period. Part 2: Personal care product ingredients. *Toxicology Letters*, *231*(2), 261–269. https://doi.org/https://doi.org/10.1016/j.toxlet.2014.06.023

Lafit, G., Tuerlinckx, F., Myin-Germeys, I., & Ceulemans, E. (2019). A Partial Correlation Screening Approach for Controlling the False Positive Rate in Sparse Gaussian Graphical Models. *Scientific Reports*, *9*(1), 17759. https://doi.org/10.1038/s41598-019-53795-x

Lesliam, Q.-A., & Boyd, B. D. (2023). Invited Perspective: Mixtures—Are They Worth the Risk (Assessment)? *Environmental Health Perspectives*, *131*(4), 41301. https://doi.org/10.1289/EHP12596

Lin, E. Z., Esenther, S., Mascelloni, M., Irfan, F., & Godri Pollitt, K. J. (2020). The Fresh Air Wristband: A Wearable Air Pollutant Sampler. *Environmental Science & Technology Letters*, *7*(5), 308–314. https://doi.org/10.1021/acs.estlett.9b00800

Loh, M. M., Schmidt, P., Christopher de Vries, Y., Vogel, N., Kolossa-Gehring, M., Vlaanderen, J., … Luijten, M. (2023). Toxicity Weighting for Human Biomonitoring Mixture Risk Assessment: A Proof of Concept. *Toxics*, *11*(5), 408. https://doi.org/10.3390/toxics11050408

Longnecker, M. P., Rogan, W. J., & Lucier, G. (1997). The human health effects of DDT (dichlorodiphenyltrichloroethane) and PCBs (polychlorinated biphenyls) and an overview of organochlorines in public health. *Annual Review of Public Health*, *18*(1), 211–244. https://doi.org/10.1146/annurev.publhealth.18.1.211

McCord, J. P., Groff, L. C., & Sobus, J. R. (2022a). Quantitative non-targeted analysis: Bridging the gap between contaminant discovery and risk characterization. *Environment International*, *158*, 107011. https://doi.org/https://doi.org/10.1016/j.envint.2021.107011

McCord, J. P., Groff, L. C., & Sobus, J. R. (2022b). Quantitative non-targeted analysis: Bridging the gap between contaminant discovery and risk characterization. *Environment International*, *158*, 107011. https://doi.org/https://doi.org/10.1016/j.envint.2021.107011

Meinshausen, N., & Bühlmann, P. (2006). High-dimensional graphs and variable selection with the Lasso. *The Annals of Statistics*, *34*(3), 1436–1462. https://doi.org/10.1214/009053606000000281

Mustafa, E., Valente, M. J., & Vinggaard, A. M. (2023). Complex chemical mixtures: Approaches for assessing adverse human health effects. *Current Opinion in Toxicology*, *34*, 100404. https://doi.org/https://doi.org/10.1016/j.cotox.2023.100404

O'Connell, S. G., Kincl, L. D., & Anderson, K. A. (2014). Silicone Wristbands as Personal Passive Samplers. *Environmental Science & Technology*, *48*(6), 3327–3335. https://doi.org/10.1021/es405022f

Ougier, E., Ganzleben, C., Lecoq, P., Bessems, J., David, M., Schoeters, G., … Vicente, J. L. (2021). Chemical prioritisation strategy in the European Human Biomonitoring Initiative (HBM4EU) – Development and results. *International Journal of Hygiene and Environmental Health*, *236*, 113778. https://doi.org/https://doi.org/10.1016/j.ijheh.2021.113778

Palla, G., Barabási, A.-L., & Vicsek, T. (2007). Quantifying social group evolution. *Nature*, *446*(7136), 664–667. https://doi.org/10.1038/nature05670

Palla, G., Derényi, I., Farkas, I., & Vicsek, T. (2005). Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, *435*(7043), 814–818. https://doi.org/10.1038/nature03607

Pellizzari, E. D., Woodruff, T. J., Boyles, R. R., Kannan, K., Beamer, P. I., Buckley, J. P., … Bennett, D. H. (2019). Identifying and Prioritizing Chemicals with Uncertain Burden of Exposure: Opportunities for Biomonitoring and Health-Related Research. *Environmental Health Perspectives*, *127*(12), EHP5133. https://doi.org/10.1289/EHP5133

Perrier, F., Giorgis-Allemand, L., Slama, R., & Philippat, C. (2016). Within-subject Pooling of Biological Samples to Reduce Exposure Misclassification in Biomarker-based Studies. *Epidemiology*, *27*(3), 378–388. https://doi.org/10.1097/EDE.0000000000000460

Petrovic, D., Bodinier, B., Dagnino, S., Whitaker, M., Karimi, M., Campanella, G., … Chadeau-Hyam, M. (2022). Epigenetic mechanisms of lung carcinogenesis involve differentially methylated CpG sites beyond those associated with smoking. *European Journal of Epidemiology*, *37*(6), 629–640. https://doi.org/10.1007/s10654-022-00877-2

Pons, P., & Latapy, M. (2005). Computing Communities in Large Networks Using Random Walks (pp. 284–293). https://doi.org/10.1007/11569596_31

Pourchet, M., Debrauwer, L., Klanova, J., Price, E. J., Covaci, A., Caballero-Casero, N., … Antignac, J.-P. (2020). Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: From promises to challenges and harmonisation issues. *Environment International*, *139*, 105545. https://doi.org/10.1016/j.envint.2020.105545

Price, E. J., Vitale, C. M., Miller, G. W., David, A., Barouki, R., Audouze, K., … Klánová, J. (2022). Merging the exposome into an integrated framework for "omics" sciences. *IScience*, *25*(3), 103976. https://doi.org/10.1016/j.isci.2022.103976

Rempelos, L., Wang, J., Barański, M., Watson, A., Volakakis, N., Hoppe, H.-W., … Leifert, C. (2022). Diet and food type affect urinary pesticide residue excretion profiles in healthy individuals: results of a randomized controlled dietary intervention trial. *The American Journal of Clinical Nutrition*, *115*(2), 364–377. https://doi.org/10.1093/ajcn/nqab308

Rísová, V. (2019). The pathway of lead through the mother's body to the child. *Interdisciplinary Toxicology*, *12*(1), 1–6. https://doi.org/10.2478/intox-2019-0001

Samon, S. M., Hammel, S. C., Stapleton, H. M., & Anderson, K. A. (2022). Silicone wristbands as personal passive sampling devices: Current knowledge, recommendations for use, and future directions. *Environment International*, *169*, 107339. https://doi.org/10.1016/j.envint.2022.107339

Santos, S., Maitre, L., Warembourg, C., Agier, L., Richiardi, L., Basagaña, X., & Vrijheid, M. (2020). Applying the exposome concept in birth cohort research: a review of statistical approaches. *European Journal of Epidemiology*, *35*(3), 193–204. https://doi.org/10.1007/s10654-020-00625-4

Savitz., D. A., & Hattersley., A. M. (2023). Evaluating Chemical Mixtures in Epidemiological Studies to Inform Regulatory Decisions. *Environmental Health Perspectives*, *131*(4), 45001. https://doi.org/10.1289/EHP11899

Schwikowski, B., Uetz, P., & Fields, S. (2000). A network of protein–protein interactions in yeast. *Nature Biotechnology*, *18*(12), 1257–1261. https://doi.org/10.1038/82360

Schymanski, E. L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H. P., & Hollender, J. (2014). Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environmental Science & Technology*, *48*(4), 2097–2098. https://doi.org/10.1021/es5002105

Tamayo-Uria, I., Maitre, L., Thomsen, C., Nieuwenhuijsen, M. J., Chatzi, L., Siroux, V., … Basagaña, X. (2019). The early-life exposome: Description and patterns in six European countries. *Environment International*, *123*, 189–200. https://doi.org/10.1016/j.envint.2018.11.067

Tarazona, J. V., Cattaneo, I., Niemann, L., Pedraza-Diaz, S., González-Caballero, M. C., de Alba-Gonzalez, M., … Santonen, T. (2022). A Tiered Approach for Assessing Individual and Combined Risk of Pyrethroids Using Human Biomonitoring Data. *Toxics*, *10*(8), 451. https://doi.org/10.3390/toxics10080451

Taylor, K. W., Joubert, B. R., Braun, J. M., Dilworth, C., Gennings, C., Hauser, R., … Carlin, D. J. (2016). Statistical Approaches for Assessing Health Effects of Environmental Chemical Mixtures in Epidemiology: Lessons from an Innovative Workshop. *Environmental Health Perspectives*, *124*(12). https://doi.org/10.1289/EHP547

Travis, S. C., Kordas, K., & Aga, D. S. (2021). Optimized workflow for unknown screening using gas chromatography high-resolution mass spectrometry expands identification of contaminants in silicone personal passive samplers. *Rapid Communications in Mass Spectrometry*, *35*(8). https://doi.org/10.1002/rcm.9048

Vahter, M. (2008). Health Effects of Early Life Exposure to Arsenic. *Basic & Clinical Pharmacology & Toxicology*, *102*(2), 204–211. https://doi.org/10.1111/j.1742-7843.2007.00168.x

van den Berg, M., Kypke, K., Kotz, A., Tritscher, A., Lee, S. Y., Magulova, K., … Malisch, R. (2017). WHO/UNEP global surveys of PCDDs, PCDFs, PCBs and DDTs in human milk and benefit–risk evaluation of breastfeeding. *Archives of Toxicology*, *91*(1), 83–96. https://doi.org/10.1007/s00204-016-1802-z

Vermeulen, R. (2022). Invited Perspective: Inroads to Biology-Informed Exposomics. *Environmental Health Perspectives*, *130*(11). https://doi.org/10.1289/EHP12224

Vermeulen, R., Schymanski, E. L., Barabási, A.-L., & Miller, G. W. (2020a). The exposome and health: Where chemistry meets biology. *Science*, *367*(6476), 392–396. https://doi.org/10.1126/science.aay3164

Vermeulen, R., Schymanski, E. L., Barabási, A.-L., & Miller, G. W. (2020b). The exposome and health: Where chemistry meets biology. *Science*, *367*(6476), 392–396. https://doi.org/10.1126/science.aay3164

Vernet, C., Philippat, C., Calafat, A. M., Ye, X., Lyon-Caen, S., Siroux, V., … Slama, R. (2018). Within-Day, Between-Day, and Between-Week Variability of Urinary Concentrations of Phenol Biomarkers in Pregnant Women. *Environmental Health Perspectives*, *126*(3), 37005. https://doi.org/10.1289/EHP1994

Vitale, C. M., Lommen, A., Huber, C., Wagner, K., Garlito Molina, B., Nijssen, R., … Antignac, J.-P. (2022). Harmonized Quality Assurance/Quality Control Provisions for Nontargeted Measurement of Urinary Pesticide Biomarkers in the HBM4EU Multisite SPECIMEn Study. *Analytical Chemistry*. https://doi.org/10.1021/acs.analchem.2c00061

Wacławik, M., Rodzaj, W., & Wielgomas, B. (2022a). Silicone Wristbands in Exposure Assessment: Analytical Considerations and Comparison with Other Approaches. *International Journal of Environmental Research and Public Health*, *19*(4), 1935. https://doi.org/10.3390/ijerph19041935

Wacławik, M., Rodzaj, W., & Wielgomas, B. (2022b). Silicone Wristbands in Exposure Assessment: Analytical Considerations and Comparison with Other Approaches. *International Journal of Environmental Research and Public Health*, *19*(4), 1935. https://doi.org/10.3390/ijerph19041935

Wang, A., Gerona, R. R., Schwartz, J. M., Lin, T., Sirota, M., Morello-Frosch, R., & Woodruff, T. J. (2018). A Suspect Screening Method for Characterizing Multiple Chemical Exposures among a Demographically Diverse Population of Pregnant Women in San Francisco. *Environmental Health Perspectives*, *126*(7), 077009. https://doi.org/10.1289/EHP2920

Wild, C. P. (2005). Complementing the Genome with an "Exposome": The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology. *Cancer Epidemiology, Biomarkers & Prevention*, *14*(8), 1847–1850. https://doi.org/10.1158/1055-9965.EPI-05-0456

Wild, C. P. (2012). The exposome: from concept to utility. *International Journal of Epidemiology*, *41*(1), 24–32. https://doi.org/10.1093/ije/dyr236

Ye, M., Beach, J., Martin, J. W., & Senthilselvan, A. (2015). Associations between dietary factors and urinary concentrations of organophosphate and pyrethroid metabolites in a Canadian general population. *International Journal of Hygiene and Environmental Health*, *218*(7), 616–626. https://doi.org/10.1016/j.ijheh.2015.06.006

Zhang, D., & Lu, S. (2022). Human exposure to neonicotinoids and the associated health risks: A review. *Environment International*, *163*, 107201. https://doi.org/https://doi.org/10.1016/j.envint.2022.107201

Appendices

ENGLISH SUMMARY

Humans are exposed to a large number of chemicals from various sources in their environment, including cosmetics, household products, diet, medication, and occupational sources. To properly assess the risks associated with these chemical exposures, it is necessary to estimate the level of exposure (how much) and the toxicity posed by these chemicals. As the number of chemicals available on the market increases and new applications for existing chemicals arise, there is a need for efficient high throughput methods for the risk assessment of a large number of chemicals. Current risk assessment approaches often focus on individual chemicals or single sources, which not sufficiently account for possible combinations of chemicals and overlooks potential additional risks associated with chemical mixtures. Moreover, for those chemicals that act together in the same toxicological pathway, combined risk assessment is key. To accurately assess the risks of chemical mixtures in relation to human health, three main themes can be identified: i) measuring combined exposures to chemical mixtures, ii) accurately assessing the health risks posed by these chemical mixtures, and iii) translating these risks to the policy and regulatory fields to manage mixture health risks. In this thesis I focus on the first challenge, which involves measuring and describing exposure patterns of chemical mixtures in the general population.

Measuring exposure to chemical mixtures in real-life situations at a population level is a complex task. The preferably individual measurements can be obtained either externally (outside the human body), using silicone wristbands for instance, or internally (inside the human body), through human biomonitoring, for which exposure is measured in human tissue, such as urine or blood. In this thesis, chemical mixtures are described as combinations of manufactured chemicals that co-occur within the same individual or sample.

This thesis

My thesis revolves around three different approaches to identify chemical mixtures in the general population.

1. The first approach focusses on describing co-occurrence patterns of chemicals in existing human biomonitoring datasets, using a combination of graphical correlation network model and a clustering algorithm. This method visualizes and summarizes these co-occurrence patterns by identifying highly correlated groups or clusters of chemicals.
   These correlation networks are explored in the first half of this thesis (**chapters 2 and 3**) and applied to existing human biomonitoring datasets.
2. The second approach involves measuring chemical mixtures at individual level, employing external measurements (silicone wristbands) and internal measurements (human biomonitoring in urine samples).

3. The third approach refers to the analytical measurement of chemical mixtures in urine samples, for which a suspect screening approach based on high resolution mass spectrometry was applied, enabling detection of a broad range of biomarkers in a single sample.

The second and third approach make up the second half of this thesis (**chapters 4, 5 and 6**), specifically examining pesticide mixtures. Samples were collected to assess exposure to multiple pesticides, including silicone wristbands as part of the Dutch OBO study (**chapter 4**), and urine samples combined with suspect screening as part of the European initiative HBM4EU (**chapters 5 and 6**).

The second half of this thesis focusses on pesticides as an example of a chemical mixture. Pesticides are commonly used as mixtures by farmers (potential environmental exposure to pesticide mixtures for residents living in agricultural areas), and pesticides are present in various food items (dietary exposure to pesticide mixtures). Besides, pesticides are a relevant chemical mixture due to the attention and concerns they generate in society. Despite the extensive body of literature on human exposure to pesticides, it remains unclear which pesticide mixtures the general population is exposed to.

After **Chapter 1** as a general introduction of this thesis, **Chapter 2** describes a correlation network approach to visualize chemical mixtures based on human biomonitoring data. We demonstrated that presenting biomonitoring data in networks facilitates the identification of exposure patterns that contribute to the observed exposure levels in the samples. This approach was applied to cord blood samples collected during three measurement campaigns of the Flemish Environment and Health Studies (FLEHS). By utilizing a clustering algorithm, highly connected groups of biomarkers, referred to as communities, were identified. These communities mostly consisted of biomarkers belonging to the same chemical family, such as PFASs and metals. Links between chemical families were also found, such as between PCBs and p,p'-DDE. The detected communities highlight which biomarkers should be considered collectively in the analysis and interpretation of epidemiological studies or in the design of toxicological mixture studies. The comparison of networks between sampling campaigns, smoking status, and BMI provided insights into the extent of changes in networks and communities across these covariates. The applied network methodology benefits over e.g. heatmaps by the intuitive visual identification, formalization of the procedure to identify communities and providing a structural approach for comparison of exposure patterns.

In **Chapter 3** the same correlation network visualization was applied to four existing human biomonitoring datasets from Belgium, Czech Republic, Germany and Spain. The added value of network analyses was demonstrated at a larger scale and showed potential for a cross-European mixture exposure and risk assessment. This was the first, to our knowledge, application of a harmonized and standardized network approach on multiple datasets from different countries, incorporating a substantial number of biomarkers and individuals. The four included datasets varied in study population, study design and chemicals analyzed. Compared to the

data in **chapter 2**, the datasets included in this chapter were on a larger scale, both in terms of the number of participants and the number of biomarkers measured. The application of community detection with a clustering algorithm was instrumental in identifying patterns within and between chemical families. Within each country, the majority of the detected communities reflected a single chemical family. Cross chemical family connections were reflected in two out of the nine detected communities, one out of four, two out of ten, and three out of six, for the four countries respectively. Other observed differences between countries, such as a variation in the phthalate community distribution between Germany and Spain (split into seven and four communities respectively), maybe be attributed to differences in study population, design, chemical analytical procedures and actual exposure differences. It is plausible that the detected communities reflect shared sources, routes of exposure, or physiochemical properties among the biomarkers.

The assessment of pesticide mixtures begins with **Chapter 4**, where silicone wristbands were utilized as an add-on to a Dutch pesticide exposure study. Nineteen residents living close to flower bulb fields in the Netherlands wore the wristbands for an average of sixty days. Out of forty-six pesticides measured, thirty-one were identified in the wristbands. On average, nineteen pesticides were detected, with azoxystrobin, carbendazim and pymetrozine detected in all wristbands. We found highly individual exposure profiles, which is similar to findings of other studies. The study showed that wristbands are able to capture pesticide exposure profiles over sixty days, which is substantial longer than the usual application of several days or a week. However, it should be noted that they only reflect dermal and inhalation routes of exposure, which may lead to an underestimation of the pesticide mixtures for pesticides that are also present in the diet.

In **Chapter 5** the study protocol and results of a European pesticide exposure study (SPECIMEn study) are presented. The study employed a suspect screening based protocol using first morning void urine samples from parent-child pairs in different seasons and residential areas. The objective was to investigate the impact of proximity to agricultural areas, age (child vs adult), and sampling season on the likelihood of detecting pesticides and their metabolites. Biomonitoring samples were collected of 1,345 participants across six countries: Latvia, Hungary, Czech Republic, Spain, the Netherlands and Switzerland (with Switzerland having a slightly different study design). Samples were analyzed by five different laboratories across Europe. Forty pesticide biomarkers, including pesticide metabolites and parent pesticides, corresponding to 29 different pesticides were confidently identified across all study sites. The most frequently detected biomarkers were associated with the parent pesticides acetamiprid and chlorpropham. In 84% of the samples, at least two different pesticides were detected. The median number of detected pesticides in the urine samples was three. Some variation but no consistent pattern in the probability of detection of pesticide biomarkers was observed based on living in an agricultural area or season of urine sampling. Notably, differences in detection were observed between adults and children, suggesting a different exposure and/or elimination patterns between these age groups. This chapter demonstrated that a harmonized

pan-European sample collection, combined with suspect screening provided valuable new insights into the presence of pesticide mixture exposure in the European population.

Chapter 6 is also based on the SPECIMEn study, focusing on the samples collected in the Netherlands and Switzerland. Since both countries' samples were analyzed by the same laboratory, the pesticide biomarkers could be analyzed by their semi-quantitative levels between the two countries. The objective of this chapter was to describe the pesticide mixture patterns in the adult populations of the Netherlands and Switzerland, and to identify related exposure determinants. A total 400 adults were included in this study, and questionnaires between both study sites were harmonized. Among the 37 detected biomarkers, only three were found in at least 40% of the samples. The most frequent combination, acetamiprid with flonicamid, was detected in 22 samples (5.5%). Regression models revealed an inverse association between high organic vegetable and fruit consumption and exposure to urinary concentrations of acetamiprid and chlorpropham. Other exposure determinants assessed related to pesticide usage, exposed household members, distance to forest or agriculture, and consumption of organic rice, meat, eggs, diary, bread and homegrown foods, did not reveal any effect in the regression models. Within-individual correlations in repeated samples (summer/winter) from the Netherlands were low (≤ 0.3), and no significant seasonal differences in average exposures were observed in Switzerland. In both countries, detection rates and co-occurrence of pesticides in the same urine sample were typically low.

Chapter 7 provides a comprehensive review of the main findings presented in this thesis and discusses the applicability of the three different approaches employed to identify chemical mixtures. The work conducted in this thesis represents significant steps towards describing and measuring exposure to chemical mixtures. Understanding real-life exposure patterns is crucial for conducting mixture risk assessment. Co-occurring chemicals that share a common mode of action should be collectively considered together in risk assessments.

Two directions for future progress can be identified: 1) I recommend to incorporate correlation networks to investigate interdependencies between chemicals measured in biomonitoring data. The network approach as applied in this thesis can be improved by incorporating overlapping clusters and identifying temporal changes to facilitate addressing specific research questions. 2) High-resolution screening methods should be integrated in HBM studies to build a comprehensive database to assess exposure and risks to chemical mixtures. To improve the collection and sharing of exposure data from chemical mixtures, I believe it is important that agreements are made at the European level regarding the minimum information that must be shared when publishing a biomonitoring dataset that can contribute to risk assessment. A good description of the analytical method used and an overview of the observed correlations with the dataset are important points.

NEDERLANDSE SAMENVATTING

Mensen worden blootgesteld aan een groot aantal chemische stoffen uit verschillende bronnen vanuit hun omgeving, vanuit cosmetica, huishoudelijke producten, voeding, medicijnen, en tijdens het werk. Om de risico's van de blootstelling aan chemische stoffen goed te kunnen beoordelen is het nodig om het niveau van blootstelling (hoeveelheid) en de toxiciteit van deze stoffen in te schatten. Aangezien het aantal beschikbare chemische stoffen toeneemt en er continue nieuwe toepassingen voor bestaande chemische stoffen ontstaan, is er behoefte aan beoordelingsmethodes die efficiënt op een groot aantal stoffen tegelijk toepasbaar zijn. Binnen de huidige kaders van risicobeoordeling wordt er vaak naar afzonderlijke chemische stoffen of afzonderlijke bronnen gekeken, waardoor er onvoldoende rekening wordt gehouden met mogelijke combinaties van chemische stoffen en de mogelijke extra risico's van deze chemische mengsels. Daarnaast is een gecombineerde risicobeoordeling van cruciaal belang voor chemische stoffen die mogelijk hetzelfde toxicologische effect hebben. Het inschatten van gezondheidsrisico's van chemische mengsels kan worden samengevat in drie belangrijke thema's: i) het meten van gecombineerde blootstelling aan chemische mengsels, ii) het beoordelen van de gezondheidsrisico's van deze mengsels, en iii) het vertalen van deze risico's naar het beleid en de regelgeving om zo de gezondheidsrisico's van mengsels te verlagen. Dit proefschrift richt zich op het eerste thema, die bestaat uit het meten en beschrijven van blootstellingspatronen van chemische mengsels met een specifieke nadruk op het gezamenlijk voorkomen van chemische stoffen in de algemene bevolking.

Het meten van de daadwerkelijke blootstelling aan chemische mengsels is een complexe taak. De bij voorkeur individuele metingen kunnen extern (buiten het lichaam) worden uitgevoerd, bijvoorbeeld met siliconen polsbandjes, of intern (in het lichaam) via bijvoorbeeld biomonitoring, waarbij blootstelling wordt gemeten in menselijk lichaamsmateriaal zoals urine of bloed. In dit proefschrift worden chemische mengsels beschreven als de combinaties van door de mens gemaakte chemische stoffen die samen in hetzelfde individu of monster gevonden worden.

Dit proefschrift

Mijn proefschrift beschrijft drie verschillende benaderingen om chemische mengsels in de algemene bevolking te identificeren.

1. De eerste methode richt zich op het beschrijven van patronen van chemische stoffen die gelijktijdig voorkomen in bestaande biomonitoring datasets. Deze patronen worden beschreven met behulp van een grafisch correlatienetwerkmodel en een clusteralgoritme. Met deze methode kunnen de patronen van stoffen die gelijktijdig voorkomen worden gevisualiseerd en samengevat door middel van identificatie van sterk gecorreleerde groepen of clusters van chemische stoffen.

Deze correlatienetwerken worden beschreven en toegepast op bestaande biomonitoring data in de eerste helft van dit proefschrift (**hoofdstukken 2 en 3)**.

2. De tweede benadering is het meten van chemische mengsels op individueel niveau, waarbij gebruik wordt gemaakt van externe metingen (siliconen polsbandjes) en interne metingen (biomonitoring door middel van urinemonsters).
3. De derde benadering verwijst naar de analytische meting van mengsel in urinemonsters, waarvoor een screening methode ("suspect screening") op basis van hoge resolutie massaspectrometrie werd toegepast. Deze op screening gebaseerde analyse kan een breed scala aan biomarkers detecteren in een enkel urinemonster.

De tweede en derde benadering vormen samen de tweede helft van dit proefschrift (**hoofdstukken 4, 5 en 6**), waarbij specifiek naar mengsels van bestrijdingsmiddelen werd gekeken. Er werden monsters verzameld om de blootstelling aan meerdere bestrijdingsmiddelen in kaart te brengen, waaronder siliconen polsbandjes als onderdeel van het Nederlandse OBO-onderzoek (**hoofdstuk 4**), en urinemonsters gecombineerd met suspect screening analyse als onderdeel van het Europese initiatief HBM4EU **(hoofdstukken 5 en 6)**.

De tweede helft van dit proefschrift richt zich op bestrijdingsmiddelen als een voorbeeld van een chemisch mengsel. Bestrijdingsmiddelen worden vaak als een mengsel toegepast door boeren (potentiële mengsel blootstelling voor omwonenden van landbouwgebieden), en bestrijdingsmiddelen zijn aanwezig in verschillende voedingsmiddelen (potentiële mengsel blootstelling via de voeding). Daarnaast zijn bestrijdingsmiddelen een relevant chemisch mengsel vanwege de maatschappelijke aandacht en zorgen over mogelijke risico's. Ondanks een groot aantal studies naar de blootstelling aan bestrijdingsmiddelen, is de vraag aan welke combinaties van bestrijdingsmiddelen de algemene bevolking wordt blootgesteld nog niet beantwoord.

Na **hoofdstuk 1** als algemene inleiding van dit proefschrift, beschrijft **hoofdstuk 2** een correlatienetwerkmethode om chemische mengsels te visualiseren op basis van biomonitoringsgegevens. We laten hier zien dat deze netwerken de identificatie van patronen tussen biomarkers kan vergemakkelijken. De netwerkmethode werd in dit hoofdstuk toegepast op navelstrengbloedmonsters die verzameld werden tijdens drie meetcampagnes van de Vlaamse Milieu- en Gezondheidsstudies (FLEHS). Door gebruik te maken van een clusteralgoritme werden sterk verbonden groepen biomarkers, "communities" genoemd, geïdentificeerd. Deze communities bestonden meestal uit biomarkers die tot dezelfde chemische familie behoren, zoals PFASs en metalen. Er werden echter ook verbanden tùssen chemische families werden gevonden, zoals tussen de PCB's en p,p'-DDE. De gevonden communities geven inzicht in welke chemische stoffen gezamenlijk moeten worden meegenomen bij de analyse en interpretatie van epidemiologische studies, of bij het ontwerp van toxicologische mengsel studies. Door netwerken te vergelijken tussen meetcampagnes, wel/niet roken, en normaal/hoog BMI verkregen we inzicht in veranderingen in de netwerkstructuur en bijbehorende communities. De voordelen van de toegepaste netwerkmethode ten opzichte van bijvoorbeeld "heatmaps", zijn de intuï-

tieve visuele interpretatie, formalisering van de procedure om communities te identificeren en een gestructureerde methode voor het vergelijken van blootstellingspatronen.

In **hoofdstuk 3** werd dezelfde netwerk visualisatie methode toegepast op vier bestaande biomonitoring datasets uit België, Tsjechië, Duitsland en Spanje. De toegevoegde waarde van netwerken werd gedemonstreerd op een grotere schaal, naast het potentieel van netwerken voor een pan-Europese mengsel blootstellings- en risicobeoordeling. Dit was, voor zover wij weten, de eerste toepassing van een geharmoniseerde en gestandaardiseerde netwerkmethode op meerdere biomonitoring datasets uit verschillende landen. De vier gebruikte datasets varieerden in studiepopulatie, onderzoeksopzet en geanalyseerde stoffen. Vergeleken met de dataset uit **hoofdstuk 2,** waren de vier datasets grootschaliger, zowel wat betreft het aantal deelnemers als het aantal gemeten biomarkers. De toepassing van het cluster algoritme was een waardevol hulpmiddel om patronen tussen en binnen chemische families te identificeren. Binnen elk land reflecteerde de meerderheid van de gedetecteerde communities patronen binnen dezelfde chemische familie. Kruisverbanden tussen chemische families werden gevonden in twee van de negen communities, één van de vier, twee van de tien, en drie van de zes, respectievelijk voor de vier landen. Andere gedetecteerde verschillen tussen landen, zoals een verschil in spreiding tussen weekmakers tussen Duitsland en Spanje (opgesplitst in respectievelijk zeven en vier communities), kunnen worden toegeschreven aan verschillen in studiepopulatie, design, chemische analyseprocedures, en feitelijke blootstellingsverschillen. Het is aannemelijk dat de gedetecteerde communities een gemeenschappelijk bron, blootstellingsroute of fysiochemische eigenschappen van de biomarkers weerspiegelen.

De beoordeling van mengsels van bestrijdingsmiddelen begint met **hoofdstuk 4**, waarin siliconen polsbandjes werden gedragen als aanvulling op een Nederlands onderzoek naar blootstelling aan bestrijdingsmiddelen. Negentien bewoners die in de buurt van bloembollenvelden in Nederland woonden, droegen de polsbandjes gedurende gemiddeld zestig dagen. Van de zesenveertig gemeten bestrijdingsmiddelen werden er eenendertig gevonden in de polsbandjes. Gemiddeld werden negentien bestrijdingsmiddelen gemeten, waarbij azoxystrobin, carbendazim en pymetrozine in alle polsbandjes werden gevonden. We vonden zeer individuele blootstellingsprofielen, wat vergelijkbaar is met bevindingen uit andere studies. Het onderzoek toonde aan dat blootstellingspatronen over een langere periode van zestig dagen gemeten konden worden door middel van polsbandjes, wat aanzienlijk langer is dan gebruikelijk (enkele dagen of een week). Er dient de kanttekening gemaakt te worden dat de polsbandjes alleen blootstellingsroutes via de huid en luchtwegen reflecteren, wat kan leiden tot een onderschatting van het mengel van bestrijdingsmiddelen ook aanwezig in voeding.

In **hoofdstuk 5** worden het onderzoeksprotocol en de resultaten van een Europees onderzoek naar blootstelling aan bestrijdingsmiddelen (SPECIMEn-studie) beschreven. Het onderzoek maakte gebruik van suspect screening analyse, waarbij urinemonsters werden geanalyseerd van ouder-kind paren in verschillende seizoenen en woongebieden. Het doel van deze studie was om de invloed van het wonen nabij landbouwgebieden, leeftijd (kind vs. volwassene)

en seizoen op de detectie van bestrijdingsmiddelen te beoordelen. Er werden urinemonsters verzameld van 1.345 deelnemers in zes landen: Letland, Hongarije, Tsjechië, Spanje, Nederland en Zwitserland (waarbij Zwitserland een iets andere onderzoeksopzet had). De monsters werden geanalyseerd door vijf verschillende laboratoria verspreid over Europa. Er werden met hoge mate van zekerheid 40 biomarkers (metabolieten en oorspronkelijke stoffen) geïdentificeerd, die zijn terug te voeren tot 29 verschillende bestrijdingsmiddelen. De meest frequent gedetecteerde biomarkers zijn gerelateerd aan de bestrijdingsmiddelen acetamiprid en chloorprofam. In 84% van de monsters werden ten minste twee verschillende bestrijdingsmiddelen gedetecteerd. De mediaan van het aantal gedetecteerde bestrijdingsmiddelen in de urinemonsters was drie. Er werd wel enige variatie maar geen consistent patroon in de detectie van bestrijdingsmiddelen waargenomen op basis van wonen nabij landbouwgebied of het seizoen van de urinemonstering. Er werden met name verschillen in detectie waargenomen tussen volwassenen en kinderen, wat duidt op verschillende blootstellings- en/of eliminatiepatronen tussen deze leeftijdsgroepen. Dit hoofdstuk heeft laten zien dat een geharmoniseerde pan-Europese monsterverzameling, gecombineerd met suspect screening, waardevolle nieuwe inzichten kan opleveren over welke bestrijdingsmiddelenmengsels aanwezigheid zijn in de Europese bevolking.

Hoofdstuk 6 is ook gebaseerd op de SPECIMEn-studie en heeft betrekking op de monsters die in Nederland en Zwitserland zijn verzameld. Aangezien de urinemonsters van beide landen door hetzelfde laboratorium werden geanalyseerd, konden semi-kwantitatieve niveaus van de gemeten biomarkers worden vergeleken tussen de twee landen. Het doel van dit hoofdstuk was om de patronen in blootstelling aan meerdere bestrijdingsmiddelen in de volwassen populaties van Nederland en Zwitserland te beschrijven, en om gerelateerde blootstellingsdeterminanten te identificeren. In totaal werden 400 volwassenen geïncludeerd in dit onderzoek en de vragenlijsten van beide onderzoek locaties werden geharmoniseerd. Van de 37 gedetecteerde biomarkers werden er slechts drie in minstens 40% van de monsters aangetroffen. De meest voorkomende combinatie, acetamiprid met flonicamid, werd aangetroffen in 22 monsters (5,5%). Regressiemodellen toonden een associatie tussen een hoge consumptie van biologische groente en fruit en lagere blootstelling aan acetamiprid en chloorprofam. Andere onderzochte blootstellingsdeterminanten, gerelateerd aan bestrijdingsmiddelen gebruik, huisgenoten die werkzaam zijn met bestrijdingsmiddelen, afstand tot bos of landbouw, en consumptie van biologische rijst, vlees, eieren, zuivel, brood en zelfgekweekt voedsel, lieten geen effect zien in de regressiemodellen. Intra-individuele correlaties in herhaalde steekproeven (zomer/winter) uit Nederland waren laag (≤ 0,3), en in Zwitserland werden geen significante seizoensgebonden verschillen in gemiddelde blootstellingen waargenomen. In beide landen waren de detectiepercentages en de gelijktijdige aanwezigheid van bestrijdingsmiddelen in dezelfde urinemonsters laag.

Hoofdstuk 7 geeft een overzicht van de belangrijkste bevindingen in dit proefschrift en belicht de toepasbaarheid van de drie verschillende benaderingen die zijn gebruikt om chemische mengsels te identificeren. De resultaten uit dit proefschrift zijn een belangrijke stap

naar het beschrijven en meten van blootstelling aan chemische mengsels. Inzicht in de daadwerkelijke is cruciaal om het risico van mengsels te kunnen inschatten. De tegelijk voorkomende chemische stoffen met een gemeenschappelijk werkingsmechanisme dienen samen in beschouwing te worden genomen in risicobeoordelingen.

Twee aanbevelingen voor toekomstig onderzoek kunnen worden geïdentificeerd: 1) Ik adviseer om correlatienetwerken mee te nemen in de om het onderlinge verband tussen chemische stoffen te onderzoeken in biomonitoringsdata. De netwerkmethode zoals toegepast in deze thesis kan voor bepaalde onderzoeksvragen worden verfijnd door bijvoorbeeld overlappende clusters en identificatie van veranderingen in de tijd te beschrijven. 2) Hoge-resolutie screeningmethoden moeten worden geïntegreerd in biomonitoringstudies om een uitgebreide database op te bouwen, om zo de blootstelling aan en risico's van mengsels van chemische stoffen te kunnen beoordelen. Om het verzamelen en delen van blootstellingsgegevens van chemische mengsels te verbeteren, vind ik het belangrijk dat er op Europees niveau afspraken gemaakt worden wat betreft de minimale informatie die gedeeld moet worden bij het publiceren van een biomonitoring dataset die kan bijdragen aan risicobeoordeling. Een goede beschrijving van de gebruikte analytische methode en een overzicht van de geobserveerde correlaties binnen de dataset zijn daarbij belangrijke punten.

SCIENTIFIC PUBLICATIONS

Bekker, C., Fransman, W., Boessen, R., Oerlemans, A., **Ottenbros, I. B.**, & Vermeulen, R. (2017). Assessment of Determinants of Emission Potentially Affecting the Concentration of Airborne Nano-Objects and Their Agglomerates and Aggregates. *Annals of Work Exposures and Health*, *61*(1), 98–109.

Ottenbros, I. B., Boerleider, R. Z., Jubitana, B., Roeleveld, N., & Scheepers, P. T. J. (2019). Knowledge and awareness of health effects related to the use of mercury in artisanal and small-scale gold mining in Suriname. *Environment International*, *122*, 142–150.

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. *Frontiers in Public Health*, *9*.

Huber, C., Nijssen, R., Mol, H., Antignac, J.P., Krauss, M., Brack, W., Wagner, K., Debrauwer, L., Vitale, C.M., Price, E.J., Klanova, J., Garlito Molina, B., Leon, N., Pardo, O., Fernández, S.F., Szigeti, T., Középesy, S., Šulc, L., Čupr, P., Mārtiņsone, I., Akülova, L., **Ottenbros, I.,** Vermeulen, R., Vlaanderen, J., Luijten, M., & Lommen, A. (2022). A large scale multi-laboratory suspect screening of pesticide metabolites in human biomonitoring: From tentative annotations to verified occurrences. *Environment International*, *168*, 107452.

Ottenbros, I., Lebret, E., Huber, C., Lommen, A., Antignac, J.-P., Čupr, P., Šulc, L., Mikeš, O., Szigeti, T., Középesy, S., Martinsone, I., Martinsone, Z., Akulova, L., Pardo, O., Fernández, S. F., Coscollá, C., Pedraza-Diaz, S., Krauss, M., Debrauwer, L., Wagner, K., Nijssen, R., Mol, H., Vitale, C.M., Klanova, J., Garlito Molina, B., León, N., Vermeulen, R., Luijten, M., & Vlaanderen, J. (2023). Assessment of exposure to pesticide mixtures in five European countries by a harmonized urinary suspect screening approach. *International Journal of Hygiene and Environmental Health*, *248*, 114105.

Šulc, L., Janoš, T., Figueiredo, D., **Ottenbros, I.**, Šenk, P., Mikeš, O., Huss, A., & Čupr, P. (2022). Pesticide exposure among Czech adults and children from the CELSPAC-SPECIMEn cohort: Urinary biomarker levels and associated health risks. *Environmental Research*, *214*, 114002.

Janoš, T., **Ottenbros, I.,** Bláhová, L., Šenk, P., Šulc, L., Pálešová, N., Sheardová, J., Vlaanderen, J., & Čupr, P. (2023). Effects of pesticide exposure on oxidative stress and DNA methylation urinary biomarkers in Czech adults and children from the CELSPAC-SPECIMEn cohort. *Environmental Research*, *222*, 115368.

Luijten, M., Vlaanderen, J., Kortenkamp, A., Antignac, J.-P., Barouki, R., Bil, W., van den Brand, A., den Braver-Sewradj, S., van Klaveren, J., Mengelers, M., **Ottenbros, I.**, Rantakokko, P., Kolossa-Gehring, M., & Lebret, E. (2023). Mixture risk assessment and human biomonitoring: Lessons learnt from HBM4EU. *International Journal of Hygiene and Environmental Health*, *249*, 114135.

Rodriguez Martin, L., **Ottenbros, I.**, Vogel, N., Kolossa-Gehring, M., Schmidt, P., Řiháčková, K., Juliá Molina, M., Varea-Jiménez, E., Govarts, E., Pedraza-Diaz, S., Lebret, E., Vlaanderen, J., & Luijten, M. (2023). Identification of Real-Life Mixtures Using Human Biomonitoring Data: A Proof of Concept Study. *Toxics*, *11*(3), 204.

Tarazona, J. V., Cattaneo, I., Niemann, L., Pedraza-Diaz, S., González-Caballero, M. C., de Alba-Gonzalez, M., Cañas, A., Dominguez-Morueco, N., Esteban-López, M., Castaño, A., Borges, T., Katsonouri, A., Makris, K. C., **Ottenbros, I.**, Mol, H., De Decker, A., Morrens, B., Berman, T., Barnett-Itzhaki, Z., Probst-Hensch, N., Fuhrimann, S., Snoj Tratnik, J., Horvat, M., Rambaud, L., Riou, M., Schoeters, G., Govarts, E., Kolossa-Gehring, M., Weber, T., Apel, P., Namorado, S., & Santonen, T. (2022). A Tiered Approach for Assessing Individual and Combined Risk of Pyrethroids Using Human Biomonitoring Data. *Toxics*, *10*(8), 451.

DANKWOORD

Wetenschappelijk onderzoek wordt, gelukkig maar, niet alleen gedaan en ik ben enorm dankbaar voor alle samenwerkingen die ik heb kunnen aangaan in de afgelopen jaren. Het was een avontuur met vele uitdagingen, maar ik kan met trots terugkijken nu ik dit boekje in mijn handen heb. Daarnaast heb deze PhD nooit tot een goed einde kunnen brengen zonder de nodige afleiding en support van familie en vrienden om mij heen.

Allereerst mijn promotoren en copromotor. Roel, bedankt voor je scherpe blik tijdens onze afspraken en je continue doortastende wetenschappelijke houding door steeds opnieuw de juiste vragen te stellen. Het hield mij scherp en heeft mij enorm laten ontwikkelen als wetenschapper. Bedankt ook voor de gesprekken toen ik zoekende was naar de rollen tussen de twee instituten binnen mijn project. Erik, het was mij een eer om je allerlaatste promovenda te zijn. Ik heb het geprobeerd te rekken met een zwangerschap, hersenschudding en pandemie tussendoor, maar het is nu toch echt afgerond. Je expertise in het vakgebied gecombineerd met een menselijke en toegankelijke houding waren van onschatbare waarde voor mijn motivatie en kwaliteit van het werk. Jelle, bedankt voor je support en kritische houding tijdens de analyses en het schrijven, het heeft mij geleerd om niet zomaar genoegen te nemen en door te gaan tot het optimale resultaat is bereikt. Vanaf het derde jaar spraken we elkaar regelmatiger en jouw deur stond altijd open voor alle vragen die ik had.

Dank ben ik daarna verschuldigd aan de vele coauteurs van mijn publicaties. Specifiek Eva, voor het samen opstellen van het FLEHS manuscript, bedankt voor de gastvrijheid bij VITO in Mol, het was een groot plezier om daar tijdelijk te verblijven. Priska, I also would like to thank you for our collaboration on the last manuscript of my thesis and the first of yours. Although we've never met in person, I really think we build a bond and cooperated without any issues with a clear and joyful division of tasks. All SPECIMEn WP15 colleagues: it was my pleasure coordinating the fieldwork for the entire study and the results we have established wouldn't have been possible without all your continuous efforts.

De verzameling van nieuwe gegevens was niet mogelijk geweest zonder de medewerking van alle ouders en hun kinderen binnen de SPECIMEn-NL studie. Bedankt dat de deur openstond en voor alle huisbezoeken die wij hebben mogen doen. Deze bezoeken kon ik onmogelijk zelf allemaal doen, bedankt voor jullie assistentie tijdens het veldwerk Paul S, Karin, Paul F en Daan.

Mirjam en Jan-Paul, bedankt voor jullie ondersteuning vanuit het RIVM tijdens mijn werk. Het was een hele fijne samenwerking met korte lijntjes. Irene, Leo en Rik, bedankt voor jullie luisterend oor toen het nodig was. Irene, het was mij ook een groot plezier om Equal-Life te

ondersteunen naast de laatste thesis loodjes. We zijn goed in lange gesprekken, maar zoals je zelf al schreef op mijn afscheidskaart: 'speechless'.

Het grootste gedeelte van mijn kantoortijd (wanneer het mogelijk was) spendeerde ik op het IRAS. De afdelingsuitjes (online en offline), de lunch en koffiemomenten waren een welkome afwisseling tussen het soms solitaire PhD bestaan. Het warme bad van 3.19b heeft ervoor gezorgd dat ik deze thesis daadwerkelijk tot een goed einde heb kunnen brengen. Collega-PhDs zijn in mijn ogen van cruciaal belang, op sociaal vlak, maar vooral ook omdat je allemaal in hetzelfde schuitje zit en frustraties en successen kunt delen. The 'old'-3.19b roommates: Calvin, Daniel, Jules, Marije, Anran; thanks for listening to my frustrations, all the dinners, drinks, lunches, coffees and barbecues. The 'new'-3.19b roommates: Wenxin, Esther, Haykanush, Bea and Jie; thanks for your office support during my writing phase. Warner, Myrna, bedankt dat jullie deur ook altijd openstond, de support tijdens het schrijven en de feestjes als afleiding. Kees, fijn om af en toe wat Thaise herinneringen op te halen en mijn PhD proces met je te delen, geniet de vrije tijd straks. Yujia and Youchen, thanks for joining during the writing retreat including all the nice food. Fleur, bedankt voor alle samen schrijf uurtjes en alle afleiding eromheen, jij bent er ook bijna.

Bedankt ook aan alle TNO collega's waardoor ik mij meteen thuis voelde aan de andere kant van de Uithof. Calvin, bedankt voor het overhalen om toch te solliciteren, het was een goede keus en ik ben blij dat we weer collega's zijn. Speciale dank naar Eelco, Carina en Wouter voor jullie begrip en steun tijdens het schrijfproces naast een nieuwe baan, ik ben blij om met jullie in de toekomst verder te mogen werken.

Constanza, Anna, Julia, Lieke and Emma, thanks for sharing drinks, rooms in Zeeland, PhD stories and much more. It is so much nicer to have a group of like-minded girls around when needed.

Sonia, Kiran, Vera, Kim, Evelyn, van BeeVee feestjes toen we 18 waren naar baby's. Wat is ons leven veranderd maar wat is het fijn om nog steeds zo veel met jullie te kunnen delen. Dank-julliewel voor de vriendschap en nodige afleiding tijdens dit hele promotieproces.

Iris, Lisa, van Suriname tot Utrecht tot Nijverdal, wat een plezier om jullie in mijn leven te hebben, te sporten, goede gesprekken te hebben en tegenwoordig een gemeenschappelijke deler te hebben met allemaal een kindje.

Anran, thanks for being a roommate, a friend, being always enthusiastic, introducing me to the Epi girls and for being my paranymph, this chapter in life can close with style with you on my side. Kim, van collega's aan het begin van mijn promotie bij het RIVM tot aan collega's bij het IRAS aan het einde van mijn promotie. Ik ben heel blij dat wij elkaar hebben leren kennen en vriendinnen zijn; je enthousiasme en altijd zin in een feestje zijn zeer aanstekelijk, dankjewel dat ook jij mijn paranimf wilt zijn.

Jan, Anne, bedankt voor jullie steun afgelopen jaren. Niet alleen de vele oppasuurtjes, maar ook het begrip en mij thuis laten voelen binnen de Ogink familie.

Papa, mama, ondanks dat promoveren aan de universiteit voor jullie in eerste instantie vrij onbekend terrein was, kon ik altijd rekenen op jullie onvoorwaardelijke steun. Dankjulliewel voor de motivatie om toch een bètastudie te kiezen, heeft het toch goed uitgepakt he. Wat een plezier om jullie dochter te zijn en jullie nu als opa en oma te zien.

Helena, mama worden van jou was niet alleen een enorme verandering in mijn leven maar zette ook mijn werk in perspectief. Blijf je vrolijke eigenwijze zelf.

Vincent, van twee-keer-bellen naar getrouwd met straks twee kindjes, wat een geluk dat ik jou aan mij zij mag hebben. Dankjewel voor het eindeloze vertrouwen in mij, zelfs als ik het zelf soms even kwijt was. Op naar het volgende hoofdstuk samen, ik hou van je.

Dankjewel voor jou als lezer dat je tot zover bent gekomen met het lezen van deze thesis.

CURRICULUM VITAE

Ilse Ottenbros was born on September 26, 1991, in Alkmaar, the Netherlands. In 2009, she started her bachelor's in Biology at Radboud University Nijmegen, the Netherlands. She subsequently followed a master's in Toxicology and Environmental Health at the Institute of Risk Assessment Sciences (IRAS), Utrecht University, the Netherlands. As part of her master's she did a 6-month internship on indoor air pollution in Lampang, Thailand. Her major 9-month internship focused on exposure determinants of nanomaterials at TNO Zeist, the Netherlands. In 2015, Ilse commenced her career as a junior researcher at Radboud University Medical Center Nijmegen. Over a span of two years, she was engaged in multiple projects, including a study on mercury exposure among small-scale gold miners in Suriname. In December 2017, she started her PhD research at IRAS and the Dutch Institute for Public Health and the Environment (RIVM) under the European project HBM4EU, investigating mixtures of exposure measured by human biomonitoring. The findings of her research are presented in this thesis. While finalizing her PhD, she worked part-time for 9 months at RIVM as a project manager for the European project Equal-Life, contributing to project reporting and review activities. In May 2023, Ilse joined TNO Utrecht as a scientist specializing in occupational exposures. In this current role she is involved in developing and enhancing methods for aggregated exposures.

Universiteit Utrecht

Doctoral thesis, Utrecht University