HBM4EU science and policy for a healthy future

HORIZON2020 Programme
Contract No. 733032 HBM4EU

Case study reports on mixture health effects

Deliverable Report

D15.5

WP15 - Mixtures, HBM and human health risks

Deadline: April 2021

Upload by Coordinator: 29 June 2021

Entity	Name of person responsible	Short name institution	Date [Received]
Coordinator	Marike Kolossa-Gehring	UBA	09/06/2021
Grant Signatory	Ovnair Sepai	DH	

Entity	Name of person responsible	Short name institution	Date [Approved]
Pillar Leader	Robert Barouki	INSERM	15/05/2021
Work Package Leader	Mirjam Luijten	RIVM	04/05/2021
Task leader	Andreas Kortenkamp	BRUNEL	30/04/2021

Responsible author	Andreas Kortenkamp
Short name of institution	BRUNEL
Co-authors	Anne Marie Vinggaard, Marcel Mengelers, Remy Slama, Maria João Silva, Henriqueta Louro, Susana Viegas, Ana Tavares, Thomas Goen, Sibylle Ermler, Olwenn Martin, Annick van den Brand, Gerda van Donkersgoed, Tiina Santonen, Juha Tuovila, Yanying Ma, Anna Kjerstine Rosenmai, Andrea Rodríguez-Carrillo, Vicente Mustieles, Camilla Taxvig, Lena Reiber, Anja Kiesow, Madlen David, Nathalie Michelle Löbl, Mariana F. Fernández, Mousumi Chatterjee, Lydiane Agier, Xavier Basagna, Mirjam Luijten

Table of contents

1 Authors and acknowledgements

Lead author

Andreas Kortenkamp, Brunel University London

Contributors

Anne Marie Vinggaard, Marcel Mengelers, Rémy Slama, Maria João Silva, Henriqueta Louro, Susana Viegas, Ana Tavares, Thomas Goen, Sibylle Ermler, Olwenn Martin, Annick van den Brand, Gerda van Donkersgoed, Tiina Santonen, Juha Tuovila, Yanying Ma, Anna Kjerstine Rosenmai, Andrea Rodríguez-Carrillo, Vicente Mustieles, Camilla Taxvig, Lena Reiber, Anja Kiesow, Madlen David, Nathalie Michelle Löbl, Mariana F. Fernández, Mousumi Chatterjee, Lydiane Agier, Xavier Basagna, Mirjam Luijten

2 Glossary

ADI	Acceptable Daily Intake
AF	Assessment Factor
AOP	Adverse Outcome Pathway
AR	Androgen Receptor
BBzP	Butyl benzylphthalate
BMD	Benchmark Dose
CA	Concentration Addition
CAG	Cumulative Assessment Group
DA	Dose Addition
DAP	Dialkylphosphate
DBP	Dibutyl phthalate
DEHP	Diethyl hexylphthalate
DIBP	Di-isobutylphthalate
DINP	Di-isononylphthalate
DNT	Developmental Neurotoxicants or Developmental Neurotoxicity
ECx	Effect Concentration at Response x
ECHA	European Chemicals Agency
EDx	Effect Dose at Response x
EFSA	European Food Safety Authority
HBGV	Health-based Guidance Value(s)
HED	Human Equivalent Dose
HI	Hazard Index
IA	Independent Action
LOEL	Lowest Observed Effect Level
LOAEL	Lowest Observed Adverse Effect Level
MCR	Maximum Cumulative Ratio
MoA	Mode of Action
MOE	Margin of Exposure
MRA	Mixture Risk Assessment
NOEL	No Observed Effect Level
NOAEC	No Observed Adverse Effect Concentration
NOAEL	No Observed Adverse Effect Level

PAH	Polycyclic Aromatic Hydrocarbon
PBDE	Poly -Brominated Diphenyl Ethers
PNEC	Predicted No Effect Concentration
POD	Point of Departure
PODI	Point-of-Departure Index
RfD	Reference dose, for specific common toxicities
RPF	Relative Potency Factor
RQ	Risk Quotient
RV	Regulatory Value
TDI	Tolerable Daily Intake
TEF	Toxicity Equivalency Factor
TRV	Toxicological Reference Value
TTC	Threshold of Toxicological Concern
TUS	Toxic Unit Summation
TWI	Tolerable Weekly Intake
UF	Uncertainty Factor
USEPA	US Environmental Protection Agency

3 Abstract/Summary

The aim of our work in WP15 Task 15.3 was to evaluate a proof-of-concept for the identification of mixture health effects. To realise this aim we conducted selected case studies on human health endpoints of concern. The following human health endpoints were chosen for these case studies:

- Endocrine disruption with a focus on disruptions of male reproductive health
- Developmental neurotoxicity
- Cancers, including e.g. lung cancer
- Nephrotoxicity

By focusing on PBDEs, chemicals with anti-androgenic properties, heavy metals with nephrotoxicity (lead, mercury, cadmium) and occupational carcinogens (chromium (VI), nickel and PAHs) we aligned the pollutants and groups of chemicals in the case studies to the priority pollutants chosen for HBM4EU. As an issue relevant to all health endpoints and all chemicals, we addressed exposure misclassification resulting from the biomonitoring of multiple chemicals in spot samples. We developed a strategy for minimizing bias towards the null (false negatives) based on collecting repeated biospecimens per subject, especially for non-persistent chemicals. Pooling within-subject biospecimens is an efficient way of limiting bias without increasing assay costs.

We built up a consistent assessment framework for all case studies, by elaborating criteria for the grouping of chemicals to be subjected to mixture risk analyses, based on Adverse Outcome Pathway considerations if possible. We utilized mixture risk assessment methods that apply the principle of dose addition by summing up risk quotients as in the Hazard Index (HI) or the Point-of-Departure Index (PODI). We adopted tiering rules, for both the exposure assessment and hazard assessment arms of the mixture risk assessment. These rules are based on the principle of advancing mixture risk assessment in a stepwise manner with the aim of minimizing bias by approximating the scientific principles of experimental mixture assessments based on Dose Addition or Independent Action. We introduced the concept of drivers of mixture effects.

The number of chemicals investigated in the case studies ranged from 3 (heavy metals and nephrotoxicity, carcinogenicity from chromium (VI), nickel and PAHs) to 61 (disruption of male reproductive health).

All case studies identified exceedances of combined acceptable approaches, sometimes by quite large margins, suggestive of mixture risks that do not become apparent with the standard focus on chemical-by-chemical risk assessments.

However, many of the single chemicals included in the assessments already exceeded their individual acceptable exposures. But even if all single chemicals were compliant with their individual exposure limits, mixture risks would still be discernable. Often, a small number of chemicals contributed disproportionately to the sum of risk quotients, close to Pareto's 20:80 rule (20 % of the components produce 80 % of the sum of risk quotients). The nature of these "drivers" of mixture risks varied with the health endpoint considered in mixture risk assessments, but some drivers were common to several health endpoints (e.g. heavy metals for nephrotoxicity and developmental neurotoxicity). Some chemicals that have attracted a great deal of public and regulatory attention (parabens, phthalates, certain pesticides) were not among drivers of mixture risks.

Most case studies had to be based on exposure assessments for single chemicals conducted separately in different study cohorts. It was not possible to directly derive which chemicals, and at

what levels, occur together. This is a shortcoming that can only be overcome by developing dedicated monitoring strategies for mixture exposures.

We propose that cost-effective Human Biomonitoring strategies aimed at capturing combined exposures should be developed by adopting multi-chemical analyses that focus on measuring identified drivers of mixture risks.

Our case studies show that the traditional focus on single pollutants has led to a fragmented view of the extent of the mixtures issue. Chemicals from different regulatory domains (pesticides, industrial chemicals, pharmaceuticals) work together to produce considerable exceedances of combined acceptable exposures. The regulatory instruments necessary to deal with this issue are not currently available. Populations in Europe are insufficiently protected from the risks associated with unintentional mixtures.

4 Introduction

Humans and wildlife are continuously exposed to multiple chemicals from different sources and via different routes, both simultaneously and in sequence. Scientific evidence for heightened toxicity from such mixtures is mounting, yet risk assessment approaches viable in regulatory practice are not available and need to be developed, evaluated and implemented. As much as possible, Human Biomonitoring should integrate knowledge about mixture effects and begin to develop strategies that can cope with the reality of multi-chemical, multi-pathway exposures of humans.

This Deliverable describes the outcome of several mixture risk assessment case studies that were developed to describe and assess mixture health effects.

All case studies dealt with the question: Do predicted mixture risks from combined chemical exposures exceed the levels regarded as acceptable for humans? They addressed the following specific aspects:

- Are there human health endpoints for which acceptable combined exposures are exceeded?
- Are there pollutants that contribute most to defined health endpoints of interest and are therefore drivers of mixture risks?
- Is it possible to develop priorities for Human Biomonitoring strategies that could capture combined exposures to those pollutants?

Our intention was to focus on several methodological issues, especially:

- Identify methods for the prediction of mixture effects that can be used consistently for human risk assessments and can inform biomonitoring strategies in terms of chemicals to be monitored together,
- Define properties of data required as input for mixture effect predictions.

Accordingly, this Deliverable is structured as follows. We will first provide a brief overview of the general mixture risk assessment concepts and principles that we utilized in the case studies.

We then provide detailed reports of each case study, which are structured as follows:

Grouping criteria – a description of the criteria used to select and group together the chemicals to be subjected to mixture risk assessment.

Exposure assessment – compilations of relevant human exposure data, with a focus on Europe and EU countries. As much as possible, this relies on Human Biomonitoring data, but other exposure assessment methodologies such as those based on occurrence in food and food consumption are also considered.

Hazard assessment – compilations of health-based guidance values or reference doses for the health endpoints considered in each case study. Where relevant guidance values or reference doses for the health endpoints in question were not available, efforts were made to derive reference doses for MRA.

Mixture risk assessment – an integration of exposure and hazard assessments by using established MRA methods, followed by uncertainty considerations.

Conclusions and lessons learnt – a brief discussion of the extent of the mixtures issue for each case study, with a focus on implications for Human Biomonitoring.

5 Mixture risk assessment (MRA): General considerations, approaches, and methods

5.1 General considerations on combined toxicity and its assessment

In an earlier Deliverable (D15.4) we detailed the fundamental approaches for assessing experimental mixture data, dose addition (DA) and independent action (IA), together with their assumptions and applicability domains. DA is traditionally allied with combinations of substances that exert their toxicity through similar modes of action (MoA); IA is reserved for combinations composed of chemicals with a diversity of MoAs.

5.2 Dose addition as a default assumption

In the light of the practical difficulties encountered when using considerations of similar or dissimilar action as the starting point for MRA (lack of mechanistic data, ambiguities associated with key terms such as mode of action, mechanism of action), a dichotomous approach to MRA that seeks to establish similarity or dissimilarity before the assessment can commence is increasingly considered as problematic in cases where the MRA is based on complex hazard endpoints covering several MoAs. There is a consensus that MRAs should start from the default assumption of DA for all mixture components, regardless of MoAs. If this indicates a significant risk, refined MoA-based assessments may be conducted where the necessary data are available. Alternatively, precautionary measures may be taken. This way of thinking has guided ecotoxicological MRAs for quite some time and is now also gaining increased acceptance in the human arena. This development opens the way for using consistent and coherent approaches across the disciplinary borders.

5.3 Data requirements and the distinction between scientific concepts of additivity and pragmatic simplifications for regulatory MRAs

The data requirements for DA-based MRAs are easier to meet than those for IA-based MRAs. Proper application of IA requires knowledge about the slope of dose response functions *F* of the individual mixture components. In contrast, under the assumption of DA, the prediction of effect concentrations (or effect doses) of mixtures only requires that equivalent effect concentrations of single substances are known. Furthermore, the algorithm for DA can be directly transformed into a risk quotient for mixtures. The algebraic equivalent of the DA formula usually used for this purpose is the so-called toxic unit summation (TUS) as explained in more detail below.

DA and TUS require ECx or EDx values that refer to the same endpoint in the same species under comparable conditions. In a regulatory context, even this relatively simple data requirement may be impossible to meet. As a result, several pragmatic approaches have been derived from the original DA concept. Some prominent examples are the Hazard Index (HI), Point of Departure Index (PODI), Relative Potency Factors (RPF) and the concept of Toxic Equivalency Factors (TEF). All these MRA methods are simplifications of the DA concept. As a common feature, they make use of the DA formula as a calculation rule but use input data that deviate more or less from the strict requirements of the original concept or make additional simplifying assumptions about the individual dose response curves. As detailed in D15.4, the two most widely used MRA methods are the **Hazard Index (HI)** and the **Point-of-Departure Index (PODI)**. Both these methods will be considered in terms of their applicability in the case studies below.

5.4 The principal assumptions and simplifications behind mixture risk assessment methods

In common with the MRA approaches applied in regulatory practice, the practical application of mixture risk assessment approaches is based on a number of assumptions and simplifications which must be made explicit. These assumptions were detailed in D15.4, but are briefly repeated here:

- **The possibility of synergisms or antagonisms is disregarded**. This assumption is the direct consequence of the fact that the degree of synergism or antagonism cannot be predicted quantitatively based on the toxicity of the mixture components. All mixture effect prediction methods and accordingly, all MRA methods, assume additivity. Considering that the likelihood of synergisms is relatively small when multiple toxicants are present at low regulatory acceptable levels the disregard for toxic interactions may be considered as sufficiently protective.
- **Simultaneous exposure to multiple chemicals is assumed.** In numerous settings encountered by human populations there is simultaneous exposure to multiple chemicals. For example, there is consumption of single food items that contain multiple chemicals and, even when food items are consumed sequentially, the subsequent exposure of body tissues to the chemicals contained within the items may be simultaneous. Strictly sequential exposures are also a reality, but the risk assessment methods available for MRA are not applicable to sequential exposure to multiple chemicals. In theory and concept, the development of methods for the assessment of sequential exposures is still in its infancy (Altenburger and Greco 2009).
- **Potency estimates for mixture components may be derived from different endpoints.** Application of DA requires the use of potency estimates for the same adverse outcome as input values. However, such input values are often not available because chemical safety testing is geared towards identifying critical toxic effects which can be used for the establishment of reference doses (ADI, TDI) and health-based guidance values (HBGV). In practice, this means that toxicity information for chemicals that occur together in mixtures often derives from disparate endpoints. To enable assessments of cumulative risks, the demand for potency estimates for the same endpoints is therefore relaxed, especially for simplified analyses at lower tiers of MRA.
- **It is assumed that the potency estimates entered into MRA methods (e.g. ADIs, NOAELs, benchmark doses) describe doses associated with the same effect magnitude.** The equations for DA are based on single chemical effect doses for identical effect magnitudes. When applied to the PODs that enter the mathematical expressions used in MRA methods such as HI or PODI, this means that all PODs should describe effect doses for the same effect levels. In practice however, this demand cannot always be met, except in the case of benchmark doses which are defined in relation to the same effect levels. In human toxicology, the effect magnitude associated with NOAELs are normally not known. To make cumulative risk assessment methods workable despite these knowledge gaps, ADIs and similar metrics are taken as if they described effect doses for the same effect magnitude.
- **Potency estimates can be derived from different tests, performed under different conditions.** In the interest of consistency, the evaluation of experimental mixture effects by using the concept of DA should utilize effect data for all the mixture components that were gathered under the same experimental conditions, with the same test organisms. If this

condition is not fulfilled, a bias may be introduced into the analysis, leading to erroneous determinations of mixture effects in terms of additivity, synergy or antagonism. MRA however usually rely on data that were produced in the context of single chemical testing, under widely varying experimental conditions, even when the same organisms were used, so that the demand of consistency of data cannot be realized in practice. To allow continuation of MRA, this demand therefore has to be relaxed.

- **Data on exposures and potency must be recorded by using the same dose metric.** To allow utilization of the formula for HI or PODI, for all mixture components both exposure and toxicity data must be expressed in the same unit and must refer to the same route or matrix, either in terms of an external concentration in an environmental medium such as water or food, or in terms of an amount taken up per unit of time and biomass via a defined route such as oral, dermal, inhalation etc., or in terms of an aggregated dose via different routes, or as an internal concentration in defined tissue or body fluids such fat or blood, or a total body burden.

5.5 Principles of a stepwise (tiered) analysis in mixture risk assessment and development of a common assessment framework

Tiered approaches to MRA can avoid unnecessary expenditure of resources by offering the possibility of discontinuing the analysis based on crude and simple assumptions about exposures and hazards when cumulative exposures are judged to be tolerable or acceptable. In this way, lengthy, but largely unproductive efforts of refining the analysis can be avoided.

As detailed in D15.4, we developed a framework and workflow for all case studies which includes the following:

- Clear **rules for tiering** of the exposure and hazard assessment arms of the framework
- Clear **decision rules** that define when further refinement of the analysis is to be conducted, and when the analysis can be stopped
- Consistent **application of assessment factors** in each of the tiers
- A concept that will remove in a step-wise fashion the distortions introduced through the use of different assessment factors in regulatory values for pollutants, with the aim of **approaching the scientific principles of Dose Addition** in higher tiers (points of departures for the same adverse outcome, representative of similar effect magnitudes, e.g. benchmark doses for 5 % effects and based on the same test species)
- A consistent approach to **bridging toxicity data gaps** (e.g. by using the Thresholds of Toxicological Concern (TTC) concept, or by read-across methods)
- Clear **grouping criteria**, based on AOP considerations where possible and on common adverse outcomes

In the **hazard assessment arm** of the mixture risk assessment, three Tiers can be used. Data quality and availability permitting, the analysis can skip lower Tiers and commence at higher Tiers:

Tier 1 will check compliance with Health-based Guidance Values (HBGV) that have been defined for the chemicals to be subjected to mixture risk assessment. It is recognized that these HBGV can be based on different critical toxicities; thus, at this stage the assessment mixes different health endpoints and assessment factors. The Hazard Index (HI) is the appropriate mixture risk

assessment method for this Tier of the analysis. If non-compliance with the sum of risk quotients based on these HBGV is detected, the analysis is refined by moving to Tier 2.

Tier 2 improves the certainty of the assessment by removing the distortions that were introduced in Tier 1 through the mixing of HBGV based on different health endpoints, with different assessment factors. As much as possible, dose estimates for common adverse outcomes are used, here termed Reference doses (RfD) for the specific endpoint under consideration. To deal with distortions arising from the use of different assessment factors behind these RfDs there are several options: Either the same assessment factors are employed for all chemicals, or, if this is not possible, the risk assessment will be conducted on the basis of points-of-departures.

Tier 3 introduces further refinements, if Tier 2 reveals unacceptable combined exposures. These refinements can focus on using hazard data on chemicals with the same mode of action and potency values associated with identical effect magnitudes (as in benchmark modelling).

It is not necessary to carry the analysis through Tier 1 in all cases. If data of sufficient quality are available, Tier 1 can be skipped and the analysis begins with Tier 2.

The **exposure assessment arm** of the procedure can also be refined in a stepwise manner, as follows:

Tier 1 will often use exposure data derived from census methods (combining e.g. food consumption data with levels of pollutants in food) for the population of relevance.

Tier 2 introduces refinements by utilizing Human Biomonitoring data for the pollutants of interest. In this way aggregated exposure independent of exposure source can be included in the MRA.

Tier 3 uses Human Biomonitoring data for multiple pollutants measured in the same individuals.

6 Case study 1: A mixture risk assessment for male reproductive health with a focus on semen quality

Andreas Kortenkamp, Sibylle Ermler, Olwenn Martin

Brunel University London

6.1 Introduction

In many countries there have been increases in testicular non-descent (Main *et al.*, 2010) and penile malformations, hypospadias. The incidence of testicular germ cell cancers has risen steadily in Caucasian men (Chia *et al.*, 2010), while semen quality continues to decline in Western countries (Levine *et al.*, 2017).

Together, these disorders form a syndrome termed testicular dysgenesis syndrome (TDS), thought to arise from insufficient androgen action in fetal life (Skakkebaek *et al.*, 2016). The TDS hypothesis predicts that exposures to chemicals capable of disrupting androgen signaling in fetal life, so-called anti-androgens, are an etiological factor. Certain phthalates, bisphenols and anti-androgenic pesticides play a role, but the overall contours of chemical exposures contributing to declines in male reproductive health have not yet come into view.

In this case study, we assess the role of multiple chemical exposures in the decline of male reproductive health. The primary aim of our work is to assess whether combined acceptable exposures of chemicals known to compromise male reproductive health are exceeded. Secondly, we wish to identify those chemicals that make the strongest contribution to combined health risks.

We focus on a risk assessment for the general population. Considering that the period of heightened sensitivity for risks to male reproductive health is in fetal life, we have paid particular attention to exposures in expectant mothers. Our assessment applies to exposures experienced in Europe.

6.2 Grouping criteria and cumulative assessment group

Several experimental studies have shown that multiple phthalates act together to produce adverse reproductive and developmental effects through a hormonal MoA (Howdeshell, Hotchkiss and Gray, 2017; Kortenkamp, 2020). These combinations show stronger effects than any single phthalate in the mixture on its own. Accordingly, MRAs involving phthalates have been conducted in the context of assessing chemical exposures relevant to male reproductive health (Kortenkamp and Koch, 2020), but MRAs that also include other anti-androgenic chemicals are few and far between.

However, experimental evidence has accumulated to show that multiple anti-androgenic chemicals with diverse chemical characteristics, and not only phthalates, can act together to disrupt androgen signaling and male sexual differentiation (Kortenkamp 2020). Already in 2008, the US National Academy of Sciences recommended that human mixture risk assessments should not stop with phthalates but should include a multitude of other anti-androgenic chemicals. That this recommendation has not been taken up more widely can be attributed to a lack of clarity about which chemicals, beyond phthalates, should be grouped together for the purpose of a comprehensive MRA.

With the aim of providing criteria for the inclusion of additional chemicals in mixture risks assessments for male reproductive health, we have examined the mechanisms of action of various chemicals capable of disrupting male sexual differentiation. We constructed an Adverse Outcome

Pathway (AOP) network for disruption of the male reproductive system (Kortenkamp 2020) and used this network to identify pathways that converge at critical nodal points to produce down-stream adverse effects (Figure 6.1). The AOP network enabled us to predict combinations of chemicals with different mechanisms of action expected to lead to cumulative effects of disrupted male sexual development. We then mapped these predictions against evidence from experimental mixture studies with relevant combinations of chemicals. This analysis revealed that phthalates work together with androgen receptor (AR) antagonists, chemicals capable of disrupting steroid synthesis, InsL3 production, prostaglandin signaling and co-planar polychlorinated dibenzo-dioxins to disrupt male sexual differentiation after gestational exposures. Thus, **cumulative assessment groups** (CAG) for male reproductive health risks should not only include phthalates but also comprise pesticides such as vinclozolin, prochloraz, procymidone, linuron, the pain-killers paracetamol, aspirin and ibuprofen, pharmaceuticals such as finasteride, ketoconazole, and the lipid-lowering drug simvastin, poly-chlorinated dibenzo-dioxins and other dioxin-like pollutants and phenolics such as bisphenol A and butylparaben.

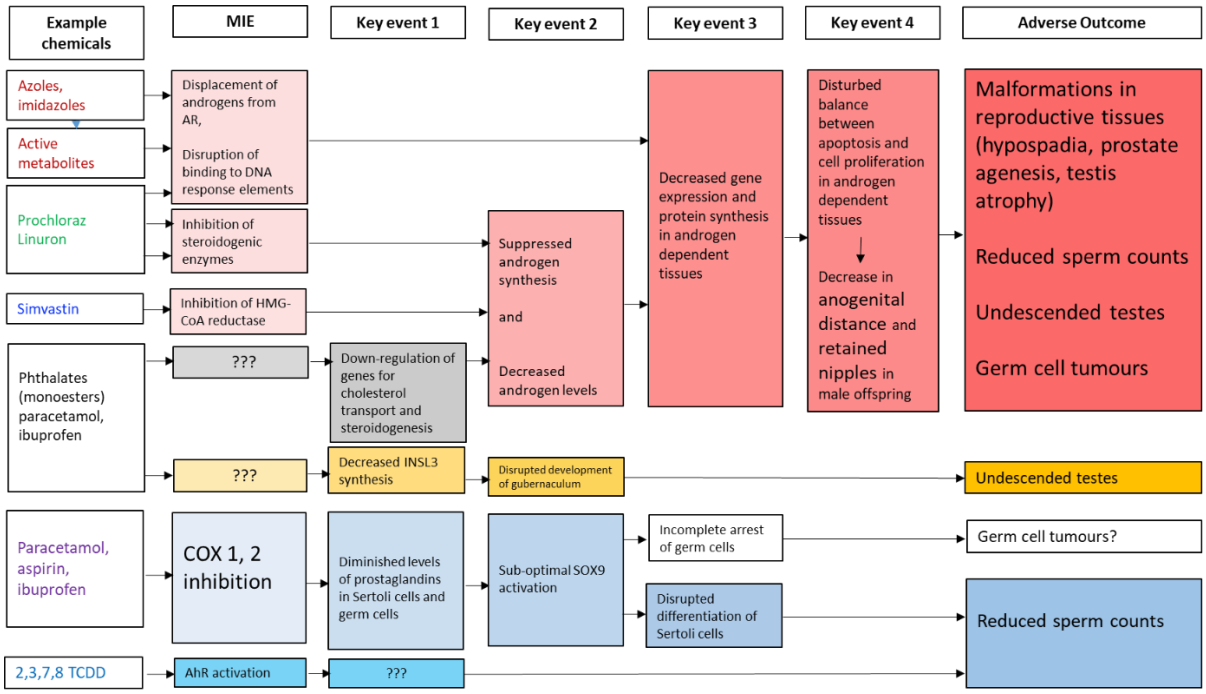


Figure 6.1: AOP network for the induction of male reproductive malformations (from Kortenkamp 2020)

In humans, disruption of the signaling processes important for male sexual development materializes as a constellation of effects that includes changes in anogenital distance, non-descending testes, hypospadias, increased testis cancer risks and poor semen quality, the testicular dysgenesis syndrome (TDS). This effect spectrum has parallels with the changes observed in experimental animals. Although the elements of the spectrum are strongly related to each other - they are the result of convergences of molecular signaling chains at certain critical nodes (see Figure 6.1) - there are differences in the shape and position of dose-response curves for each of the component effects. For example, increases in retained nipples seen with an AR-antagonist such as vinclozolin usually become apparent at lower doses than changes in anogenital distance or malformations of the penis (Christiansen *et al.*, 2009). As a result, effect doses derived for the same chemical for different component effects of the syndrome show slight differences. This is of consequence for establishing risk quotients needed for the MRA methods of HI and

PODI. Potency comparisons of members of the cumulative assessment group will be distorted to a certain degree if the corresponding effect doses do not relate to the same effect of the syndrome. To avoid such inconsistencies, we have focused our MRA for male reproductive effects on **declines in semen quality**.

Accordingly, we used the following **criteria** for the selection of members of the corresponding cumulative assessment group (CAG):

- Evidence of induction of declines in semen quality (sperm numbers, concentration, morphology, motility) either from human epidemiological studies or experimental studies in mammals, together with
- Evidence of a hormonal MoA (AR antagonism, suppression of testosterone synthesis, inhibition of enzymes involved in androgen synthesis, suppression of InsL3 production, disruption of prostaglandin synthesis)

We gave preference to studies demonstrating declines in semen quality after gestational exposures, either in epidemiological studies or in animal studies. Where data on declining semen quality in animal studies was not available, we used evidence from human studies. In cases where data from gestational exposures in animal experiments were missing, but data of declines in semen quality was available from studies in adult animals, we only included substances that had a clear hormonal MoA.

Accordingly, we selected the following 29 chemicals or chemical groups for a MRA relevant to male reproductive health:

- **AR antagonists**: Bisphenols A, F, S; n-butyl paraben; polybrominated diphenyl ethers BDE 99, 100, 183, 209; PCB 118, 126; chlorpyrifos, vinclozolin, procymidone, fenitrothione
- **Disruption of prostaglandin signaling and InsL3 production**: Paracetamol
- **Suppression of testosterone synthesis**: Phthalates DEHP, DBP, BBzP, DINP; acrylamide
- **Inhibition of steroidogenic enzymes**: linuron
- **AhR activation**: polychlorinated dibenzodioxins and -furans (PCDD/F, 17 congeners), PCB 118, 126, 169

6.3 Hazard characterization

For all the substances included in the CAG, we collated quantitative dose estimates for declines in semen quality, **Reference Doses for male reproductive health,** from here on referred to as RfD. Declines in semen quality do not always equate with the critical toxicity used to derive Health-based Guidance Values (HBGV) for single chemical risk assessments (i.e. the toxicity that appears first at doses > 0). In many cases, declines in semen quality become evident at doses higher than those associated with the critical toxicity. Unless declines in semen quality represent the critical toxicity of a substance, as is the case e.g. for PCDD/F, **RfDs used for this case study are therefore not applicable to single chemical risk assessments**.

As much as possible, we retrieved RfD from existing evaluations of competent authorities. Often, however, it was necessary to conduct separate reviews to derive the respective RfDs de novo. Not only did this require quantitative analyses of published data, it also necessitated examinations of the strength of evidence and study quality.

We applied the principles of a tiered assessment which we defined in the previous Deliverable D 15.4. Because we selected RfDs related to a specific common effect, i.e. declines in semen quality,

we were able to skip Tier 1 of the workflow and to commence the assessment in Tier 2 of the hazard assessment arm of the MRA.

Table 6.1: RfDs for declines in semen quality and related toxicity

Chemical	Reference dose RfD	Dose unit	Endpoint for RfD	Derived from animal studies?	Derived from human studies?	Comments
Bisphenol A	0.01	µg/kg d	semen quality	X		supporting evidence from epidemiology
Bisphenol F	0.014	µg/kg d	semen quality	X		
Bisphenol S	0.01	µg/kg d	semen quality	X		
n-butylparabe	30	µg/kg d	semen quality	X		
BDE 28	150	ng/kg d				derived by read-across from BDE 47
BDE 47	150	ng/kg d	semen quality	X		
BDE 99	2.9	ng/kg d	semen quality	X		
BDE 100	2.9	ng/kg d				derived by read-across from BDE 99
BDE 154	2.9	ng/kg d				derived by read-across from BDE 99
BDE 153	2.9	ng/kg d				derived by read-across from BDE 99
BDE 183	1000000	ng/kg d				derived by read-across from BDE 209
BDE 209	1000000	ng/kg d	semen quality	X		
PCDD/F	0.28	pg/kg d	semen quality		X	from EFSA (2018) TWI
PCB 118	2900	pg/kg d	semen quality	X		supporting evidence from epidemiology
PCB 126	73	pg/kg d	semen quality	X		supporting evidence from epidemiology
PCB 169	5330	pg/kg d	semen quality	X		supporting evidence from epidemiology
DEHP	10	µg/kg d	dysgenesis of genitalia	X		supporting evidence from epidemiology
DBP	6.7	µg/kg d	spermatocyte development	X		supporting evidence from epidemiology
BBzP	10	µg/kg d	T suppression	X		supporting evidence from epidemiology
DINP	59	µg/kg d	T suppression	X		
DIBP	100	µg/kg d	T suppression	X		
Acrylamide	8.3	µg/kg d	semen quality	X		
Chlorpyrifos	10	µg/kg d	semen quality	X		
Vinclozolin	50	µg/kg d	retained nipples	X		
Procymidone	100	µg/kg d	retained nipples	X		
Linuron	100	µg/kg d	retained nipples	X		
Prochloraz	160	µg/kg d	retained nipples	X		
Fenitrothione	200	µg/kg d	AGD changes	X		
Paracetamol	1	mg/kg d	semen quality	X		

The RfD values that we derived de novo did not rely on the use of assessment factors larger than those normally used for LOAEL to NOAEL extrapolations and for inter-species extrapolations. For substances that accumulate in adipose tissue we adopted body burden considerations and estimated equivalent human daily intakes (EHDI) associated with the body burden at the point of departure in animal experiments. In line with established practice, in such cases we used an inter-species assessment factor of 2.5.

An overview of the RfDs used in this case study is shown in Table 6.1.

6.3.1 Bisphenol A

EFSA (2015a) derived a HBGV for bisphenol A based on changes in kidney weight observed in rodents. After publication of the EFSA document, new evidence has come to light showing that bisphenol A produces declines in semen quality in animal studies after exposure during gestation. Bisphenol A exposures have also been associated with declines in semen quality in adult men.

We conducted a systematic review (to be published elsewhere) to retrieve relevant data on declines in semen quality. After exclusion of studies where bisphenol A was administered to adult

animals, or by subcutaneous injection, we identified 26 studies in rats and mice eligible for data extraction. We rated the quality of these studies and excluded 8 where there were concerns about reliability due the use of soy-containing diets or lack of effects with positive controls. We further excluded studies that used fewer than 2 doses of bisphenol A, as this would complicate dose-response analyses. This left 8 mouse studies (of which 1 did not find any effects of bisphenol A on semen quality) and 8 rat studies (of which 4 did not discover relevant effects). We selected the studies that reported the lowest LOAELs or NOAELs (Salian *et al.*, 2009; Hass *et al.*, 2016; Shi *et al.*, 2018, 2019; Ullah *et al.*, 2019) and, where necessary, extrapolated LOAELS to NOAELs by applying an assessment factor of 3. The resulting NOAELs were converted to Human Equivalent Doses (HED) by using the procedures detailed in EFSA (2015). HED were divided by 25 (2.5 for inter-species differences, and 10 for intra-species differences) to derive RfD. This procedure yielded RfDs ranging from 0.0004 µg/kg d (Shi et al. 2018, 2019) to 0.24 µg/kg d (Hass et al. 2016), with Salian et al. (2009) and Ullah et al. (2019) occupying the middle ground (0.014 and 0.011 µg/kg d, respectively).

There is evidence from multiple epidemiological studies that bisphenol A is associated with declines in semen quality after exposures in adult life.

We adopted 0.01 µg/kg d as RfD for bisphenol A as this value is near the midpoint of the above range of possible RfDs.

6.3.2 Bisphenol F, S

We followed the evaluation of male reproductive health animal studies conducted in HBM4EU Deliverable 5.9, Derivation of BPS and BPF HBM Guidance Values.

The only study of bisphenol F effects on spermatogenesis after gestational exposure is by Ullah et al. (2019) who observed alterations of spermatogenesis and decreases in sperm motility at 1.5 µg/kg d in rats. Conversion of this LOAEL to a NOAEL yields 0.5 µg/kg d. Assuming toxicokinetics similar to bisphenol A, this leads to a HED of 0.014 µg/kg d as RfD for bisphenol F.

For bisphenol S, the studies by Ullah et al. (2019) produced the same LOAELs as for bisphenol F. Shi et al. (2018) observed a three-fold lower LOAEL of 0.5 µg/kg d. We chose 0.01 µg/kg d as RfD for bisphenol S.

Epidemiological studies of associations between bisphenol F, S exposures and semen quality in men could not be located.

6.3.3 n-Butylparaben

Kang *et al.*, (2002) observed decreased sperm numbers and motility in the offspring of rats dosed during gestation with 100 and 200 mg/kg d n-butylparaben. Decreased sperm counts were also reported by Zhang *et al.*, (2016) who administered 400 and 1000 mg/kg d during gestation. NOAELs cannot be established from these studies.

Boberg *et al.*, (2016) administered a wider dose range to pregnant rats and observed reduced sperm counts at all tested doses (> 10 mg/kg d). We estimate that 10 mg/kg d is a LOAEL and extrapolated a NOAEL of 3 mg/kg d by application of an assessment factor of 3. With the standard assessment factor of 100 for animal to human extrapolation, this yields a RfD of 30 µg/kg d.

This value is slightly larger than the 20 µg/kg d derived by SCCS (2005), based on a NOAEL for estrogenic effects of n-butylparaben in rats.

To our knowledge, there are no epidemiological studies that show statistically significant associations between n-butyl paraben exposure and declining semen quality.

6.3.4 Poly-brominated diphenyl ethers BDE 99, 100, 183 and 209

Based on the systematic review of male reproductive effects of PBDEs by Zhang *et al.*, (2020), we conducted a systematic review for the time period after publication of this paper and located an additional 11 records which, together with the 12 studies from Zhang et al. (2020) gave a total of 23 studies on semen quality in animals eligible for data extraction. We evaluated the reliability of the 23 retrieved studies in terms of the purity of the dosed BDE congeners (to exclude issues of contamination with PCDD/F). We also considered blinding, appropriate numbers of animals and successful demonstration of effects with positive control substances.

Based on the studies we rated as being of sufficient quality, we identified the following points of departure (NOAELs) and RfDs:

BDE 47: The studies by Zhang *et al.*, (2013) and Khalil *et al.*, (2017) produced NOAELs of 0.03 and 0.066 mg/kg/d, equivalent to body burdens of 0.5 and 1.44 mg/kg body weight, respectively. We adopted the toxicokinetic approach in EFSA (2011) to convert these body burdens into equivalent human daily intakes of 374 and 1079 ng/kg d. By allocation of an assessment factor of 2.5, this yields 150 and 430 ng/kg d as possible RfDs. In the interest of conservatism, we selected 150 ng/kg d for BDE 47.

BDE 99: Kuriyama *et al.*, (2005) allows identification of a NOAEL of 0.02 mg/kg d, with an estimated equivalent human intake of 7.2 ng/kg d which by application of an assessment factor of 2.5 produces 2.9 ng/kg d as RfD for BDE 99.

BDE 209: Based on the studies by Tseng *et al.*, (2006) and Sarkar *et al.*, (2016) we derive NOAELs of 100 and 750 mg/kg d which by application of an assessment factor of 100 give RfDs of 1 and 7.5 mg/kg d, respectively. For the purposes of this MRA we selected 1 mg/kg d.

Other relevant BDE congeners: Data about the ability of other abundant BDE congeners, such as BDE 28, 100, 153, 154 and 183 to produce declines in semen quality are missing. Exclusion of these congeners from consideration in the MRA will lead to underestimations of risk. We therefore adopted the read-across approach used by Martin *et al.*, (2017) to assign RfDs to these untested congeners. In this read-across approach we assumed that the potency of untested congeners is similar to that of the nearest tested congeners. Accordingly, we chose the RfD of BDE 47 for BDE 28 (150 ng/kg d). The RfD of BDE 99 (2.9 ng/kg d) was selected to also cover BDE 100, 153 and 154. Finally, we assumed that the untested BDE 183 is as potent as BDE 209.

There is epidemiological evidence of associations of PBDE exposure with poor semen quality, but these studies do not allow congener-specific evaluations, or if they do, have monitored PBDEs in tissues that make it difficult to relate the data to daily intakes (e.g. PBDEs in hair).

6.3.5 PCDD/F

We adopted the HBGV of 0.28 pg/kg d (equivalent to the tolerable weekly intake value of 2 pg/kg d) established by EFSA (2018) for PCDD/F. This is based on the associations between PCDD/F exposures in early childhood and later declines in semen quality observed in the Russian Children's Cohort. Since these associations were only observed for PCDD/F congeners, and not for dioxin-like PCBs, we applied this value only to TCDD equivalents for 17 PCDD/F congeners (assessment group B in EFSA 2018).

6.3.6 PCBs

PCB 118, 126, 169: To establish congener-specific RfDs for PCBs we conducted a systematic review of human epidemiological and animal studies of PCB exposures and declines in semen

quality (to be published elsewhere). Our keyword searches identified 6897 records, of which 175 qualified for full-text analysis. Of these, 68 papers were subjected to data extraction. Based on the studies we rated as being of sufficient quality we estimated the following RfDs:

PCB 118: He *et al.*, (2020) derived a NOAEL of 6 µg/ kg d in rats which results in a body burden of 0.036 mg/kg bw. The human equivalent daily intake associated with this rodent body burden can be estimated as 7.25 ng/kg d which we converted to a RfD of 2.9 ng/kg d by application of an inter-species assessment of 2.5.

PCB 126: Based on the study by Wakui *et al.*, (2010) a NOAEL of 25 ng/kg d can be estimated, resulting in a body burden of 0.15 mg/kg bw, equivalent to an estimated human daily intake of 0.18 ng/kg d. We applied an inter-species assessment factor of 2.5 and obtained 0.073 ng/kg d as the RfD.

PCB 169: The study by Xiao et al. (2010) allows estimation of a NOAEL of 0.0083 mg/kg d. We calculated that this dose produces a body burden of 0.051 mg/kg bw. The human equivalent dose associated with this rodent body burden is 13.33 ng/kg d which gives an RfD of 5.33 ng/kg d (inter-species assessment factor of 2.5).

6.3.7 Phthalates

We adopted the values proposed by Kortenkamp and Koch 2020 which were derived based on suppressions of testosterone synthesis and other effects related to the phthalate syndrome in rats. There is evidence to associate exposure to phthalates in adult life with reductions in semen quality (Radke *et al.*, 2018). In this MRA, we used the following RfD values:

DEHP: 10 µg/kg d

DBP: 6.7 µg/kg d

BBzP: 10 µg/kg d

DINP: 59 µg/kg d

DIBP: 100 µg/kg d

6.3.8 Acrylamide

Acrylamide produces reductions in sperm production in rodents dosed in adulthood (Kermani-Alghoraishi *et al.*, 2010; Kalaivani *et al.*, 2018; Ivanski *et al.*, 2020) through a mechanism involving disturbance of testosterone levels. The lowest reported LOAEL of 2.5 mg/kg d is by Ivanski et al. (2020). Kermani-Alghoraishi et al. 2010, Kalaivani et al. 2018 observed 5 and 6.5 mg/kg d, respectively. By application of a LOAEL to NOAEL extrapolation factor of 3, and the default factor of 100, we obtained a RfD of 8.3 µg/kg d.

6.3.9 Chlorpyrifos

Chlorpyrifos acts as an AR antagonist and produces declines in semen quality in laboratory animals (Ubaid ur Rahman *et al.*, 2021). In the studies we retrieved, chlorpyrifos was administered at only one dose level; studies that permit a dose-response analysis could not be located. Babazadeh and Najafi (2017) observed declines in semen quality in rats that received 37 mg/kg d, Alaa-Eldin, *et al.* (2017) and Hassan *et al.*, (2021) saw similar effects at 6.5 mg/kg d. The lowest reported dose for declines in semen quality was 3 mg/kg d, given to adult rats over 20 weeks (Li *et al.*, 2019). We deemed this dose a LOAEL and estimated a RfD of 10 µg/kg d (assessment factor of 3 for LOAEL to NOAEL extrapolation, plus a default factor of 100).

6.3.10 Vinclozolin, procymidone, linuron, prochloraz and fenitrothion

We adopted the values used by Kortenkamp and Faust, (2010) for vinclozolin, procymidone, linuron and fenitrothione. The value for prochloraz is taken from Laier *et al.*, (2006). All values are based on alterations of landmarks of male sexual development in the rat (retained nipples in male offspring, changes in anogenital distance, AGD). The estimates were used as proxy's to bridge the absence of data on declines in semen quality. The following RfD were used:

Vinclozolin: 50 µg/kg d

Procymidone: 100 µg/kg d

Linuron: 100 µg/kg d

Prochloraz: 100 µg/kg d

Fenitrothione: 200 µg/kg d

6.3.11 Paracetamol

Paracetamol interferes with COX signaling and suppresses InsL3, a factor required for the first phase of testis descent. It also impairs semen quality. We could not locate dose-response studies of declines in semen quality and paracetamol in laboratory animals. The only studies that administered paracetamol and examined parameters of semen quality are of Axelstad *et al.*, (2018) in rats and Rossitto *et al.*, (2019) in mice. Rossitto et al. (2019) dosed paracetamol twice daily in mice during gestation and saw reductions in sperm number at 30 mg/kg d. Axelstad et al. (2018) gave 360 mg/kg d through-out gestation and lactations and observed declines in semen quality. We used the latter study to derive a RfD of 1 mg/kg d, by applying a factor of 3 for LOAEL to NOAEL extrapolation, and an additional 100 for inter- and intra-species extrapolation.

6.4 Exposure assessment

The exposure assessment arm of a MRA has to rely on information about the chemicals that occur together in relevant exposure media. Ideally, this should be based on measurements of multiple chemicals in one and the same sample. However, with the substances considered in this case study, this requirement could not be met completely. Only for subsets of chemicals was it possible to retrieve publications where multiple substances were measured in the same subjects, e.g. bisphenols A, F, S and phthalates (Frederiksen *et al.*, 2020). We gave preference to such exposure data. For the remainder of the substances considered here, we collated single chemical exposure data from different sources, making the tacit assumption that the general population is exposed to all members of the CAG simultaneously, most of the time. We further assumed that basing the assessment on median exposure levels for multiple substances gives a reasonable approximation of the cumulative exposure patterns at a population level. To construct a worst-case scenario, we also investigated cumulative exposures with all substances at their 95th percentile.

Table 6.2: Exposure levels for substances in the CAG for male reproductive health

Chemical	Median	95th percentile	Dose unit	Comments
Bisphenol A	0.048	0.159	µg/kg d	Frederiksen 2020
Bisphenol F	0.006	0.04	µg/kg d	Frederiksen 2020
Bisphenol S	0.002	0.016	µg/kg d	Frederiksen 2020
n-butylparaben	0.6	8.7	µg/kg d	Moos et al 2016, median and P95 in women, Tab 3
BDE 28	0.02	0.29	ng/kg d	EFSA 2011, median lower bound, P95 upper bound, adults, Tab 17, p 63
BDE 47	0.58	1.97	ng/kg d	EFSA 2011, median lower bound, P95 upper bound, adults, Tab 17, p 63
BDE 99	0.18	0.67	ng/kg d	EFSA 2011, median lower bound, P95 upper bound, adults, Tab 17, p 63
BDE 100	0.15	0.64	ng/kg d	EFSA 2011, median lower bound, P95 upper bound, adults, Tab 17, p 63
BDE 154	0.05	0.53	ng/kg d	EFSA 2011, median lower bound, P95 upper bound, adults, Tab 17, p 63
BDE 153	0.04	0.48	ng/kg d	EFSA 2011, median lower bound, P95 upper bound, adults, Tab 17, p 63
BDE 183	0.02	0.42	ng/kg d	EFSA 2011, median lower bound, P95 upper bound, adults, Tab 17, p 63
BDE 209	0.61	3.02	ng/kg d	EFSA 2011, median lower bound, P95 upper bound, adults, Tab 17, p 63
PCDD/F	0.25	1.31	pg/kg d	EFSA 2018, mean median lower bound adults, Tab 43, p 186
PCB 118	575	575	pg/kg d	EFSA 2018, based on data in Fig 19, p 187, mean median LB
PCB 126	3.5	3.5	pg/kg d	EFSA 2018, based on data in Fig 19, p 187, mean median LB
PCB 169	3.5	3.5	pg/kg d	EFSA 2018, based on data in Fig 19, p 187, mean median LB
DEHP	2.78	9.5	µg/kg d	Frederiksen 2020
DBP	1.61	5.34	µg/kg d	Frederiksen 2020
BBzP	0.33	1.42	µg/kg d	Frederiksen 2020
DINP	0.79	3.18	µg/kg d	Frederiksen 2020
DIBP	1.72	6.57	µg/kg d	Frederiksen 2020
Acrylamide	0.4	2	µg/kg d	EFSA 2015, lower bound mean, p 60
Chlorpyrifos	0.07	0.07	µg/kg d	EFSA 2014, exposure 7% ADI (ADI: 1 ug/kg d)
Vinclozolin	0.35	0.35	µg/kg d	EFSA 2011, 2009 food monitoring report
Procymidone	0.25	0.25	µg/kg d	EFSA 2011, 2009 food monitoring report
Linuron	0.069	0.069	µg/kg d	EFSA 2011, 2009 food monitoring report
Prochloraz	0.34	0.34	µg/kg d	EFSA 2011, 2009 food monitoring report
Fenitrothione	0.06	0.06	µg/kg d	EFSA 2011, 2009 food monitoring report
Paracetamol	7	21	mg/kg d	Therapeutic doses

As much as possible we relied on exposure data from Human Biomonitoring studies. Where this was not possible, we retrieved exposure data derived from pathway analyses and food consumption and -prevalence data.

Over the years, there have been changing exposure patterns, due to the substitution of certain chemicals by alternative substances. An example is the phthalate DBP which has been used less and less since the 1990s. There have been corresponding increases in exposures to less toxic phthalates such as DINP (Apel *et al.*, 2020). Similarly, Frederiksen et al. (2020) observed decreases in exposures to phenolic substances between 2009 and 2017. Such changes are likely to have an impact on the outcome of our MRA. To avoid biasing the MRA by using exposure data from different years, we focused on one time period as much as possible and selected the years 2009 / 2010.

Table 6.2 shows a compilation of the exposure levels we used in our MRA case study.

6.4.1 Bisphenol A, F, S, phthalates

Frederiksen et al. (2020) measured urinary levels of multiple chemicals in young Danish men, among them bisphenol A, F, S and the phthalates DIBP, DBP, BBzP, DEHP and DINP in samples drawn in 2009. We converted the urinary level to estimated daily intakes, by using the approaches

described in EFSA (2015a) and Koch *et al.*, (2012). The estimated daily intakes of the bisphenols agreed well with the values reported by Karrer et al. (2020). The daily intakes for DIBP, DBP and BBzP were up to twice as high as those in Apel et al. (2020). There was good agreement with the values for DEHP and DINP reported by Apel et al. (2020).

6.4.2 n-Butylparaben

We selected the daily intakes estimated by Moos *et al.*, (2017) based on Human Biomonitoring of urinary n-butylparaben concentrations from women in samples from 2009. There was a noticeable gender influence of exposures to n-butylparaben, with higher exposures in women, possibly due to the more frequent use of cosmetics and personal care products.

6.4.3 Poly-brominated diphenyl ethers

Extensive exposure estimations for PBDEs are available from the CONTAM Panel of EFSA (EFSA, 2011). We based our evaluation on the median lower bound intake estimates for adults for BDE 28, 47, 99, 100, 153, 154, 183 and 209, as well as the corresponding 95th percentiles (upper bound).

6.4.4 PCDD/F

We adopted the intake estimates reported in EFSA (2018) for adults (median, lower bound and 95th percentile, lower bound). We based our assessment on TCDD equivalents for 17 PCDD/F congeners (assessment group B in EFSA 2018).

6.4.5 PCBs

The most recent intake estimates for PCBs 118, 126 and 169 are available in EFSA (2018). That report details the percentage contribution of certain PCB congeners (weighted with TEFs) to the overall lower bound mean exposure to 29 PCDD/F and dioxin-like PCB congeners (EFSA 2018, p 187, Fig 19). We used this information to reconstruct daily intakes for PCBs 118, 126 and 169.

6.4.6 Acrylamide

EFSA (2015b) estimated the mean dietary exposures to acrylamide for adults which we used, together with the estimates for the 95th percentile.

6.4.7 Chlorpyrifos

The 2014 EFSA (EFSA 2014) peer review of chlorpyrifos found dietary exposures in the EU to be 7 % of the ADI. As the ADI is 1 µg/kg d, we took the exposure to be 0.07 µg/kg d.

6.4.8 Vinclozolin, procymidone, linuron, prochloraz and fenitrothion

Exposure estimated for these pesticides were taken from the 2009 food monitoring report by EFSA (2011).

6.4.9 Paracetamol

Quantifying expectant mothers' exposures to over-the-counter analgesics such as paracetamol is not straightforward. Zafeiri *et al.*, (2021) have reviewed the issue and cite studies showing that the prevalence of self-medication during pregnancy is 32 % worldwide. In Western Europe and the USA this is closer to 50-60 %. In Europe, paracetamol, ibuprofen and aspirin are widely used by pregnant women. It is unclear whether the use of these analgesics in pregnancy occurs through self-medication, physicians' prescription, or both. Information about the dose levels and the timing of medication is also not readily available. This is of importance as paracetamol use towards the end of the first trimester and the beginning of the second trimester is associated with increased

risks of non-descending testes in newborn boys. Medication at other times during pregnancy does not affect the risk of cryptorchidism.

We therefore considered an exposure scenario where an expectant mother takes one 500 mg paracetamol dose during the critical time window, equivalent to 7 mg/kg d. As paracetamol can be used up to three times a day at a dose of 1000 mg, this can easily rise to 42 mg/kg d. In our high exposure scenario (Tier 1) we considered 14 mg/kg d.

Paracetamol is used in poultry and pork farming and as a result has entered the food chain. However, reliable exposure assessments for this route of uptake could not be located.

6.5 Mixture risk assessment

We followed the workflow defined for the hazard assessment arm of the MRA in the previous Deliverable D 15.4. We were able to skip Tier 1 of the workflow and commenced the assessment in Tier 2 of the hazard assessment arm of the MRA. Because the RfDs we derived for declines in semen quality were based on human epidemiological evidence and data from animal studies, it was difficult to apply the PODI approach. We therefore used the HI method.

For the exposure assessment arm, we used two Tiers. In the first Tier, we employed exposure estimates at the 95th percentile for all chemicals in the CAG. Because it is unlikely that exposures for all included chemicals at the 95th percentile will be experienced simultaneously all the time, the estimates derived from this Tier define the upper limit of a worst-case scenario.

To derive more realistic estimates of exceedances of combined exposures, we used median exposure estimates in the second Tier of the assessment.

6.5.1 Tier 1 assessment: Exposures at the 95th percentile

The outcome of the MRA based on exposures at the 95th percentile is shown in Table 6.3. With a value of 42.6, the sum of RQs (HI), by far exceeded the value of 1. For bisphenol A, paracetamol, PCDD/F, bisphenol F and S, the individual RQs are already above 1, indicating risk management issues. The HI for these chemicals alone amounted to 39, equivalent to 91 % of the HI. Thus, 5 out of 29 chemicals, or 17 % of the substances in this CAG, explain 91 % of the RQ, close to Pareto's 20-80 rule.

Table 6.3: Risk quotients and sum of risk quotients for Tier 1 of the MRA, with exposures at the 95th percentile

Tier 1: 95th percentile exposures					
Chemical	Exposure	Unit	Reference dose	Unit	Risk Quotient
Bisphenol A	0.159	µg/kg d	0.01	µg/kg d	15.9
Paracetamol	14	mg/kg d	1	mg/kg d	14
PCDD/F	1.31	pg/kg d	0.28	pg/kg d	4.7
Bisphenol F	0.04	µg/kg d	0.014	µg/kg d	2.86
Bisphenol S	0.016	µg/kg d	0.01	µg/kg d	1.6
DEHP	9.5	µg/kg d	10	µg/kg d	0.950
DBP	5.34	µg/kg d	6.7	µg/kg d	0.797
n-butylparaben	8.7	µg/kg d	30	µg/kg d	0.29
Acrylamide	2	µg/kg d	8.3	µg/kg d	0.241
BDE 99	0.67	ng/kg/d	2.9	ng/kg/d	0.231
BDE 100	0.64	ng/kg/d	2.9	ng/kg/d	0.221
PCB 118	575	pg/kg/d	2900	pg/kg/d	0.198
BDE 154	0.53	ng/kg/d	2.9	ng/kg/d	0.183
BDE 153	0.48	ng/kg/d	2.9	ng/kg/d	0.166
BBzP	1.42	µg/kg d	10	µg/kg d	0.142
DIBP	6.57	µg/kg d	100	µg/kg d	0.066
DINP	3.18	µg/kg d	59	µg/kg d	0.054
PCB 126	3.5	pg/kg/d	73	pg/kg/d	0.048
BDE 47	1.97	ng/kg/d	150	ng/kg/d	0.013
Chlorpyrifos	0.07	µg/kg d	10	µg/kg d	0.007
Vinclozolin	0.35	µg/kg d	50	µg/kg d	0.007
Procymidone	0.25	µg/kg d	100	µg/kg d	0.0025
Prochloraz	0.34	µg/kg d	160	µg/kg d	0.002
BDE 28	0.29	ng/kg/d	150	ng/kg/d	0.002
Linuron	0.069	µg/kg d	100	µg/kg d	0.00069
PCB 169	3.5	pg/kg/d	5330	pg/kg/d	0.00066
Fenitrothione	0.06	µg/kg d	200	µg/kg d	0.0003
BDE 209	3.02	ng/kg/d	1000000	ng/kg/d	0.000
BDE 183	0.42	ng/kg/d	1000000	ng/kg/d	0.000
Sum of RQ					**42.66**

If exposure to all chemicals which exceeded their RfDs were equal to their RfDs, i.e. if these RQs were close to 1, and all the pother RQs stayed as they are, the HI would reduce to 8.6.

6.5.2 Tier 2 assessment: Median exposures

We refined our analysis by using median exposures (Table 6.4). With a value of 14.4, the corresponding HI was still considerably larger than 1. Exposures of two substances, paracetamol and bisphenol A already on their own exceeded RQs of 1. Together, paracetamol and bisphenol A made up 82 % of the HI. Thus, based on median exposures, 2 out of 29 chemicals explained more than 80 % of the HI.

If paracetamol and bisphenol A complied with their RfDs, the HI would decrease to 4.6.

Table 6.4: Risk quotients and sum of risk quotients for Tier 2 of the MRA, with median exposures

Tier 2: Median exposures					
Chemical	**Exposure**	**Unit**	**Reference dose**	**Unit**	**Risk Quotient**
Paracetamol	7	mg/kg d	1	mg/kg d	7
Bisphenol A	0.048	µg/kg d	0.01	µg/kg d	4.8
PCDD/F	0.25	pg/kg d	0.28	pg/kg d	0.9
Bisphenol F	0.006	µg/kg d	0.014	µg/kg d	0.43
DEHP	2.78	µg/kg d	10	µg/kg d	0.278
DBP	1.61	µg/kg d	6.7	µg/kg d	0.240
Bisphenol S	0.002	µg/kg d	0.01	µg/kg d	0.2
PCB 118	575	pg/kg/d	2900	pg/kg/d	0.198
BDE 99	0.18	ng/kg/d	2.9	ng/kg/d	0.062
BDE 100	0.15	ng/kg/d	2.9	ng/kg/d	0.052
Acrylamide	0.4	µg/kg d	8.3	µg/kg d	0.048
PCB 126	3.5	pg/kg/d	73	pg/kg/d	0.048
BBzP	0.33	µg/kg d	10	µg/kg d	0.033
n-butylparaben	0.6	µg/kg d	30	µg/kg d	0.02
BDE 154	0.05	ng/kg/d	2.9	ng/kg/d	0.017
DIBP	1.72	µg/kg d	100	µg/kg d	0.017
BDE 153	0.04	ng/kg/d	2.9	ng/kg/d	0.014
DINP	0.79	µg/kg d	59	µg/kg d	0.013
Chlorpyrifos	0.07	µg/kg d	10	µg/kg d	0.007
Vinclozolin	0.35	µg/kg d	50	µg/kg d	0.007
BDE 47	0.58	ng/kg/d	150	ng/kg/d	0.004
Procymidone	0.25	µg/kg d	100	µg/kg d	0.0025
Prochloraz	0.34	µg/kg d	160	µg/kg d	0.002
Linuron	0.069	µg/kg d	100	µg/kg d	0.00069
PCB 169	3.5	pg/kg/d	5330	pg/kg/d	0.00066
Fenitrothione	0.06	µg/kg d	200	µg/kg d	0.0003
BDE 28	0.02	ng/kg/d	150	ng/kg/d	0.000
BDE 209	0.61	ng/kg/d	1000000	ng/kg/d	0.000
BDE 183	0.02	ng/kg/d	1000000	ng/kg/d	0.000
Sum of RQ					**14.39**

6.5.3 Drivers of mixture risks

Following on from the principles developed in D 15.4, we identified as drivers of mixture risks those chemicals that increase the HI beyond 1, when the RQs of all chemicals are arrayed in an ascending order. As shown in Figure 6.2 for the Tier 1 MRA, the RQs of all PBDEs, pesticides, PCBs, and the phthalates DINP, DIBP and BBzP collectively stay below 1. Adding the RQs of acrylamide, n-butylparaben, DBP, DEHP, bisphenols A, F, S, PCDD/F and paracetamol increases the HI to values above 1. Accordingly, these 9 chemicals (or chemical groups) can be considered drivers of mixture risks at the 95 % percentile of exposures.

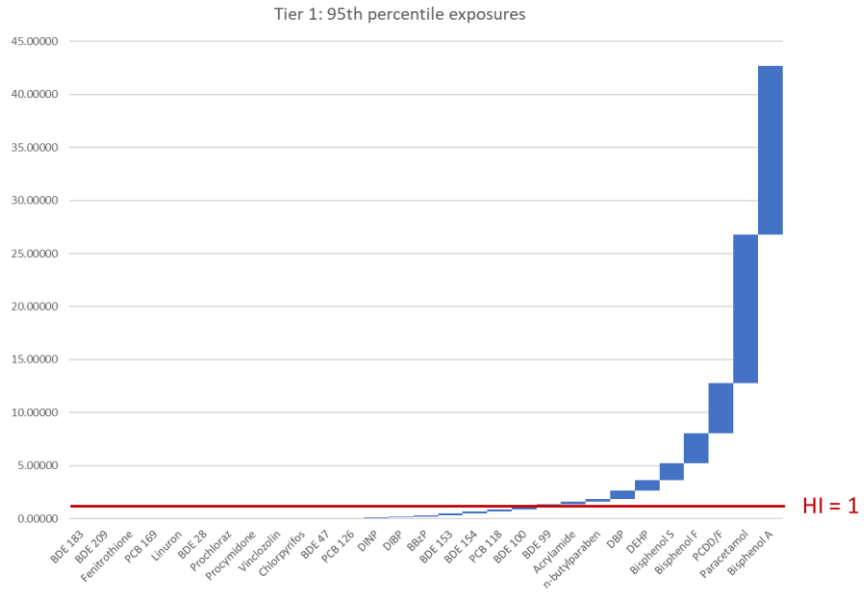


Figure 6.2: Drivers of mixture risk at Tier 1 of the MRA (exposures at the 95th percentile). The risk quotients of all chemicals in the CAG were arrayed in ascending order. The graph shows cumulative HIs. The red horizontal line depicts HI = 1; the substances above that line contribute to combined exposures above HI = 1.

Based on median exposures, the number of drivers decreased to 5 substances, DEHP, bisphenols A and F, PCDD/F, and paracetamol. By targeting these chemicals with exposure reduction measures, mixture risks to male reproductive health could be mitigated significantly (Figure 6.3).

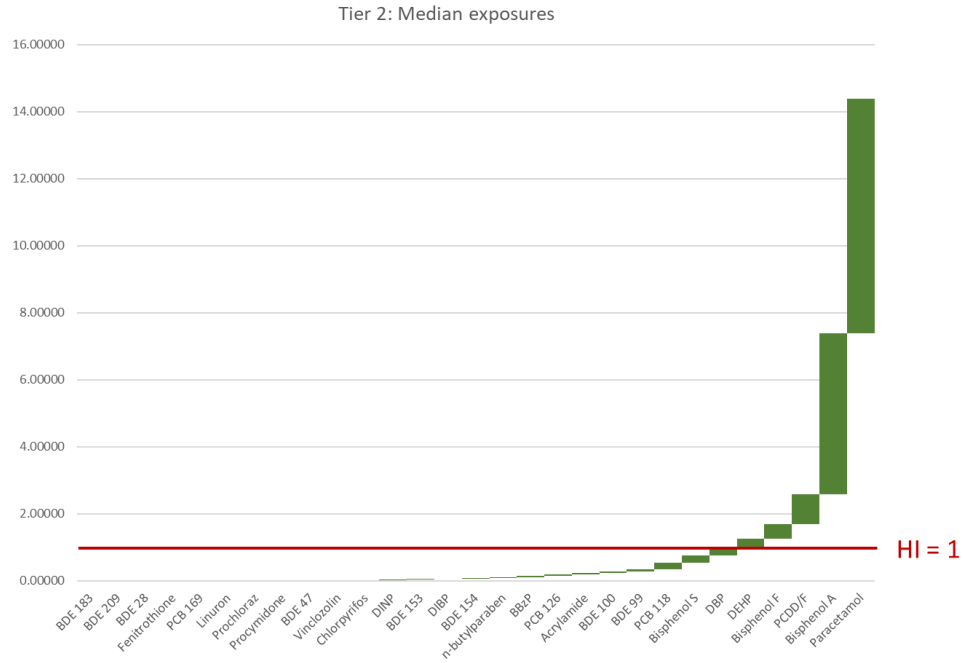


Figure 6.3: As Figure 6.2, but for Tier 2 of the MRA (median exposures).

6.5.4 Interpretation of results and uncertainty considerations

As in single chemical risk assessments, HI larger than 1 indicate fold-exceedances of combined acceptable exposures. Since position and gradient of the underlying dose-response relationships are unknown, it is not possible to infer the extent of risks in terms of effect magnitudes from these data. However, as in single chemical risk assessment, exceedances of acceptable exposure indicate the need for risk management measures.

There are some uncertainties associated with our evaluation, and these stem from both the exposure assessment and hazard characterization components of our analysis.

6.5.4.1 Uncertainties in the hazard characterization step

Because data about the ability of **BDE 28, 100, 153, 154 and 183** to produce declines in semen quality are missing altogether, we had to bridge these data gaps by read-across approaches. However, the corresponding RQs were rather small. Accordingly, even if the necessary extrapolation step over-estimates risks, the impact on the overall HI is small.

Similar considerations apply to all the **pesticides** in the CAG for which we could not locate data about declines in semen quality and instead had to use data about their ability to interfere with hallmarks of male sexual differentiation in rodents. The contribution of these pesticides to the sum of RQ is negligible.

Of greater consequence to the outcome of our assessment are the uncertainties in the hazard characterization for some of the chemicals we identified as drivers of mixture risks. Prominent among them is **paracetamol**. There are currently no epidemiological data about associations between paracetamol exposure in fetal life and declines in semen quality. Consequently, we had to infer such hazards from animal data. However, the few laboratory studies available (Axelstad et al. 2018, Rossitto et al. 2018) used only one dose level; dose-response analyses are not possible. The paracetamol RfD we used to build the RQ may strongly underestimate the potency of paracetamol. Detailed dose-response studies in rodents would greatly reduce these uncertainties. It will also be worthwhile to investigate the impact of gestational exposures to paracetamol on semen quality in epidemiological studies.

We could locate only one experimental study of declines in semen quality with **bisphenol F** (Ullah et al. 2019). This reduces the confidence we can attach to our RfD estimate for this substance. It is recommended to conduct further studies to substantiate this estimate.

In contrast, we have high confidence in the RfDs used for bisphenol A, PCDD/F and DEHP.

The value for bisphenol A was derived de novo based on a systematic review of animal studies of declines in semen quality after gestational exposure. The RfD we estimated is in the range of exposures associated with poor semen quality in epidemiological studies.

The RfD we adopted for PCDD/F is the HBGV derived by EFSA (2018), which is based on the latest insights into associations of PCDD/F exposures and poor semen quality in the Russian Children's study.

Although the RfD estimated for DEHP is based on testicular dysgenesis in the rat (Christiansen et al. 2010), and not on declines in semen quality, the numerical value is in the range of exposures found to be associated with semen quality declines in epidemiological studies (Radke et al. 2018).

There are indications that exposures to **perfluorinated compounds, phytoestrogens**, **pyrethroids** and **triclosan** are also associated with poor semen quality after gestational exposures. However, within the resources available, we could not locate data of sufficient quality

that would have allowed us to include these substances in our assessment. This omission may have led to underestimations of risks.

Furthermore, we were unable to evaluate the contribution of the other widely used analgesics, **ibuprofen** and **aspirin**. This data gap will bias our MRA towards underestimations of risks.

6.5.4.2 Uncertainties in the exposure assessment arm of the MRA

The greatest source of uncertainty in the exposure assessment in this MRA is in the lack of information about the co-exposures that occur on an individual basis. Even where multiple chemical exposures were measured in the same individuals, as in Frederiksen et al. (2020), the way the data are reported does not allow us to reconstruct individual exposure patterns. The use of median exposures can approximate the patterns experienced on a population level but is not capable of identifying individuals that suffer high combined exposures.

With the exception of paracetamol, we can attach high confidence in the exposure estimates that we used for all the drivers of mixture risks, i.e. acrylamide, n-butylparaben, DBP, DEHP, bisphenols A, F, S, and PCDD/F. The estimates for n-butylparaben, DBP, DEHP, bisphenols A, F, S were based on Human Biomonitoring studies, and the values we used for DBP, DEHP, bisphenols A, F, S even came from a monitoring exercise where these chemicals were measured in the same study subjects (Frederiksen et al. 2020). The estimates for acrylamide and PCDD/F were derived from large-scale European food monitoring (EFSA 2015, 2018).

In contrast, the exposure estimates we used for paracetamol are highly uncertain, due to a lack of data about the precise usage by expectant mothers. It is very likely that our assumptions represent significant underestimations of the true risks.

6.5.5 Prospects of further refinements of the MRA

Our assessment could be improved by filling the data gaps identified above. However, there is little prospect that the distorting influence introduced by utilizing data from animal studies and epidemiological studies can be removed in the foreseeable future. Addressing these shortcomings would mean to conduct time-consuming epidemiological studies, at least of paracetamol and bisphenol F.

The uncertainties associated with estimating paracetamol exposures could be addressed by better Human Biomonitoring strategies for this substance.

6.6 Conclusions

In Europe, combined exposures that can be deemed acceptable in terms of avoiding risks of declining semen qualities are exceeded by very large margins. In the worst case, assuming simultaneous exposures at the 95th percentile of all the single chemicals entered into the assessment, acceptable combined exposures are exceeded 42-fold. Based on a more realistic scenario of exposures at the median of all single chemicals, we estimate 14-fold exceedances of combined acceptable levels. It can be assumed that a large fraction of the population experiences combined risks between these extremes.

It is striking that the resulting mixture risks seem to be driven by a relatively small number of chemicals, no more than 5 to 9 of the 29 substances (or in the case of PCDD/F substance groups) that we considered in this case study. At median exposures of all chemicals, paracetamol and bisphenol A alone contributed 82 % of the overall HI.

The problem is a mixed constellation of true combined risks and non-compliance of single chemicals with their individual RfD, signaling the need for risk management measures at the single chemical level. However, even if the RQs of all the substances considered here stayed below 1, the sum of RQ would still exceed 1, highlighting the need for dedicated mixture risk mitigation strategies.

Our findings provide solid guidance for future Human Biomonitoring studies fit for the purpose of supporting mixture risk assessments. We propose that methods for the simultaneous monitoring of acrylamide, n-butylparaben, DBP, DEHP, bisphenols A, F, S, PCDD/F and paracetamol should be developed and implemented. These 9 chemicals and substance groups drive the combined risks to male reproductive health. A biomonitoring strategy focusing on these chemicals could underpin better epidemiological studies of possible associations with declines in semen quality and lead to proper health impact assessments.

6.7 References

Alaa-Eldin, E. A., El-Shafei, D. A. and Abouhashem, N. S. (2017) 'Individual and combined effect of chlorpyrifos and cypermethrin on reproductive system of adult male albino rats', *Environmental Science and Pollution Research*. Environmental Science and Pollution Research, 24(2), pp. 1532–1543. doi: 10.1007/s11356-016-7912-6.

Apel, P. *et al.* (2020) 'Time course of phthalate cumulative risks to male developmental health over a 27-year period : Biomonitoring samples of the German Environmental Specimen Bank', *Environment International*. Elsevier, 137(December 2019), p. 105467. doi: 10.1016/j.envint.2020.105467.

Axelstad, M. *et al.* (2018) 'EDC IMPACT: Reduced sperm counts in rats exposed to human relevant mixtures of endocrine disrupters', *Endocrine Connections*, 7(1), pp. 139–148. doi: 10.1530/EC-17-0307.

Babazadeh, M. and Najafi, G. (2017) 'Effect of chlorpyrifos on sperm characteristics and testicular tissue changes in adult male rats.', *Veterinary research forum : an international quarterly journal*, 8(4), pp. 319–326. Available at: http://www.ncbi.nlm.nih.gov/pubmed/29326791%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5756252.

Boberg, J. *et al.* (2016) 'Multiple endocrine disrupting effects in rats perinatally exposed to butylparaben', *Toxicological Sciences*, 152(1), pp. 244–256. doi: 10.1093/toxsci/kfw079.

Chia, V. M. *et al.* (2010) 'International trends in the incidence of testicular cancer, 1973-2002', *Cancer Epidemiology Biomarkers and Prevention*, 19(5), pp. 1151–1159. doi: 10.1158/1055-9965.EPI-10-0031.

Christiansen, S. *et al.* (2009) 'Synergistic disruption of external male sex organ development by a mixture of four antiandrogens', *Environmental Health Perspectives*, 117(12), pp. 1839–1846. doi: 10.1289/ehp.0900689.

EFSA (2011) 'Scientific Opinion on Polybrominated Diphenyl Ethers (PBDEs) in Food', *EFSA Journal*, 9(5), pp. 1–274. doi: 10.2903/j.efsa.2011.2156.

EFSA (2014) 'Conclusion on the peer review of the pesticide human health risk assessment of the active substance chlorpyrifos', *EFSA Journal*, 12(4), pp. 1–34. doi: 10.2903/j.efsa.2014.3640.

EFSA (2015a) 'Scientific Opinion on the risks to public health related to the presence of bisphenol A (BPA) in foodstuffs', *EFSA Journal*, 13(1). doi: 10.2903/j.efsa.2015.3978.

EFSA (2015b) 'Scientific Opinion on acrylamide in food', *EFSA Journal*, 13(6). doi: 10.2903/j.efsa.2015.4104.

EFSA (2018) 'Risk for animal and human health related to the presence of dioxins and dioxin-like PCBs in feed and food', *EFSA Journal*, 16(11). doi: 10.2903/j.efsa.2018.5333.

Frederiksen, H. *et al.* (2020) 'Changes in urinary excretion of phthalates, phthalate substitutes, bisphenols and other polychlorinated and phenolic substances in young Danish men; 2009–2017', *International Journal of Hygiene and Environmental Health*. Elsevier, 223(1), pp. 93–105. doi: 10.1016/j.ijheh.2019.10.002.

Hass, U. *et al.* (2016) 'Low-dose effect of developmental bisphenol A exposure on sperm count and behaviour in rats', *Andrology*, 4(4), pp. 594–607. doi: 10.1111/andr.12176.

Hassan, M. A. *et al.* (2021) 'Iprodione and chlorpyrifos induce testicular damage, oxidative stress, apoptosis and suppression of steroidogenic- and spermatogenic-related genes in immature male albino rats', *Andrologia*, 53(4), pp. 1–14. doi: 10.1111/and.13978.

He, Q. L. *et al.* (2020) 'Effects of intrauterine exposure to 2,3′,4,4′,5-pentachlorobiphenyl on the reproductive system and sperm epigenetic imprinting of male offspring', *Journal of Applied Toxicology*, 40(10), pp. 1396–1409. doi: 10.1002/jat.3992.

Howdeshell, K. L., Hotchkiss, A. K. and 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*. Elsevier GmbH., 220(2), pp. 179–188. doi: 10.1016/j.ijheh.2016.11.007.

Ivanski, F. *et al.* (2020) 'Prepubertal acrylamide exposure causes dose-response decreases in spermatic production and functionality with modulation of genes involved in the spermatogenesis in rats', *Toxicology*. Elsevier, 436(March), p. 152428. doi: 10.1016/j.tox.2020.152428.

Kalaivani, M. *et al.* (2018) 'Effect of acrylamide ingestion on reproductive organs of adult male wistar rats', *Journal of Clinical and Diagnostic Research*, 12(11). doi: 10.7860/JCDR/2018/38170.12364.

Kang, K. S. *et al.* (2002) 'Decreased sperm number and motile activity on the F1 offspring maternally exposed to butyl p-hydroxybenzoic acid (butyl paraben)', *Journal of Veterinary Medical Science*, 64(3), pp. 227–235. doi: 10.1292/jvms.64.227.

Kermani-Alghoraishi, M. *et al.* (2010) 'The effects of acrylamide on sperm parameters and membrane integrity of epididymal spermatozoa in mice', *European Journal of Obstetrics and Gynecology and Reproductive Biology*. Elsevier Ireland Ltd, 153(1), pp. 52–55. doi: 10.1016/j.ejogrb.2010.07.008.

Khalil, A. *et al.* (2017) 'Perinatal exposure to 2,2′,4′4′ −Tetrabromodiphenyl ether induces testicular toxicity in adult rats', *Toxicology*. Elsevier, 389(August 2016), pp. 21–30. doi: 10.1016/j.tox.2017.07.006.

Karrer, C *et al.* (2020) 'The Euromix human biomonitoring study: Source-to-dose modeling of cumulative and aggregate exposure for the bisphenols BPA, BPS, and BPF and comparison with measured urinary levels', *Environment International*, 136, 105397.

Koch, H. M. *et al.* (2012) 'Bisphenol A in 24 h urine and plasma samples of the German Environmental Specimen Bank from 1995 to 2009: A retrospective exposure evaluation', *Journal of Exposure Science and Environmental Epidemiology*, 22(6), pp. 610–616. doi: 10.1038/jes.2012.39.

Kortenkamp, A. (2020) 'Molecular and Cellular Endocrinology Which chemicals should be grouped together for mixture risk assessments of male reproductive disorders ?', *Molecular and Cellular Endocrinology*. Elsevier, 499(September 2019), p. 110581. doi: 10.1016/j.mce.2019.110581.

Kortenkamp, A. and Faust, M. (2010) 'Combined exposures to anti-androgenic chemicals: Steps towards cumulative risk assessment', in *International Journal of Andrology*, pp. 463–472.

Kortenkamp, A. and Koch, H. M. (2020) 'International Journal of Hygiene and Refined reference doses and new procedures for phthalate mixture risk assessment focused on male developmental

toxicity', *International Journal of Hygiene and Environmental Health*. Elsevier, 224(November 2019), p. 113428. doi: 10.1016/j.ijheh.2019.113428.

Kuriyama, S. N. *et al.* (2005) 'Developmental exposure to low-dose PBDE-99: Effects on male fertility and neurobehavior in rat offspring', *Environmental Health Perspectives*, 113(2), pp. 149–154. doi: 10.1289/ehp.7421.

Laier, P. *et al.* (2006) 'Mechanisms of action underlying the antiandrogenic effects of the fungicide prochloraz', *Toxicology and Applied Pharmacology*, 213(2), pp. 160–171. doi: 10.1016/j.taap.2005.10.013.

Levine, H. *et al.* (2017) 'Temporal trends in sperm count: A systematic review and meta-regression analysis', *Human Reproduction Update*, 23(6), pp. 646–659. doi: 10.1093/humupd/dmx022.

Li, J. *et al.* (2019) 'Chlorpyrifos-induced reproductive toxicity in rats could be partly relieved under high-fat diet', *Chemosphere*. Elsevier Ltd, 229, pp. 94–102. doi: 10.1016/j.chemosphere.2019.05.020.

Main, K. M. *et al.* (2010) 'Genital anomalies in boys and the environment', *Best Practice and Research: Clinical Endocrinology and Metabolism*. Elsevier Ltd, 24(2), pp. 279–289. doi: 10.1016/j.beem.2009.10.003.

Martin, O. V. *et al.* (2017) 'A human mixture risk assessment for neurodevelopmental toxicity associated with polybrominated diphenyl ethers used as flame retardants', *Environmental Health Perspectives*, 125(8). doi: 10.1289/EHP826.

Moos, R. K. *et al.* (2017) 'Daily intake and hazard index of parabens based upon 24h urine samples of the German environmental specimen bank from 1995 to 2012', *Journal of Exposure Science and Environmental Epidemiology*, 27(6), pp. 591–600. doi: 10.1038/jes.2016.65.

Radke, E. G. *et al.* (2018) 'Phthalate exposure and male reproductive outcomes : A systematic review of the human epidemiological evidence', *Environment International*. Elsevier, 121(October), pp. 764–793. doi: 10.1016/j.envint.2018.07.029.

Rossitto, M. *et al.* (2019) 'Intergenerational effects on mouse sperm quality after in utero exposure to acetaminophen and ibuprofen', *FASEB Journal*, 33(1), pp. 339–357. doi: 10.1096/fj.201800488RRR.

Salian, S., Doshi, T. and Vanage, G. (2009) 'Perinatal exposure of rats to Bisphenol A affects the fertility of male offspring'. Elsevier Inc., 85, pp. 742–752. doi: 10.1016/j.lfs.2009.10.004.

Sarkar, D., Chowdhury, J. P. and Singh, S. K. (2016) 'Effect of polybrominated diphenyl ether (BDE-209) on testicular steroidogenesis and spermatogenesis through altered thyroid status in adult mice', *General and Comparative Endocrinology*. Elsevier Inc., 239, pp. 50–61. doi: 10.1016/j.ygcen.2015.11.009.

Shi, M. *et al.* (2018) 'Prenatal Exposure to Bisphenol A Analogues on Male Reproductive Functions in Mice', 163(2), pp. 620–631. doi: 10.1093/toxsci/kfy061.

Shi, M. *et al.* (2019) 'Prenatal Exposure to Bisphenol A, E, and S Induces Transgenerational Effects on Male Reproductive Functions in Mice', *Toxicological Sciences*, 172(2), pp. 303–315. doi: 10.1093/toxsci/kfz207.

Skakkebaek, N. E. *et al.* (2016) 'Male Reproductive Disorders and Fertility Trends: Influences of Environment and Genetic Susceptibility', *Physiological Reviews*, 96(1), pp. 55–97. doi: 10.1152/physrev.00017.2015.

Tseng, L. H. *et al.* (2006) 'Postnatal exposure of the male mouse to 2,2′,3,3′,4,4′,5,5′,6,6′-decabrominated diphenyl ether: Decreased epididymal sperm functions without alterations in DNA content and histology in testis', *Toxicology*, 224(1–2), pp. 33–43. doi: 10.1016/j.tox.2006.04.003.

Ubaid ur Rahman, H. *et al.* (2021) 'A comprehensive review on chlorpyrifos toxicity with special reference to endocrine disruption: Evidence of mechanisms, exposures and mitigation strategies',

Science of the Total Environment. Elsevier B.V. doi: 10.1016/j.scitotenv.2020.142649.

Ullah, A. *et al.* (2019) 'Prenatal BPA and its analogs BPB , BPF , and BPS exposure and reproductive axis function in the male offspring of Sprague Dawley rats'. doi: 10.1177/0960327119862335.

Wakui, S. *et al.* (2010) 'Testicular spermiation failure in rats exposed prenatally to 3,3′,4,4′,5-pentachlorobiphenyl', *Journal of Toxicological Sciences*, 35(5), pp. 757–765. doi: 10.2131/jts.35.757.

Xiao et al. 2010 Influence of persistent thyroxine reduction on spermatogenesis in rats neonatally exposed to 2,2′,4,4′,5,5′-hexa-chlorobiphenyl, *Birth Defects Research*, 89, pp. 18-25.

Zafeiri, A. *et al.* (2021) 'Over-the-counter analgesics during pregnancy: A comprehensive review of global prevalence and offspring safety', *Human Reproduction Update*, 27(1), pp. 67–95. doi: 10.1093/humupd/dmaa042.

Zhang, Linyuan *et al.* (2016) 'N-Butylparaben Induces Male Reproductive Disorders Via Regulation of Estradiol and Estrogen Receptors', *Journal of Applied Toxicology*, 36(9), pp. 1223–1234. doi: 10.1002/jat.3291.

Zhang, T. *et al.* (2020) 'Toxicity of polybrominated diphenyl ethers (PBDEs) on rodent male reproductive system: A systematic review and meta-analysis of randomized control studies', *Science of the Total Environment*, 720(300). doi: 10.1016/j.scitotenv.2020.137419.

Zhang, Z. *et al.* (2013) 'Cytochrome P450 3A1 Mediates 2,2′,4,4′-Tetrabromodiphenyl Ether-Induced Reduction of Spermatogenesis in Adult Rats', *PLoS ONE*, 8(6). doi: 10.1371/journal.pone.0066301.

7 Case study 2: Mixture risk assessment of antiandrogenic chemicals based on human-derived hazard and exposure data

Yanying Ma[1], Anna Kjerstine Rosenmai[1], Andrea Rodríguez-Carrillo[2,3], Vicente Mustieles[2,3], Camilla Taxvig[1], Lena Reiber[5], Anja Kiesow[5], Madlen David[5], Nathalie Michelle Löbl[1], Mariana F. Fernández[2,3,4], and Anne Marie Vinggaard[1]

[1]National Food Institute, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[2]Center for Biomedical Research (CIBM), University of Granada, Spain
[3]Instituto de Investigación Biosanitaria Ibs GRANADA, Spain
[4]Consortium for Biomedical Research in Epidemiology & Public Health (CIBERESP), 18100, Spain
[5] German Environment Agency (UBA), Berlin, Germany

7.1 Rationale for the case study

The overall aim of this study was to explore a novel way of performing mixture risk assessment based on human data only. As a case we evaluated whether combined exposure to antiandrogenic chemicals that antagonize the human androgen receptor (AR) in vitro might have an impact on male reproductive health, focusing primarily on exposure during fetal and early life.

Multiple studies have shown that mixtures of chemicals can induce effects greater than those induced by single chemicals on male reproductive health disorders e.g. anogenital distance, nipple retention, external sex organ malformations, sex organ weights, and dysgenesis and hypoplasia in male reproductive organs (Christiansen et al., 2009; Hass et al., 2007; Metzdorff et al., 2007). Similar observations have been reported in vitro on the androgen (Birkhøj et al., 2004; Ermler et al., 2011; Kjærstad et al., 2010; Orton et al., 2014) and estrogen receptor (Rajapakse et al., 2002; Silva et al., 2002), as well as sex hormone synthesis (Kjærstad et al., 2010).

Androgen insufficiency in the developing male fetus, which is a result of AR antagonism, will lead to shortening of anogenital distance which is considered a unique, early and non-invasive marker of male reproductive health effects in general (Scholze et al., 2020). It is plausible that this specific mechanism of action, AR antagonism, constitutes the most important molecular initiating event in the adverse outcome pathway to male reproductive health disorders and therefore is an obvious choice for characterizing hazard in a mixture risk assessment for male reproductive health disorders.

The most common way of performing chemical mixture risk assessment is to base it on the dose addition principle using hazard data retrieved from in vivo rodent data and exposure data retrieved from food intake data. Due to inherent uncertainties in this approach, we have in this case study explored the possibility of using entirely human derived data for a mixture risk assessment of antiandrogenic chemicals. Our hypothesis was that relevant risk quotients can be defined based on human-based in vitro data and Human Biomonitoring (HBM) data, reflecting aggregated internal human exposure. Using internal chemical levels in human blood has the advantage that it gives aggregated exposure levels irrespective of sources and pathway of exposure.

Moreover, we focus on chemicals with antiandrogenic potential that block the AR. Chemicals with this mode of action are known to be involved in reproductive health disorders observed in boys and men when exposed during fetal life. By default, all chemicals with this specific mode of action belong to the same common assessment group and thus considerations on which grouping criteria need to be included is redundant.

7.2 Retrieving hazard data

A bucket list of substances previously reported to be AR antagonists was prepared based on a selection of publications (Danish Environmental Protection Agency, 2017; Engel et al., 2017; Ermler et al., 2011; Kojima et al., 2004, 2009, 2019; Kortenkamp & Faust, 2010; Orton et al., 2011; Rosenmai et al., 2014; O. Shen et al., 2009; Vinggaard et al., 2008) leading to a collection of approximately 231 substances. HBM data were searched for the 231 compounds according to a systematic procedure briefly described below and combining these data led to a refined list of 77 substances having both hazard and exposure data. For these substances a comprehensive literature search for additional hazard data was performed in Web of Science. An inclusion criterion for this search was that the hazard data should stem from human AR antagonism data and thereby substances for which only animal data exist and *in vitro* effects dealing with e.g. AR mediated proliferation, AR binding, AR mediated gene expression, and effects on testosterone levels were excluded from the list. After thorough evaluation of the 77 substances, all studies in which no information on cytotoxicity was available were omitted. In addition, all substances for which the majority of the studies reported 'no effect' versus 'effect' were omitted. This resulted in a final list of 61 compounds for which reliable hazard and exposure data exist.

Inhibitory concentration of 10 % (IC_{10}) values were used as a measure of the LOEC for deriving risk quotients. In cases where no IC_{10} values were available, the following equation available in GraphPad Prism (ver.8) was used to calculate IC_{10} from the IC_{50}:

$$IC_x = \left(\frac{F}{100-F}\right)^{\frac{1}{H}} * IC_{50} \Leftrightarrow IC_{50} = \frac{IC_x}{\left(\frac{F}{100-F}\right)^{\frac{1}{H}}} \qquad Eq.1$$

where F is the percent change on the y-axis at IC_x, which can be any value above 0 and below 100. H is the Hill Slope, which was constrained to -1. Table 7.1 below summarizes IC_{10} mean and IC_{10} max values of each substance together with the corresponding references.

Table 7.1: Hazard values given as IC_{10} for human androgen receptor antagonism (measured or estimated from IC_x) given for each substance

	CAS no.	IC_{10} mean (µM)	IC_{10} max (µM)	References for IC_{10} mean	Reference for IC_{10} max
Bisphenol A	80-05-7	0.30	0.08	Araki, Ohno, Nakai, et al., 2005; Bonefeld-Jørgensen et al., 2007; Christen et al., 2012a; Di Paolo et al., 2016; Ermler et al., 2010, 2011; Krüger et al., 2008; Molina-Molina et al., 2013; Rosenmai et al., 2014; Rostkowski et al., 2011; K. Satoh et al., 2004; Satoh et al., 2005; Vinggaard et al., 2008; Wang et al., 2014; Xu et al., 2005; Zwart et al., 2017	Xu et al., 2005
Bisphenol F	620-92-8	0.55	0.33	Molina-Molina et al., 2013; Rosenmai et al., 2014; Satoh et al., 2004	Rosenmai et al., 2014
Butyl paraben	94-26-8	6.22	4.56	Ermler et al., 2011; Pop et al., 2016; Satoh et al., 2005	Ermler et al., 2011
Prochloraz	67747-09-5	0.86	0.34	Aït-Aïssa et al., 2010; Birkhøj et al., 2004; Mia B. Kjærstad et al., 2010; Roelofs et al., 2014; Vinggaard et al., 2002, 2008	Aït-Aïssa et al., 2010

	CAS no.	IC_{10} mean (µM)	IC_{10} max (µM)	References for IC_{10} mean	Reference for IC_{10} max
Procymidone	32809-16-8	0.10	0.07	Araki et al., 2005; Nellemann et al., 2003; Vinggaard et al., 2008	Nellemann et al., 2003
Butylbenzyl phthalate	85-68-7	43.04	1.44	Christen et al., 2010, 2012b; Krüger et al., 2008	Krüger et al., 2008
Dibutyl phthalate	84-74-2	19.48	0.12	Araki et al., 2005; Christen et al., 2010, 2012b; Shen et al., 2009	Shen et al., 2009
Dicyclohexyl phthalate	84-61-7	1.69	1.69	Takeuchi et al., 2005	Takeuchi et al., 2005
Diethyl phthalate	84-66-2	48.50	39.78	Christen et al., 2010, 2012b	Christen et al., 2012b
Diethylhexyl phthalate	117-81-7	30.56	>11.11	Christen et al., 2010; Shen et al., 2009	Shen et al., 2009
Di-isobutyl phthalate	84-69-5	2.07	1.38	Hu et al., 2013; Takeuchi et al., 2005	Hu et al., 2013
Dimethyl phthalate	131-11-3	85.44	85.44	Christen et al., 2010	Christen et al., 2010
Triclosan	3380-34-5	0.39	0.14	Vinggaard et al., 2008; Di Paolo et al., 2016	Di Paolo et al., 2016
PCB 138	35065-28-2	0.44	0.08	Ermler et al., 2011; Hamers et al., 2011; Misaki et al., 2015; Vinggaard et al., 2008	Misaki et al., 2015
PCB 49	41464-40-8	1.18	1.18	Vinggaard et al., 2008	Vinggaard et al., 2008
PCB 157	69782-90-7	0.40	0.40	Vinggaard et al., 2008	Vinggaard et al., 2008
PCB 156	38380-08-4	1.36	1.36	Vinggaard et al., 2008	Vinggaard et al., 2008
PCB 105	32598-14-4	0.14	0.14	Vinggaard et al., 2008	Vinggaard et al., 2008
PCB 114	74472-37-0	1.08	1.08	Vinggaard et al., 2008	Vinggaard et al., 2008
PCB 167	52668-72-6	0.59	0.59	Vinggaard et al., 2008	Vinggaard et al., 2008
PCB 118	31508-00-6	0.32	0.06	Hamers et al., 2011; Misaki et al., 2015; Vinggaard et al., 2008	Hamers et al., 2011
PCB 66	32598-10-0	0.13	0.13	Vinggaard et al., 2008	Vinggaard et al., 2008
PCB 74	32690-93-0	0.72	0.23	Hamers et al., 2011; Vinggaard et al., 2008	Hamers et al., 2011
PCB 28	7012-37-5	0.55	0.09	Hamers et al., 2011; Pěnčíková et al., 2018; Vinggaard et al., 2008	Hamers et al., 2011
PCB 126	57465-28-8	0.17	0.06	Hamers et al., 2011; Vinggaard et al., 2008	Hamers et al., 2011

	CAS no.	IC_{10} mean (µM)	IC_{10} max (µM)	References for IC_{10} mean	Reference for IC_{10} max
Ethyl paraben	120-47-8	11.78	>11.11	Ermler et al., 2011; Satoh et al., 2005	Satoh et al., 2005
Methyl paraben	99-76-3	17.11	>11.11	Ermler et al., 2011; Satoh et al., 2005	Satoh et al., 2005
n-Propyl paraben	94-13-3	8.67	7.78	Ermler et al., 2011; Satoh et al., 2005	Ermler et al., 2011
Perfluorooctane sulfonate	1763-23-1	1.48	0.52	Ermler et al., 2011; Kjeldsen & Bonefeld-Jørgensen, 2013	Kjeldsen & Bonefeld-Jørgensen, 2013
1-Aminopyrene	1606-67-3	0.32	0.32	Vinggaard et al., 2008	Vinggaard et al., 2008
1-Hydroxypyrene	5315-79-7	0.39	0.22	Rostkowski et al., 2011; Vinggaard et al., 2008	Rostkowski et al., 2011
Benzo[a]pyrene	50-32-8	0.53	0.43	Vinggaard et al., 2000, 2008	Vinggaard et al., 2000
Benzo[j]fluoranthene	205-82-3	0.22	0.22	Vinggaard et al., 2008	Vinggaard et al., 2008
Benzo[k]fluoranthene	207-08-9	0.07	0.07	Vinggaard et al., 2008	Vinggaard et al., 2008
Chrysene	218-01-9	1.11	1.07	Vinggaard et al., 2000, 2008	Vinggaard et al., 2008
Fluoranthene	206-44-0	0.22	0.21	Araki et al., 2005; Vinggaard et al., 2008	Vinggaard et al., 2008
Pyrene	129-00-0	1.51	1.51	Vinggaard et al., 2008	Vinggaard et al., 2008
Benzophenone-3	131-57-7	1.25	0.22	Ermler et al., 2011; Ma et al., 2003; Molina-Molina et al., 2008; Schreurs et al., 2005	Schreurs et al., 2005
Zearalenone	17924-92-4	1.02	1.02	Vinggaard et al., 2008	Vinggaard et al., 2008
p,p′-DDE	72-55-9	0.39	0.03	Araki et al., 2005; Kojima et al., 2003; Schrader & Cooke, 2000; Vinggaard et al., 2008; Xu et al., 2006; Misaki et al., 2015	Misaki et al., 2015
o,p-DDT	789-02-6	0.22	0.16	Aït-Aïssa et al., 2010; Araki et al., 2005; Kojima et al., 2003; Lemaire et al., 2004; Vinggaard et al., 2008	Vinggaard et al., 2008
o,p-DDE	3424-82-6	0.33	0.18	Aït-Aïssa et al., 2010; Vinggaard et al., 2008	Aït-Aïssa et al., 2010
p,p'-DDD	72-54-8	0.08	0.08	Misaki et al., 2015	Misaki et al., 2015
p,p-DDT	50-29-3	0.34	0.11	Vinggaard et al., 2008; Misaki et al., 2015	Misaki et al., 2015
TBBPA	79-94-7	1.11	1.11	Christen et al., 2010	Christen et al., 2010

	CAS no.	IC_{10} mean (µM)	IC_{10} max (µM)	References for IC_{10} mean	Reference for IC_{10} max
BDE 99	60348-60-9	0.49	0.11	Hamers et al., 2006; Suzuki et al., 2013	Suzuki et al., 2013
BDE 47	5436-43-1	0.20	0.06	Hamers et al., 2006; Liu et al., 2011; Suzuki et al., 2013	Suzuki et al., 2013
BDE 28	41318-75-6	0.34	0.34	Hamers et al., 2006	Hamers et al., 2006
BDE100	189084-64-8	0.23	0.01	Ermler et al., 2010, 2011; Hamers et al., 2006; Suzuki et al., 2013; Vinggaard et al., 2008; Zhang et al., 2011; Zwart et al., 2017	Hamers et al., 2006
4-tert-Octylphenol	27193-28-8	1.82	0.65	Araki et al., 2005; Vinggaard et al., 2008	Vinggaard et al., 2008
Methyl parathion	298-00-0	0.36	0.24	Aït-Aïssa et al., 2010; Vinggaard et al., 2008	Vinggaard et al., 2008
Methoxychlor	72-43-5	0.47	0.45	Aït-Aïssa et al., 2010; Araki et al., 2005; Vinggaard et al., 2008	Araki et al., 2005
β-Endosulfan	33213-65-9	0.96	0.69	Aït-Aïssa et al., 2010; Lemaire et al., 2004; Vinggaard et al., 2008	Vinggaard et al., 2008
Chlordane	57-74-9	1.99	1.99	Lemaire et al., 2004	Lemaire et al., 2004
Cypermethrin	52315-07-8	3.96	0.36	Vinggaard et al., 2008; Xu et al., 2008	Vinggaard et al., 2008
Dieldrin	60-57-1	0.50	0.19	Lemaire et al., 2004; Vinggaard et al., 2008	Vinggaard et al., 2008
Cyfluthrin	68359-37-5	0.04	0.04	Du et al., 2010	Du et al., 2010
λ-Cyhalothrin	91465-08-6	5.40	0.52	Du et al., 2010; Orton et al., 2011	Du et al., 2010
Alachlor	15972-60-8	1.03	1.03	Vinggaard et al., 2008	Vinggaard et al., 2008
Parathion	56-38-2	0.02	0.02	Xu et al., 2008	Xu et al., 2008
2-Phenylphenol	90-43-7	1.52	1.52	Orton et al., 2011	Orton et al., 2011

7.3 Retrieving exposure data

All 231 substances on the bucket list with hazard data were screened for availability of HBM data by reviewing reports published from relevant public organizations such as EFSA, ECHA, and US EPA, followed by a systematic literature search in PubMed. Studies were selected with the following inclusion criteria: 1) articles published in English during the last 10 years; 2) women within childbearing age as target population, and 3) studies on European populations. When no HBM data fulfilling these criteria were available, articles from other Western populations (USA or Canada) were considered. Some additional steps were applied for selected compounds. For

PCBs, only articles published during the last 3 years were searched due to the high number of references found. For compounds without any available information, such as PCB 49, PCB 66, PCB 74, methyl parathion and 4-tert-octylphenol, two reports published by NHANES were included (CDC, 2019).

HBM data concentrations were prioritized in the following order if available: geometric mean (GM), median/P50, mean, and finally concentrations closest to median values available. If no concentrations above the limit of detection/quantification were measured, the study was excluded from calculation of mean values. Four substances, namely bitertanol, fenhexamid, BDE85, and 4´-OH-BDE-17, were omitted from the refined list, since none of the identified studies reported concentrations above limit of detection.

The HBM data was further curated to focus on pregnant women, mothers, or women in the fertile age defined as between 18-45 yrs. If data from these study populations were not available for a substance, the age range was expanded to include younger and elder women. Secondly, the HBM data set was curated to ensure that study populations included recruitment after year 2000. For one substance, namely heptachlor epoxide, only one older study was available, which was included. For persistent substances blood/serum/plasma concentrations were used for internal exposure estimates if available. If not, alternative matrices were used such as breast milk. For non-persistent substances urine concentrations were used for calculation of internal exposure estimates if available.

Urine concentrations were normalized to creatinine concentrations of 1.30 g/L, which is a mean of urine concentrations across all ethnic groups of females between 20-39 years from the NHANES cohort 1988-1994 (Barr et al., 2005). Blood, serum and plasma levels measured in the lipid fraction were converted to concentrations per L matrix using a factor of 6.19 g lipid/L blood. For conversion of levels in breast milk fat to levels in breast milk a factor of 0.3 g lipid/L milk was used.

For the final 61 substances for which both hazard and HBM data were identified, 15 of them had HBM data in blood/serum/plasma, 19 of them had HBM breast milk levels and 27 had HBM urinary levels. Table 7.2 lists the mean and max HBM data for each substance. In order to conduct a mixture risk assessment, calculation of risk quotients was needed and therefore we needed to estimate blood levels of substances that had HBM data in breast milk or urine.

Table 7.2: Overview of HBM levels of all substances, for which also hazard data was identified. Values are given for levels measured in the original matrix. CAS no. can be found in Table 7.1.

	Mean HBM values (µM)	Max HBM value (µM)	References used for mean HBM values	Reference used for max HBM values
Plasma/serum/blood/cord blood				
PCB 138	5.00E-04	1.29E-03	Berghuis et al., 2018; Bjerregaard-Olesen et al., 2017; Llop et al.,	Berghuis et al., 2018; Ruel et al., 2019
PCB 49	2.61E-05	2.61E-05	CDC, 2019	CDC, 2019
PCB 156	7.39E-05	1.97E-04	Berghuis et al., 2018; Bjerregaard-Olesen et al., 2017; Nøst et al.,	Berghuis et al., 2018; Ruel et al., 2019
PCB 105	7.59E-05	7.59E-05	Berghuis et al., 2018; Ruel et al., 2019	Berghuis et al., 2018; Ruel et al., 2019
PCB 118	4.80E-03	2.31E-02	Berghuis et al., 2018; Bjerregaard-Olesen et al., 2017; Nøst et al.,	Valvi et al., 2017
PCB 66	3.18E-05	3.18E-05	CDC, 2019	CDC, 2019
PCB 74	2.00E-04	3.35E-04	CDC, 2019	CDC, 2019
BDE 99	1.54E-06	1.54E-06	Caspersen et al., 2016; Lopez-Espinosa et al., 2015	Lopez-Espinosa et al., 2015
BDE 47	2.17E-05	2.56E-05	Caspersen et al., 2016; Lopez-Espinosa et al., 2015	Lopez-Espinosa et al., 2015
BDE 28	3.04E-06	3.04E-06	Caspersen et al., 2016	Caspersen et al., 2016
BDE100	5.48E-07	5.48E-07	Caspersen et al., 2016	Caspersen et al., 2016
β-Endosulfan	4.10E-03	4.91E-03	Mariscal-Arcas et al., 2010; Monteagudo et al., 2016	Monteagudo et al., 2016
p,p′-DDE	6.92E-05	1.04E-04	Ibarluzea et al., 2011; Mørck et al., 2014	Ibarluzea et al., 2011
p,p-DDT	5.24E-05	5.24E-05	Mørck et al., 2014	Mørck et al., 2014
Perfluorooctane sulphonate	2.62E-02	5.90E-02	Inoue et al., 2019; Manzano-Salgado et al., 2019; Skogheim et al., 2020; Wikström et al., 2020	Inoue et al., 2019
Breast milk				
PCB 157	7.49E-04	1.12E-03	Antignac et al., 2016	Antignac et al., 2016
PCB 114	4.69E-04	6.21E-04	Antignac et al., 2016	Antignac et al., 2016
PCB 167	1.08E-03	1.56E-03	Antignac et al., 2016	Antignac et al., 2016

	Mean HBM values (µM)	Max HBM value (µM)	References used for mean HBM values	Reference used for max HBM values
PCB 28	2.56E-06	4.08E-06	Antignac et al., 2016	Antignac et al., 2016
PCB 126	3.51E-05	4.32E-05	Antignac et al., 2016	Antignac et al., 2016
Benzo[a]pyrene	1.84E-07	2.02E-07	Pulkrabova et al., 2016	Pulkrabova et al., 2016
Benzo[j]fluoran-thene	1.31E-07	1.31E-07	Pulkrabova et al., 2016	Pulkrabova et al., 2016
Benzo[k]fluoran-thene	1.96E-07	2.26E-07	Pulkrabova et al., 2016	Pulkrabova et al., 2016
Chrysene	3.15E-07	3.55E-07	Pulkrabova et al., 2016	Pulkrabova et al., 2016
Fluoranthene	3.11E-06	3.55E-06	Pulkrabova et al., 2016	Pulkrabova et al., 2016
Pyrene	1.11E-06	1.22E-06	Pulkrabova et al., 2016	Pulkrabova et al., 2016
o,p-DDT	3.50E-07	5.08E-07	Damgaard et al., 2006; Shen et al., 2007	Shen et al., 2007
o,p-DDE	6.60E-08	8.49E-08	Damgaard et al., 2006; Shen et al., 2007	Shen et al., 2007
p,p'-DDD	3.84E-07	4.50E-07	Damgaard et al., 2006; Shen et al., 2007	Shen et al., 2007
TBBPA	4.78E-08	9.38E-08	Harrad & Abdallah, 2015	Harrad & Abdallah, 2015
Methoxychlor	3.56E-07	3.56E-07	Damgaard et al., 2006	Damgaard et al., 2006
Chlordane	8.05E-08	8.05E-08	Damgaard et al., 2006	Damgaard et al., 2006
Dieldrin	3.27E-06	4.36E-06	Antignac et al., 2016; Damgaard et al., 2006; Shen et al., 2007	Antignac et al., 2016
Procymidone	9.50E-07	9.50E-07	Brucker-Davis et al., 2010	Brucker-Davis et al., 2010
Urine				
Bisphenol A	1.11E-02	3.35E-02	Casas et al., 2011; Covaci et al., 2015; Dereumeaux et al., 2016; Frederiksen et al., 2014; Kasper-	Ye et al., 2009
Bisphenol F	5.14E-03	2.35E-02	Derakhshan et al., 2019; Peinado et al., 2020; Philips et al., 2018,	Sanchis et al., 2019
4-tert-Octylphenol	4.97E-03	5.54E-03	CDC, 2019	CDC, 2019

	Mean HBM values (µM)	Max HBM value (µM)	References used for mean HBM values	Reference used for max HBM values
Butyl paraben	1.15E-02	6.49E-02	Bellavia et al., 2019; Egeghy et al., 2012; Fisher et al., 2017; Frederiksen, Nielsen, et al., 2013; Geer et al., 2017; Larsson et al.,	Larsson et al., 2014
Ethyl paraben	1.84E-02	4.09E-02	Fisher et al., 2017; Frederiksen, Aksglaede, et al., 2013; Geer et al., 2017; Larsson et al., 2014; Messerlian et al., 2018	Messerlian et al., 2018
Methyl paraben	7.64E-01	2.11E+00	Bellavia et al., 2019; Egeghy et al., 2012; Fisher et al., 2017; Frederiksen, Aksglaede, et al., 2013; Geer et al., 2017; Larsson et	Geer et al., 2017
n-Propyl paraben	1.31E-01	3.79E-01	Bellavia et al., 2019; Egeghy et al., 2012; Fisher et al., 2017; Frederiksen, Aksglaede, et al., 2013; Geer et al., 2017; Larsson et	Geer et al., 2017
Butylbenzyl phthalate	1.75E-01	3.27E-01	Casas et al., 2011; Myridakis et al., 2015; Philippat et al., 2012; Ye et	Myridakis et al., 2015
Dibutyl phthalate	1.11E-01	2.90E-01	Casas et al., 2011; Den Hond et al., 2015; Frederiksen et al., 2013; Kasper-Sonnenberg et al., 2012; Myridakis et al., 2015; Philippat et	Ye et al., 2008
Dicyclohexyl phthalate	9.99E-03	9.99E-03	Zeman et al., 2013	Zeman et al., 2013
Diethyl phthalate	4.01E-01	1.30E+00	Casas et al., 2011; Den Hond et al., 2015; Frederiksen et al., 2013; Kasper-Sonnenberg et al., 2012; Myridakis et al., 2015; Philippat et	Ye et al., 2008
Diethylhexyl phthalate	4.22E-02	1.45E-01	Casas et al., 2011; Frederiksen et al., 2013; Kasper-Sonnenberg et al., 2012; Myridakis et al., 2015;	Zeman et al., 2013
Diisobutyl phthalate	1.44E-01	3.21E-01	Casas et al., 2011; Den Hond et al., 2015; Frederiksen et al., 2013; Kasper-Sonnenberg et al., 2012; Myridakis et al., 2015; Philippat et	Zeman et al., 2013
Dimethyl phthalate	6.36E-03	7.36E-03	Kasper-Sonnenberg et al., 2012; Ye et al., 2009	Kasper-Sonnenberg et al., 2012
1-Aminopyrene	2.39E-03	2.75E-03	Jerzynska et al., 2017	Jerzynska et al., 2017
1-Hydroxypyrene	1.72E-03	2.38E-03	Llop et al., 2008; Polanska et al., 2014; Polańska et al., 2011	Polańska et al., 2011
Methyl parathion	1.13E-02	3.42E-02	CDC, 2019	CDC, 2019
Cypermethrin	4.83E-02	4.83E-02	Dalsager et al., 2019	Dalsager et al., 2019

	Mean HBM values (µM)	Max HBM value (µM)	References used for mean HBM values	Reference used for max HBM values
Cyfluthrin	4.97E-02	4.97E-02	Dalsager et al., 2019	Dalsager et al., 2019
λ-Cyhalothrin	5.34E-04	5.34E-04	Dalsager et al., 2019	Dalsager et al., 2019
Alachlor	1.30E-04	1.48E-04	Chevrier et al., 2011, 2014	Chevrier et al., 2014
Parathion	1.59E-03	1.96E-03	Bravo et al., 2020	Bravo et al., 2020
2-Phenylphenol	6.87E-04	6.87E-04	Frederiksen et al., 2013	Frederiksen et al., 2013
Prochloraz	1.40E-03	3.28E-03	Bravo et al., 2020; Castorina et al., 2010; Pollack et al., 2016, 2018	Bravo et al., 2020
Triclosan	5.25E-02	8.84E-02	Berghuis et al., 2018; Frederiksen, Nielsen, et al., 2013; Llop et al.,	Berghuis et al., 2018; Ruel et al., 2019
Benzophenone-3	1.74E-02	2.91E-02	Frederiksen, Nielsen, et al., 2013; Nakiwala et al., 2018; Philippat et	Sakhi et al., 2018
Zearalenone	3.14E-04	3.14E-04	Fleck et al., 2016	Fleck et al., 2016

7.4 Estimating human blood levels from breast milk levels for lipophilic compounds

For many persistent chemicals, concentrations in breast milk are assumed to mirror concentrations in blood lipids (Aylward et al., 2003). We conducted a systematic literature search to identify studies reporting conversion factors (CFs) between breast milk and the blood matrix for these substances.

Some studies (Aylward et al., 2003; DeKoning & Karmaus, 2000; LaKind et al., 2009; Mannetje et al., 2012; Todaka et al., 2010, 2011; Wittsiepe et al., 2007) measured single PCB levels in blood and breast milk, however, the majority of studies focused on the sum of PCBs or several of the most common PCBs. An overall evaluation of these studies showed that the CFs fluctuated around 1, and thus we used this value for all PCBs.

For DDT, levels are often 6-7 times higher in mothers' breast milk than in her blood (Wolff, 1983). Mes et al. (Mes et al., 1984) determined the ratio between whole blood and whole milk of 16 women at day 7 post-partum for chlordane to be 0.2, although the authors did not identify a significant correlation between the level in the blood and in the milk. The ratio of dieldrin between mother's blood and milk is around 0.5 - 3 according to WHO environmental health criteria (WHO, 1989), therefore we set a mean CF of dieldrin to 1.75.

For one compound, namely methoxychlor, only animal data was available. Chapin et al. (Chapin et al., 1997) conducted experiments in dams to study the distribution of methoxychlor in plasma and milk and found a roughly constant relationship with methoxychlor dose level, although statistical analysis was not performed. Although human levels of methoxychlor are much lower than the

lowest exposure in the animal, we used CF for the lowest dose level, which was 4.05. For procymidone it was not possible to find any CF, thus we estimated it to be 1, assuming equilibrium between blood and breast milk for lipophilic substances (Marchitti et al., 2017). In general, few studies have investigated the correlation between breast milk levels and blood levels of these pesticides in humans.

Several PAHs were measured in both maternal serum and breast milk of 20 volunteers (Tsang et al., 2011), and found CFs for benzo[k]fluoranthene, chrysene and pyrene of 0.24, 0.86 and 0.82, respectively. For the remaining PAHs, for which no individual human values were reported, we used a CF of 0.74, as this was the CF of total PAHs determined in this study.

One study (Cariou et al., 2008) examined concentrations of TBBPA in maternal serum and breast milk of 93 volunteers and presented a CF between maternal serum and breast milk of approximately 1.6.

A summary of CFs for each compound with corresponding estimated blood levels in breast milk is given in Table 7.3, together with the converted HBM data for blood.

Table 7.3: Conversion of breast milk levels to blood concentrations for lipophilic compounds using the given conversion factors

	HBM matrix	Conversion factor (CF)	References for CF	Mean blood level (µM)	Max blood level (µM)
PCB 157	Blood/Milk	1.0	Aylward et al., 2003; DeKoning & Karmaus, 2000; LaKind et al., 2009; Mannetje et al., 2012; Todaka et al., 2010, 2011; Wittsiepe et al., 2007	7.49E-04	1.12E-03
PCB 114	Blood/Milk	1.0		4.69E-04	6.21E-04
PCB 167	Blood/Milk	1.0		1.08E-03	1.56E-03
PCB 28	Blood/Milk	1.0		2.56E-06	4.08E-06
PCB 126	Blood/Milk	1.0		3.51E-05	4.32E-05
Benzo[a]pyrene	Serum/Milk	0.7	Tsang et al., 2011	1.36E-07	1.50E-07
Benzo[j]fluoranthene	Serum/Milk	0.7		9.68E-08	9.68E-08
Benzo[k]fluoranthene	Serum/Milk	0.2		4.71E-08	5.42E-08
Chrysene	Serum/Milk	0.9		2.71E-07	3.05E-07
Fluoranthene	Serum/Milk	0.7		2.31E-06	2.62E-06
Pyrene	Serum/Milk	0.8		9.06E-07	9.97E-07
o,p-DDT	Blood/Milk	0.2	Wolff, 1983	5.83E-08	8.46E-08
o,p-DDE	Blood/Milk	0.2		1.10E-08	1.41E-08
p,p'-DDD	Blood/Milk	0.2		6.41E-08	7.50E-08
TBBPA	Serum/Milk	1.6	Cariou et al., 2008	7.65E-08	1.50E-07
Methoxychlor	Plasma/Milk	4.1	Chapin et al., 1997	1.44E-06	1.44E-06
Chlordane	Blood/Milk	0.2	Mes et al., 1984	1.61E-08	1.61E-08
Dieldrin	Blood/Milk	1.8	WHO, 1989	5.73E-06	7.62E-06
Procymidone	Serum/Milk	1.0	Marchitti et al., 2017	9.50E-07	9.50E-07

Substances that have a relatively short biological half-life mostly have HBM data measured in urine. For calculating risk quotients, we needed blood levels of these substances and therefore had to estimate these from urine levels. This introduces some uncertainty as several assumptions have

to be made in order to perform these estimations. Blood levels were estimated by using these two equations. First, we estimated the daily intake (DI) of a substance by using equation 2 (Koch et al., 2007):

$$\mathrm{DI}\left[{}^{mol}/_{kg_{bw}} \cdot day\right] = \frac{\left({}^{conc_{metabolite}\left[\frac{g}{L}\right]}/_{MW_{metabolite}\left[\frac{g}{mol}\right]}\right)}{Creatinine\left[\frac{g_{crt}}{L}\right]} \cdot \frac{CE\left[\frac{g_{crt}}{kg_{bw}}{day}\right]}{F_{UE}} = \frac{UE[mol/g_{crt}] \cdot CE\left[{}^{g_{crt}}/_{kg_{bw}}/day\right]}{F_{UE}} \qquad Eq.2$$

Where:

Molecular weight of creatinine 113.12 g/mol was applied.

UE is the molar urinary excretion of the respective metabolite(s)

CE is the creatinine excretion rate normalized by body weight. We set CE to 0.023 g/kg/day for a pregnant woman according to Lioy et al. (Lioy et al., 2015).

The molar fraction F_{UE} (%) describes the molar ratio between the number of metabolite(s) excreted in urine and the amount of parent compound taken up.

Next, we calculated the blood concentrations by applying a simplified one-compartment toxicokinetic model:

$$C_p[mol/l] = \mathrm{DI}\left[{}^{mol}/_{kg_{bw}} \cdot day\right] \cdot \frac{{}^{t\,1/2}/_{day}}{0{,}693} \cdot \frac{1}{V_d[L/kg_{BW}]} \qquad Eq.3$$

This model assumes total bioavailability and resorption from the intestine as well as steady state levels (Fromme et al., 2007) to calculate blood plasma levels of chemicals. C_p is blood plasma concentration after exposure to dose X, $t_{1/2}$ is the biological half-life of the substance and V_d is the apparent volume of distribution.

To convert HBM urinary levels to blood plasma levels by using *Eq.2* and *Eq.3*, we collected all the parameters needed for the calculation, prioritizing human data. We had to use F_{UE} obtained from animal studies for 2 substances (prochloraz and alachlor) due to lack of human data.

No studies reported F_{UE} and $t_{1/2}$ for bisphenol F, and we therefore assumed these being the same as for bisphenol A, as Rochester et al. (Rochester & Bolden, 2015) and Punt et al. (Punt et al., 2019) report that these two compounds have similar properties.

Due to lack of data for dicyclohexyl phthalate we used values for diethylhexyl phthalate (mono (2-ethylhexyl) phthalate) according to Koo et al. (Koo et al., 2002), whereas for diethyl phthalate and dimethyl phthalate values for dibutyl phthalate (monobutyl phthalate) were used according to Gennings et al. (Gennings et al., 2014).

For 4-tert-octylphenol we did not find any data and estimated F_{UE} to be 0.9 based on expert judgement.

For the two metabolites of cyfluthrin, cis- and trans-DCCA, we assumed equal excretion for cis- and trans-DCCA to be able to calculate F_{UE} for each metabolite according to Hays et al. (Hays et al., 2009). A total of 0.33 mg cis- and trans- DCCA was excreted in urine after intake of 2.6 mg cyfluthrin, which results in a mass excretion fraction of 0.127 for the sum of cis- and trans-DCCA. This mass excretion fraction was further converted to F_{UE} by using *Eq.4* (Aylward et al., 2018).

$$F_{UE}(\mathrm{molar}) = F_{UE}(\mathrm{mass}) * \frac{MW\ parent}{MW\ metabolite} \qquad Eq.4$$

For many substances measured $T_{1/2}$ and V_d values could not be found and in these cases we used computationally predicted values from the CompTox Chemicals Dashboard (NCCT, US-EPA). One exception was parathion which is not part of the dashboard. For 1-aminopyrene we used V_d for its parent compound 1-nitropyrene due to lack of data.

All estimated blood levels of compounds with HBM data from urine together with the parameters used for the conversions are given in Table 7.4.

Table 7.4: Estimated mean and max blood concentrations of compounds measured in urine including the parameters used for the conversions from urine to blood

Chemical (main metabolite)	Mean blood level (µM)	Max blood level (µM)	F_{UE}	$T_{1/2}$ (hr)	V_d (L/kg)	Reference (F_{UE})	Reference ($T_{1/2}$)	Reference (V_d)
Bisphenol A	1.51E-05	4.55E-05	1.000	6.4	5.0	Koch et al., 2012	Thayer et al., 2015	NCCT,US-EPA
Bisphenol F	1.20E-05	5.46E-05	1.000	3.7	1.7	Punt et al., 2019; Rochester & Bolden, 2015	NCCT,US-EPA	NCCT,US-EPA
4-tert-Octylphenol	4.82E-05	5.37E-05	0.900	23.1	2.8	N/A[b]	NCCT,US-EPA	NCCT,US-EPA
Butyl paraben	7.10E-04	4.02E-03	0.056	7.0	2.2	Moos et al., 2016	SCCS (Scientific Committee on Consumer Safety), 2013	NCCT,US-EPA
Ethyl paraben	7.76E-04	1.73E-03	0.137	4.9	0.9	Moos et al., 2017	N/A[c]	NCCT,US-EPA
Methyl paraben	5.37E-02	1.49E-01	0.174	6.9	0.6	Moos et al., 2016	Moos et al., 2016	NCCT,US-EPA
n-Propyl paraben	3.17E-03	9.14E-03	0.097	2.9	1.3	Moos et al., 2017	Shin et al., 2019	NCCT,US-EPA
Butylbenzyl phthalate (monobutyl phthalate)	1.44E-01	2.69E-01	0.060[a]	11.9	0.4	Anderson et al., 2001	NCCT,US-EPA	NCCT,US-EPA
Dibutyl phthalate (mono-n-butyl phthalate)	1.61E-03	4.22E-03	0.690[a]	22.4	3.0	Anderson et al., 2001	NCCT,US-EPA	NCCT,US-EPA
Dicyclohexyl phthalate (monocyclohexyl phthalate)	2.12E-03	2.12E-03	0.073[a,d]	31.0	2.8	Koo et al., 2002	NCCT,US-EPA	NCCT,US-EPA

Chemical (main metabolite)	Mean blood level (µM)	Max blood level (µM)	F_{UE}	$T_{1/2}$ (hr)	V_d (L/kg)	Reference (F_{UE})	Reference ($T_{1/2}$)	Reference (V_d)
Diethyl phthalate (monoethyl phthalate)	9.50E-04	3.08E-03	0.690[a,e]	1.8	1.3	Gennings et al., 2014	NCCT,US-EPA	NCCT,US-EPA
Diethylhexyl phthalate (mono(2-ethylhexyl)phthalate)	1.32E-04	4.79E-04	0.073[a]	4.3	26.5	Koch et al., 2004	Kessler et al., 2012	NCCT,US-EPA
Di-isobutyl phthalate (monoisobutyl phthalate)	2.98E-03	6.67E-03	0.710[a]	6.3	0.6	Koch et al., 2012	NCCT,US-EPA	NCCT,US-EPA
Dimethyl phthalate (mono-methyl phthalate)	1.37E-04	1.58E-04	0.690[a,e]	21.6	1.7	Gennings et al., 2014	NCCT,US-EPA	NCCT,US-EPA
1-Aminopyrene	2.05E-04	2.35E-04	0.341	74.8	2.7	Toriba et al., 2007	NCCT,US-EPA	NCCT,US-EPA
1-Hydroxypyrene	8.39E-04	1.16E-03	0.037	4.4	0.3	Viau et al., 1995	Motorykin et al., 2016	NCCT,US-EPA
Methyl parathion (p-Nitrophenol)	4.45E-04	1.35E-03	0.511[a]	11.6	1.2	Morgan et al., 1977	NCCT,US-EPA	NCCT,US-EPA
Cypermethrin (cis-DCCA)	3.66E-03	3.66E-03	0.110[a]	16.5	4.2	Ratelle et al., 2015	Woollen et al., 1992	NCCT,US-EPA
Cyfluthrin (trans-DCCA)	4.63E-03	4.63E-03	0.132[a]	16.0	2.9	Hays et al., 2009	Williams et al., 2003	NCCT,US-EPA
λ-Cyhalothrin (3-PBA)	1.72E-05	1.72E-05	0.251[a]	17.8	4.9	Aylward et al., 2018	NCCT,US-EPA	NCCT,US-EPA
Alachlor	5.94E-07	6.79E-07	0.385	3.9	2.3	Wilson & Hall, 1986	NCCT,US-EPA	NCCT,US-EPA
Parathion	1.62E-06	1.99E-06	0.390	8.0	21.5	Katsikantami et al., 2019	Safe Work Australia, 2020	(Eyer et al., 2003)
2-Phenylphenol	5.82E-05	5.82E-05	0.005	0.8	2.0	Timchalk et al., 1998	Timchalk et al., 1998	NCCT,US-EPA

Chemical (main metabolite)	Mean blood level (µM)	Max blood level (µM)	F_{UE}	$T_{1/2}$ (hr)	V_d (L/kg)	Reference (F_{UE})	Reference ($T_{1/2}$)	Reference (V_d)
Prochloraz	4.38E-05	1.03E-04	0.410	21.7	1.8	JMPR, 2001	NCCT,US-EPA	NCCT,US-EPA
Triclosan	8.88E-04	3.14E-04	0.540	21.0	2.5	Sandborgh-Englund et al., 2006	Sandborgh-Englund et al., 2006	NCCT,US-EPA
Benzophenone-3	4.33E-03	3.13E-02	0.037	8.0	0.9	Gonzalez et al., 2006	Kasichayanula et al., 2007	NCCT,US-EPA
Zearalenone	1.23E-05	1.23E-05	0.150	11.9	2.2	Metzler et al., 2010; Mirocha, 1981	Mukherjee et al., 2014	NCCT,US-EPA

a: F_{UE} is given for the main metabolites shown in brackets; b: The F_{UE} of 4-tert-octylphenol is set to 0.9 based on expert judgement; c: Due to lack of data, we used the mean value for methyl paraben and n-propyl paraben; d: Assumed to be the same as diethylhexyl phthalate (mono (2-ethylhexyl) phthalate); e: Same value as for dibutyl phthalate (monobutyl phthalate)

7.5 Risk quotients and mixture risk assessment of antiandrogenic compounds

For calculation of a mixture risk quotient, we used the IC_{10} values for hAR antagonism as the hazard data and either the mean or the maximum exposure defined by the retrieved HBM data. In Figure 7.1 is illustrated the RQs based on mean HBM data and IC_{10} potency values. None of the compounds has got a RQ above 1 and there seem to be no obvious major mixture risk drivers. Rather it seems as if very many compounds contribute each with a small factor to the mixture risk quotient (or Hazard Index). The mixture risk quotient based on the mean HBM values and available data ends up being 0.4. The top 15 compounds include some pesticides (pyrethroids and persistent ones: β-endosulfan and p,p-DDE), phthalates (BBP, DBP, DIBP), PCBs (118, 138, 157), PFOS, the UV filter BP-3, triclosan, a paraben (MP) and a PAH metabolite (Hydroxy--pyrene).

When the calculation is based on the maximum HBM levels corresponding to a worst-case exposure scenario, the mixture risk quotient ends up being 1.1 (Figure 7.2). The top 15 compounds on the list include more or less the same compounds as when the mixture risk calculation was based on the mean HBM levels, but with a little different ranking. The only exception is that methyl-parathion is now replacing triclosan in the top 15 ranking, which has now moved a couple of numbers down.

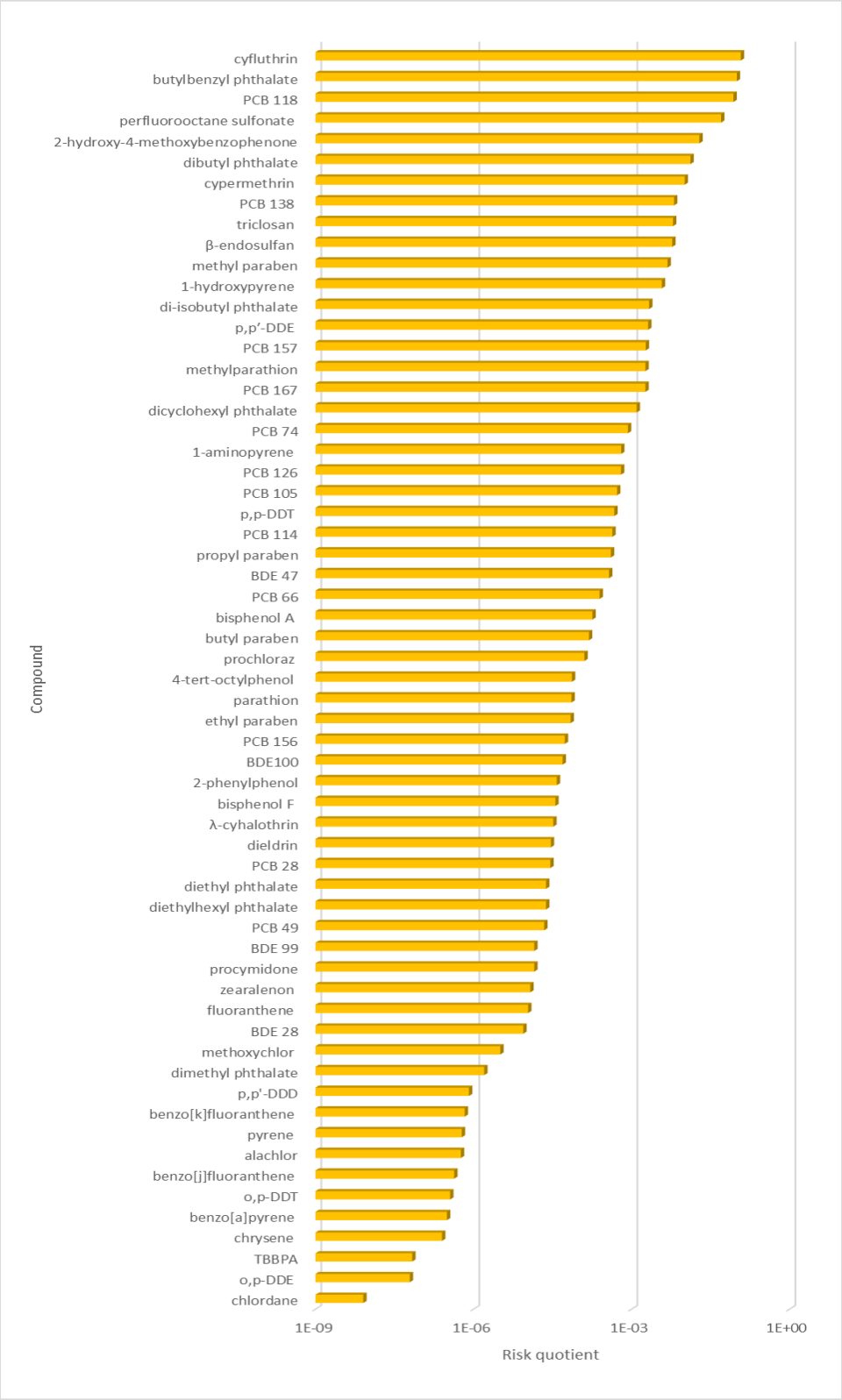


Figure 7.1: Risk quotients (RQ) of 61 substances based on the mean human exposure levels (in µM) divided by the IC_{10} for hAR antagonism (in µM) (HBM_{mean}/IC_{10}) ranked according to the magnitude of the RQ

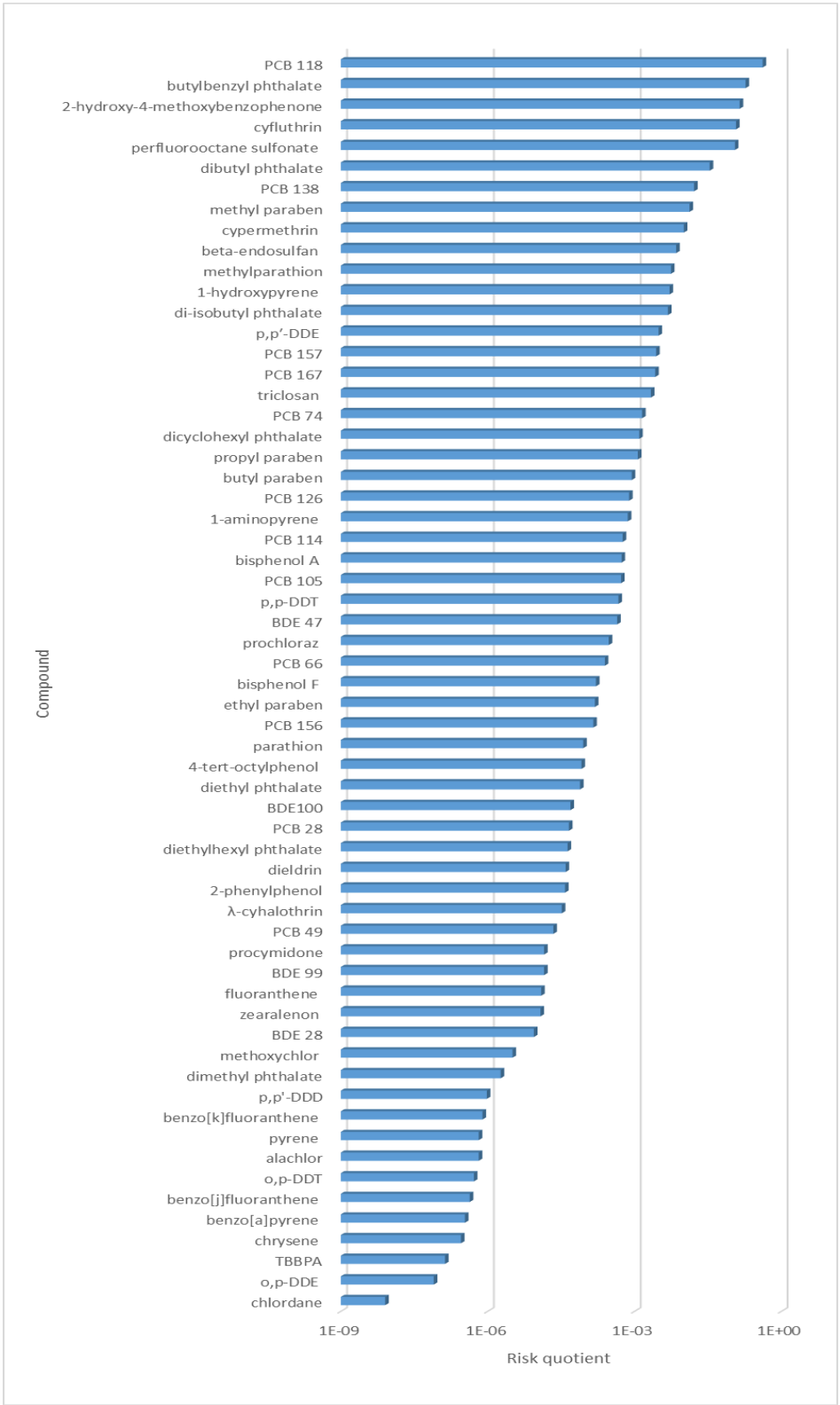


Figure 7.2: Risk quotients (RQ) of 61 substances based on the maximum human exposure levels (in µM) divided by the IC_{10} for hAR antagonism (in µM) ($HBMmean/IC_{10}$) ranked according to the magnitude of the RQ.

7.6 Discussion

Assumptions of our approach

Several assumptions were taken for this mixture risk assessment including:

1. Male reproductive health disorders arise during fetal life and the dominating mechanism of action for adverse effects on male reproductive health is AR antagonism.
2. A maternal blood level of a substance corresponding to a nominal IC_{10} for AR antagonism will translate into a small, non-detectable shortening of anogenital distance in the male fetus that contributes to the total AR blocking and appearance of a shortened anogenital distance.
3. Substances with AR antagonistic effects belong by default to a common assessment group.
4. It is possible to estimate human blood levels from HBM data in urine or breast milk within the limits of acceptable uncertainty.
5. A risk quotient for each chemical can be calculated based on an in vitro response and a HBM level in blood (in µM).
6. Antiandrogenic chemicals are acting additively and we can calculate a Mixture Risk Quotient (equivalent to Hazard Index) that gives an indication of the mixture effect based on available data.

Strengths and limitations of our approach

The advantages of using in vitro hazard data for mixture risk assessment is 1) that human data can be used circumventing the uncertainties in using hazard data from rodents, 2) more compounds can be included in the mixture risk assessment due to larger availability of in vitro compared to in vivo hazard data and 3) no grouping considerations are needed as the in vitro active compounds by default belong to the same common assessment group. On the other hand the disadvantages are that 1) only compounds acting by this specific mechanism of action are included in the mixture risk assessment and other unknown modes of action can also lead to male reproductive health disorders (e.g. 2,3,7,8-TCDD and paracetamol), 2) that uncertainties exist in extrapolating in vitro potencies to in vivo outcomes in the human fetus. Especially, the last point is critical as we do not know the magnitude of the intracellular concentrations of the chemicals in the AR assays and therefore do not know the ‘true’ potencies of the compounds.

Concerning the exposure data the advantages of using HBM data for mixture risk assessment are that 1) internal exposure levels are based on measured HBM data reflecting real-life exposure, 2) data are aggregated and cover all exposure sources. The disadvantages are that 1) mean HBM values at population level are used, which do not represent individual exposure levels 2) for many compounds urinary or breast milk levels are measured thereby causing the need of identifying or estimating toxicokinetic parameters for each substance in order to convert the levels to blood levels. This conversion of HBM data is pragmatic and encumbered with uncertainty.

Generally, great uncertainties exist in chemical mixture risk assessment and the question is if the uncertainties highlighted above will lead to an over- or an underestimation of risk. For most of the factors it is not known whether they will pull the assessment in one direction or another. However, the factors that most likely will lead to an overestimation of the mixture risk quotient are the use of maximum HBM values, which will not reflect exposure of the general consumer under normal circumstances. Moreover, mean HBM data from studies covering various populations from various

geographical areas and all seasons will not reflect individual exposure and may lead to an over- or underestimation as well.

Factors that lead to an underestimation of the mixture risk quotient include the fact that only compounds with one mechanism of action are included in the MRA. Especially the fact that 2,3,7,8-TCDD, which does not block the AR and therefore was not included, most likely leads to an underestimation of risk, as TCDD is known to have potent adverse effects on male reproductive health. Moreover, the risk will be underestimated due to lack of either hazard or exposure data for many compounds. Initially, we identified >200 AR antagonistic compounds, but the majority was excluded due to lack of available HBM data.

Probably the most important factor leading to underestimation of risk is caused by the nature of the hazard data. Our hypothesis is that the hazard levels of the compounds are underestimated when using the nominal IC_{10} values for in vitro AR antagonism. The likelihood is high that the 'true' IC_{10} values based on internal cellular concentrations will be considerably lower. Nominal effect concentrations from in vitro toxicity assays often lead to inaccurate estimations of in vivo effective doses, because the nominal concentration poorly reflects the concentration at the molecular target in cells in vitro. This intracellular concentration is responsible for initiating effects and can be referred to as the biologically effective concentration. Chemicals can differentially distribute between in vitro assay compartments, including serum constituents in exposure medium, microtitre plastic plates, headspace and extracellular matrices. The partitioning of test chemicals to these extracellular compartments reduces the concentration at the molecular target (Gilbert et al., 2015; Proença et al., 2021). If this is the case, the 'true' IC_{10} values reflecting biological activity are in reality lower, which will lead to higher RQs and a higher mixture risk quotient. An improvement of this mixture risk assessment approach could be performed by measuring or estimating the intracellular concentrations of chemicals in the AR in vitro assays.

What is the risk?

No single substance appeared with a RQ exceeding 1 and many compounds contributed each with a small RQ to the total mixture risk quotient. Thus, there was not a clear picture of a specific chemical class that drives the mixture risk. Rather it seems as if both pesticides, PCBs, phthalates, PFOS, parabens, the UV filter BP-3, the disinfectant triclosan and PAH contribute to the mixture risk.

Only in the worst-case scenario, in which we include the maximum exposure levels, the mixture risk quotient just exceeds 1. Thus, when calculations are based on this approach it needs to be investigated further if this may reflect a potential risk. Rather we would suggest to refine our approach for future mixture risk assessments.

The calculated mixture risk quotient may provide a better understanding of whether the human chemical load based on available information will have the potential to cause adverse effects on male reproductive health via AR antagonism. Further, the approach has pinpointed that it may be small contributions from compounds from several chemical classes that together contribute markedly to the mixture risk.

7.7 References

Aït-Aïssa, S., Laskowski, S., Laville, N., Porcher, J. M., & Brion, F. (2010). Anti-androgenic activities of environmental pesticides in the MDA-kb2 reporter cell line. *Toxicology in Vitro*, *24*(7), 1979–1985. https://doi.org/10.1016/j.tiv.2010.08.014

Anderson, W. A. C., Castle, L., Scotter, M. J., Massey, R. C., & Springall, C. (2001). A biomarker approach to measuring human dietay exposure to certain phthalate diesters. *Food Additives and Contaminants*, *18*(12), 1068–1074. https://doi.org/10.1080/02652030110050113

Antignac, J. P., Main, K. M., Virtanen, H. E., Boquien, C. Y., Marchand, P., Venisseau, A., Guiffard, I., Bichon, E., Wohlfahrt-Veje, C., Legrand, A., Boscher, C., Skakkebæk, N. E., Toppari, J., & Le Bizec, B. (2016). Country-specific chemical signatures of persistent organic pollutants (POPs) in breast milk of French, Danish and Finnish women. *Environmental Pollution*, *218*, 728–738. https://doi.org/10.1016/j.envpol.2016.07.069

Araki, N., Ohno, K., Nakai, M., Takeyoshi, M., & Iida, M. (2005). Screening for androgen receptor activities in 253 industrial chemicals by in vitro reporter gene assays using AR-EcoScreen™ cells. *Toxicology in Vitro*, *19*(6), 831–842. https://doi.org/10.1016/j.tiv.2005.04.009

Araki, N., Ohno, K., Takeyoshi, M., & Iida, M. (2005). Evaluation of a rapid in vitro androgen receptor transcriptional activation assay using AR-EcoScreen™ cells. *Toxicology in Vitro*, *19*(3), 335–352. https://doi.org/10.1016/j.tiv.2004.10.008

Aylward, L. L., Hays, S. M., LaKind, J. S., & Ryan, J. J. (2003). Rapid communication: Partitioning of persistent lipophilic compounds, including dioxins, between human milk lipid and blood lipid: An initial assessment. *Journal of Toxicology and Environmental Health - Part A*, *66*(1), 1–5. https://doi.org/10.1080/15287390306460

Aylward, Lesa L., Irwin, K., St-Amand, A., Nong, A., & Hays, S. M. (2018). Screening-level Biomonitoring Equivalents for tiered interpretation of urinary 3-phenoxybenzoic acid (3-PBA) in a risk assessment context. *Regulatory Toxicology and Pharmacology*, *92*(November 2017), 29–38. https://doi.org/10.1016/j.yrtph.2017.11.002

Barr, D. B., Wilder, L. C., Caudill, S. P., Gonzalez, A. J., Needham, L. L., & Pirkle, J. L. (2005). Urinary creatinine concentrations in the U.S. population: Implications for urinary biologic monitoring measurements. *Environmental Health Perspectives*, *113*(2), 192–200. https://doi.org/10.1289/ehp.7337

Bellavia, A., Chiu, Y. H., Brown, F. M., Mínguez-Alarcón, L., Ford, J. B., Keller, M., Petrozza, J., Williams, P. L., Ye, X., Calafat, A. M., Hauser, R., & James-Todd, T. (2019). Urinary concentrations of parabens mixture and pregnancy glucose levels among women from a fertility clinic. *Environmental Research*, *168*(October 2018), 389–396. https://doi.org/10.1016/j.envres.2018.10.009

Berghuis, S. A., Van Braeckel, K. N. J. A., Sauer, P. J. J., & Bos, A. F. (2018). Prenatal exposure to persistent organic pollutants and cognition and motor performance in adolescence. *Environment International*, *121*(April), 13–22. https://doi.org/10.1016/j.envint.2018.08.030

Birkhøj, M., Nellemann, C., Jarfelt, K., Jacobsen, H., Andersen, H. R., Dalgaard, M., & Vinggaard, A. M. (2004). The combined antiandrogenic effects of five commonly used pesticides. *Toxicology and Applied Pharmacology*, *201*(1), 10–20. https://doi.org/10.1016/j.taap.2004.04.016

Bjerregaard-Olesen, C., Long, M., Ghisari, M., Bech, B. H., Nohr, E. A., Uldbjerg, N., Henriksen, T. B., Olsen, J., & Bonefeld-Jørgensen, E. C. (2017). Temporal trends of lipophilic persistent organic pollutants in serum from Danish nulliparous pregnant women 2011–2013. *Environmental Science and Pollution Research*, *24*(20), 16592–16603. https://doi.org/10.1007/s11356-017-8992-7

Bonefeld-Jørgensen, E. C., Long, M., Hofmeister, M. V., & Vinggaard, A. M. (2007). Endocrine-disrupting potential of Bisphenol A, Bisphenol A dimethacrylate, 4-n-nonylphenol, and 4-n-octylphenol in vitro: New data and a brief review. *Environmental Health Perspectives*, *115*(SUPPL1), 69–76. https://doi.org/10.1289/ehp.9368

Bravo, N., Peralta, S., Grimalt, J. O., Martínez, M. Á., Rovira, J., & Schuhmacher, M. (2020). Organophosphate metabolite concentrations in maternal urine during pregnancy. *Environmental Research*, *182*(May 2019), 109003. https://doi.org/10.1016/j.envres.2019.109003

Brucker-Davis, F., Wagner-Mahler, K., Bornebusch, L., Delattre, I., Ferrari, P., Gal, J., Boda-Buccino, M., Pacini, P., Tommasi, C., Azuar, P., Bongain, A., & Fénichel, P. (2010). Exposure to selected endocrine disruptors and neonatal outcome of 86 healthy boys from Nice area (France). *Chemosphere*, *81*(2), 169–176. https://doi.org/10.1016/j.chemosphere.2010.06.068

Cariou, R., Antignac, J. P., Zalko, D., Berrebi, A., Cravedi, J. P., Maume, D., Marchand, P., Monteau, F., Riu, A., Andre, F., & bizec, B. Le. (2008). Exposure assessment of French women and their newborns to tetrabromobisphenol-A: Occurrence measurements in maternal adipose tissue, serum, breast milk and cord serum. *Chemosphere*, *73*(7), 1036–1041. https://doi.org/10.1016/j.chemosphere.2008.07.084

Casas, L., Fernández, M. F., Llop, S., Guxens, M. M., Ballester, F., Olea, N. N., Irurzun, M. B., Rodríguez, L. S. M., Riaño, I., Tardón, A., Vrijheid, M., Calafat, A. M., Sunyer, J., INMA Project, Fernandez, M. F., Llop, S., Guxens, M. M., Ballester, F., Olea, N. N., … INMA Project. (2011). Urinary concentrations of phthalates and phenols in a population of Spanish pregnant women and children. *Environment International*, *37*(5), 858–866. https://doi.org/10.1016/j.envint.2011.02.012

Caspersen, I. H., Kvalem, H. E., Haugen, M., Brantsæter, A. L., Meltzer, H. M., Alexander, J., Thomsen, C., Frøshaug, M., Bremnes, N. M. B., Broadwell, S. L., Granum, B., Kogevinas, M., & Knutsen, H. K. (2016). Determinants of plasma PCB, brominated flame retardants, and organochlorine pesticides in pregnant women and 3 year old children in The Norwegian Mother and Child Cohort Study. *Environmental Research*, *146*, 136–144. https://doi.org/10.1016/j.envres.2015.12.020

Castorina, R., Bradman, A., Fenster, L., Barr, D. B., Bravo, R., Vedar, M. G., Harnly, M. E., Mckone, T. E., Eisen, E. A., & Eskenazi, B. (2010). Comparison of current-use pesticide and other toxicant urinary metabolite levels among pregnant women in the CHAMACOS cohort and NHANES. *Environmental Health Perspectives*, *118*(6), 856–863. https://doi.org/10.1289/ehp.0901568

CDC. (2019). Fourth National Report on Human Exposure to Environmental Chemicals [WWW Document]. *U.S. Dep. Heal. Hum. Serv*. Accessed March 20, 2020 from https://www.cdc.gov/exposurereport/index.html

Chapin, R. E., Harris, M. W., Davis, B. J., Ward, S. M., Wilson, R. E., Mauney, M. A., Lockhart, A. C., Smialowicz, R. J., Moser, V. C., Burka, L. T., & Collins, B. J. (1997). The effects of perinatal/juvenile methoxychlor exposure on adult rat nervous, immune, and reproductive system function. *Fundamental and Applied Toxicology*, *40*(1), 138–157. https://doi.org/10.1006/faat.1997.2381

Chevrier, C., Limon, G., Monfort, C., Rouget, F., Garlantézec, R., Petit, C., Durand, G., & Cordier, S. (2011). Urinary biomarkers of prenatal atrazine exposure and adverse birth outcomes in the pelagie birth cohort. *Environmental Health Perspectives*, *119*(7), 1034–1041. https://doi.org/10.1289/ehp.1002775

Chevrier, C., Serrano, T., Lecerf, R., Limon, G., Petit, C., Monfort, C., Hubert-Moy, L., Durand, G., & Cordier, S. (2014). Environmental determinants of the urinary concentrations of herbicides during pregnancy: The PELAGIE mother-child cohort (France). *Environment International*, *63*, 11–18. https://doi.org/10.1016/j.envint.2013.10.010

Christen, V., Crettaz, P., Oberli-Schrämmli, A., & Fent, K. (2010). Some flame retardants and the antimicrobials triclosan and triclocarban enhance the androgenic activity in vitro. *Chemosphere*, *81*(10), 1245–1252. https://doi.org/10.1016/j.chemosphere.2010.09.031

Christen, V., Crettaz, P., Oberli-Schrämmli, A., & Fent, K. (2012a). Antiandrogenic activity of phthalate mixtures: Validity of concentration addition. *Toxicology and Applied Pharmacology*, *259*(2), 169–176. https://doi.org/10.1016/j.taap.2011.12.021

Christen, V., Crettaz, P., Oberli-Schrämmli, A., & Fent, K. (2012b). Corrigendum to “Antiandrogenic activity of phthalate mixtures: Validity of concentration addition” Toxicology and Applied Pharmacology 259 (2012) 169-176. *Toxicology and Applied Pharmacology*, *263*(3), 402–403. https://doi.org/10.1016/j.taap.2012.07.011

Christiansen, S., Scholze, M., Dalgaard, M., Vinggaard, A. M., Axelstad, M., Kortenkamp, A., & Hass, U. (2009). Synergistic disruption of external male sex organ development by a mixture of four antiandrogens. *Environmental Health Perspectives*, *117*(12), 1839–1846. https://doi.org/10.1289/ehp.0900689

Covaci, A., Den Hond, E., Geens, T., Govarts, E., Koppen, G., Frederiksen, H., Knudsen, L. E., Morck, T. A., Gutleb, A. C., Guignard, C., Cocco, E., Horvat, M., Heath, E., Kosjek, T., Mazej, D., Tratnik, J. S., Castano, A., Esteban, M., Cutanda, F., … Koch, H. M. (2015). Urinary BPA measurements in children and mothers from six European member states: Overall results and determinants of exposure. *Environmental Research*, *141*, 77–85. https://doi.org/10.1016/j.envres.2014.08.008

Dalsager, L., Fage-Larsen, B., Bilenberg, N., Jensen, T. K., Nielsen, F., Kyhl, H. B., Grandjean, P., & Andersen, H. R. (2019). Maternal urinary concentrations of pyrethroid and chlorpyrifos metabolites and attention deficit hyperactivity disorder (ADHD) symptoms in 2-4-year-old children from the Odense Child Cohort. *Environmental Research*, *176*(June), 108533. https://doi.org/10.1016/j.envres.2019.108533

Damgaard, I. N., Skakkebæk, N. E., Toppari, J., Virtanen, H. E., Shen, H., Schramm, K. W., Petersen, J. H., Jensen, T. K., Main, K. M., Boisen, K., Chellakooty, M., Schmidt, I., Kaleva, M., & Soumi, A. (2006). Persistent pesticides in human breast milk and cryptorchidism. *Environmental Health Perspectives*, *114*(7), 1133–1138. https://doi.org/10.1289/ehp.8741

Danish Environmental Protection Agency. (2017). *Exposure of children and unborn children to selected chemical substances*. https://www2.mst.dk/Udgiv/publications/2017/04/978-87-93529-84-7.pdf

DeKoning, E. P., & Karmaus, W. (2000). PCB exposure in utero and via breast milk. A review. *Journal of Exposure Analysis and Environmental Epidemiology*, *10*(3), 285–293. https://doi.org/10.1038/sj.jea.7500090

Den Hond, E., Govarts, E., Willems, H., Smolders, R., Casteleyn, L., Kolossa-Gehring, M., Schwedler, G., Seiwert, M., Fiddicke, U., Castaño, A., Esteban, M., Angerer, J., Koch, H. M., Schindler, B. K., Sepai, O., Exley, K., Bloemen, L., Horvat, M., Knudsen, L. E., … 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

Derakhshan, A., Shu, H., Peeters, R. P., Kortenkamp, A., Lindh, C. H., Demeneix, B., Bornehag, C. G., & Korevaar, T. I. M. (2019). Association of urinary bisphenols and triclosan with thyroid function during early pregnancy. *Environment International*, *133*(June), 105123. https://doi.org/10.1016/j.envint.2019.105123

Dereumeaux, C., Saoudi, A., Pecheux, M., Berat, B., de Crouy-Chanel, P., Zaros, C., Brunel, S., Delamaire, C., le Tertre, A., Lefranc, A., Vandentorren, S., & Guldner, L. (2016). Biomarkers of exposure to environmental contaminants in French pregnant women from the Elfe cohort in 2011. *Environment International*, *97*, 56–67. https://doi.org/10.1016/j.envint.2016.10.013

Di Paolo, C., Kirchner, K., Balk, F. G. P., Muschket, M., Brack, W., Hollert, H., & Seiler, T. B. (2016). Downscaling procedures reduce chemical use in androgen receptor reporter gene assay. *Science of the Total Environment*, *571*, 826–833. https://doi.org/10.1016/j.scitotenv.2016.07.059

Du, G., Shen, O., Sun, H., Fei, J., Lu, C., Song, L., Xia, Y., Wang, S., & Wang, X. (2010). Assessing hormone receptor activities of pyrethroid insecticides and their metabolites in reporter gene assays. *Toxicological Sciences*, *116*(1), 58–66. https://doi.org/10.1093/toxsci/kfq120

Egeghy, P. P., Judson, R., Gangwal, S., Mosher, S., Smith, D., Vail, J., & Cohen Hubal, E. A. (2012). The exposure data landscape for manufactured chemicals. *Science of the Total Environment*, *414*, 159–166. https://doi.org/10.1016/j.scitotenv.2011.10.046

Engel, A., Buhrke, T., Imber, F., Jessel, S., Seidel, A., Völkel, W., & Lampen, A. (2017). Agonistic and antagonistic effects of phthalates and their urinary metabolites on the steroid hormone receptors ERα, ERβ, and AR. *Toxicology Letters*, *277*(March), 54–63. https://doi.org/10.1016/j.toxlet.2017.05.028

Ermler, S., Scholze, M., & Kortenkamp, A. (2010). The sensitivity of the MDA-kb2 cell in vitro assay in detecting anti-androgenic chemicals - Identification of sources of variability and estimation of statistical power. *Toxicology in Vitro*, *24*(6), 1845–1853. https://doi.org/10.1016/j.tiv.2010.05.007

Ermler, S., Scholze, M., & Kortenkamp, A. (2011). The suitability of concentration addition for predicting the effects of multi-component mixtures of up to 17 anti-androgens with varied structural features in an in vitro AR antagonist assay. *Toxicology and Applied Pharmacology*, *257*(2), 189–197. https://doi.org/10.1016/j.taap.2011.09.005

Eyer, F., Meischner, V., Kiderlen, D., Thiermann, H., Worek, F., Haberkorn, M., Felgenhauer, N., Zilker, T., & Eyer, P. (2003). Human Parathion Poisoning. *Toxicological Reviews*, *22*(3), 143–163. https://doi.org/10.2165/00139709-200322030-00003

Fisher, M., Macpherson, S., Braun, J. M., Hauser, R., Walker, M., Feeley, M., Mallick, R., Bérubé, R., & Arbuckle, T. E. (2017). Paraben Concentrations in Maternal Urine and Breast Milk and Its Association with Personal Care Product Use. *Environmental Science and Technology*, *51*(7), 4009–4017. https://doi.org/10.1021/acs.est.6b04302

Fleck, S. C., Churchwell, M. I., Doerge, D. R., & Teeguarden, J. G. (2016). Urine and serum biomonitoring of exposure to environmental estrogens II: Soy isoflavones and zearalenone in

pregnant women. *Food and Chemical Toxicology*, *95*, 19–27. https://doi.org/10.1016/j.fct.2016.05.021

Frederiksen, H., Aksglaede, L., Sorensen, K., Nielsen, O., Main, K. M., Skakkebaek, N. E., Juul, A., & Andersson, A. M. (2013). Bisphenol A and other phenols in urine from Danish children and adolescents analyzed by isotope diluted TurboFlow-LC-MS/MS. *International Journal of Hygiene and Environmental Health*, *216*(6), 710–720. https://doi.org/10.1016/j.ijheh.2013.01.007

Frederiksen, H., Jensen, T. K., Jørgensen, N., Kyhl, H. B., Husby, S., Skakkebæk, N. E., Main, K. M., Juul, A., & Andersson, A. M. (2014). Human urinary excretion of non-persistent environmental chemicals: An overview of Danish data collected between 2006 and 2012. *Reproduction*, *147*(4), 555–565. https://doi.org/10.1530/REP-13-0522

Frederiksen, H., Nielsen, J. K. S., Mørck, T. A., Hansen, P. W., Jensen, J. F., Nielsen, O., Andersson, A. M., & Knudsen, L. E. (2013). Urinary excretion of phthalate metabolites, phenols and parabens in rural and urban Danish mother-child pairs. *International Journal of Hygiene and Environmental Health*, *216*(6), 772–783. https://doi.org/10.1016/j.ijheh.2013.02.006

Fromme, H., Schlummer, M., Möller, A., Gruber, L., Wolz, G., Ungewiss, J., Böhmer, S., Dekant, W., Mayer, R., Liebl, B., & Twardella, D. (2007). Exposure of an adult population to perfluorinated substances using duplicate diet portions and biomonitoring data. *Environmental Science and Technology*, *41*(22), 7928–7933. https://doi.org/10.1021/es071244n

Geer, L. A., Pycke, B. F. G., Waxenbaum, J., Sherer, D. M., Abulafia, O., & Halden, R. U. (2017). Association of birth outcomes with fetal exposure to parabens, triclosan and triclocarban in an immigrant population in Brooklyn, New York. *Journal of Hazardous Materials*, *323*, 177–183. https://doi.org/10.1016/j.jhazmat.2016.03.028

Gennings, C., Hauser, R., Koch, H. M., Kortenkamp, A., Lioy, P. J., Mirkes, P. E., & Schwetz, B. a. (2014). *Report to the U.S. Consumer Product Safety Commission by the CHRONIC HAZARD ADVISORY PANEL ON PHTHALATES AND PHTHALATE ALTERNATIVES* (Issue July). http://www.cpsc.gov/PageFiles/169902/CHAP-REPORT-With-Appendices.pdf

Gilbert, D., Mayer, P., Pedersen, M., & Vinggaard, A. M. (2015). Endocrine activity of persistent organic pollutants accumulated in human silicone implants - Dosing in vitro assays by partitioning from silicone. *Environment International*, *84*, 107–114. https://doi.org/10.1016/j.envint.2015.07.008

Gonzalez, H., Farbrot, A., Larkö, O., & Wennberg, A. M. (2006). Percutaneous absorption of the sunscreen benzophenone-3 after repeated whole-body applications, with and without ultraviolet irradiation. *British Journal of Dermatology*, *154*(2), 337–340. https://doi.org/10.1111/j.1365-2133.2005.07007.x

Hamers, T., Kamstra, J. H., Cenijn, P. H., Pencikova, K., Palkova, L., Simeckova, P., Vondracek, J., Andersson, P. L., Stenberg, M., & Machala, M. (2011). In vitro toxicity profiling of ultrapure non-dioxin-like polychlorinated biphenyl congeners and their relative toxic contribution to PCB mixtures in humans. *Toxicological Sciences*, *121*(1), 88–100. https://doi.org/10.1093/toxsci/kfr043

Hamers, T., Kamstra, J. H., Sonneveld, E., Murk, A. J., Kester, M. H. A., Andersson, P. L., Legler, J., & Brouwer, A. (2006). In vitro profiling of the endocrine-disrupting potency of brominated flame retardants. *Toxicological Sciences*, *92*(1), 157–173. https://doi.org/10.1093/toxsci/kfj187

Harrad, S., & Abdallah, M. A. E. (2015). Concentrations of Polybrominated Diphenyl Ethers, Hexabromocyclododecanes and Tetrabromobisphenol-A in Breast Milk from United Kingdom

Women Do Not Decrease over Twelve Months of Lactation. *Environmental Science and Technology*, *49*(23), 13899–13903. https://doi.org/10.1021/acs.est.5b00539

Hass, U., Scholze, M., Christiansen, S., Dalgaard, M., Vinggaard, A. M., Axelstad, M., Metzdorff, S. B., & Kortenkamp, A. (2007). Combined exposure to anti-androgens exacerbates disruption of sexual differentiation in the rat. *Environmental Health Perspectives*, *115*(SUPPL1), 122–128. https://doi.org/10.1289/ehp.9360

Hays, S. M., Aylward, L. L., Gagné, M., & Krishnan, K. (2009). Derivation of Biomonitoring Equivalents for cyfluthrin. *Regulatory Toxicology and Pharmacology*, *55*(3), 268–275. https://doi.org/10.1016/j.yrtph.2009.09.002

Hu, X., Shi, W., Wei, S., Zhang, X., Feng, J., Hu, G., Chen, S., Giesy, J. P., & Yu, H. (2013). Occurrence and potential causes of androgenic activities in source and drinking water in China. *Environmental Science and Technology*, *47*(18), 10591–10600. https://doi.org/10.1021/es401464p

Ibarluzea, J., Alvarez-Pedrerol, M., Guxens, M., Marina, L. S., Basterrechea, M., Lertxundi, A., Etxeandia, A., Goñi, F., Vioque, J., Ballester, F., & Sunyer, J. (2011). Sociodemographic, reproductive and dietary predictors of organochlorine compounds levels in pregnant women in Spain. *Chemosphere*, *82*(1), 114–120. https://doi.org/10.1016/j.chemosphere.2010.09.051

Inoue, K., Ritz, B., Andersen, S. L., Ramlau-Hansen, C. H., Høyer, B. B., Bech, B. H., Henriksen, T. B., Bonefeld-Jørgensen, E. C., Olsen, J., & Liew, Z. (2019). Perfluoroalkyl substances and maternal thyroid hormones in early pregnancy; findings in the danish national birth cohort. *Environmental Health Perspectives*, *127*(11), 1–11. https://doi.org/10.1289/EHP5482

Jerzynska, J., Podlecka, D., Polanska, K., Hanke, W., Stelmach, I., & Stelmach, W. (2017). Prenatal and postnatal exposure to polycyclic aromatic hydrocarbons and allergy symptoms in city children. *Allergologia et Immunopathologia*, *45*(1), 18–24. https://doi.org/10.1016/j.aller.2016.07.006

JMPR. (2001). Pesticide residues in food 2001, Toxicological evaluations PROCHLORAZ. *WHO, Geneva*. http://www.inchem.org/documents/jmpr/jmpmono/2001pr11.htm

Kasichayanula, S., House, J. D., Wang, T., & Gu, X. (2007). Percutaneous characterization of the insect repellent DEET and the sunscreen oxybenzone from topical skin application. *Toxicology and Applied Pharmacology*, *223*(2), 187–194. https://doi.org/10.1016/j.taap.2007.05.016

Kasper-Sonnenberg, M., Koch, H. M., Wittsiepe, J., & Wilhelm, M. (2012). Levels of phthalate metabolites in urine among mother-child-pairs - Results from the Duisburg birth cohort study, Germany. *International Journal of Hygiene and Environmental Health*, *215*(3), 373–382. https://doi.org/10.1016/j.ijheh.2011.09.004

Kasper-Sonnenberg, M., Wittsiepe, J., Koch, H. M., Fromme, H., & Wilhelm, M. (2012). Determination of bisphenol A in urine from mother-child pairs - Results from the Duisburg birth cohort study, Germany. *Journal of Toxicology and Environmental Health - Part A: Current Issues*, *75*(8–10), 429–437. https://doi.org/10.1080/15287394.2012.674907

Katsikantami, I., Colosio, C., Alegakis, A., Tzatzarakis, M. N., Vakonaki, E., Rizos, A. K., Sarigiannis, D. A., & Tsatsakis, A. M. (2019). Estimation of daily intake and risk assessment of organophosphorus pesticides based on biomonitoring data – The internal exposure approach. *Food and Chemical Toxicology*, *123*(October 2018), 57–71. https://doi.org/10.1016/j.fct.2018.10.047

Kessler, W., Numtip, W., Völkel, W., Seckin, E., Csanády, G. A., Pütz, C., Klein, D., Fromme, H., & Filser, J. G. (2012). Kinetics of di(2-ethylhexyl) phthalate (DEHP) and mono(2-ethylhexyl) phthalate in blood and of DEHP metabolites in urine of male volunteers after single ingestion of ring-deuterated DEHP. *Toxicology and Applied Pharmacology*, *264*(2), 284–291. https://doi.org/10.1016/j.taap.2012.08.009

Kjærstad, M. B., Taxvig, C., Andersen, H. R., & Nellemann, C. (2010). Mixture effects of endocrine disrupting compounds in vitro. *International Journal of Andrology*, *33*(2), 425–433. https://doi.org/10.1111/j.1365-2605.2009.01034.x

Kjærstad, Mia B., Taxvig, C., Nellemann, C., Vinggaard, A. M., & Andersen, H. R. (2010). Endocrine disrupting effects in vitro of conazole antifungals used as pesticides and pharmaceuticals. *Reproductive Toxicology*, *30*(4), 573–582. https://doi.org/10.1016/j.reprotox.2010.07.009

Kjeldsen, L. S., & Bonefeld-Jørgensen, E. C. (2013). Perfluorinated compounds affect the function of sex hormone receptors. *Environmental Science and Pollution Research*, *20*(11), 8031–8044. https://doi.org/10.1007/s11356-013-1753-3

Koch, H. M., Christensen, K. L. Y., Harth, V., Lorber, M., & Brüning, T. (2012). Di-n-butyl phthalate (DnBP) and diisobutyl phthalate (DiBP) metabolism in a human volunteer after single oral doses. *Archives of Toxicology*, *86*(12), 1829–1839. https://doi.org/10.1007/s00204-012-0908-1

Koch, Holger M., Becker, K., Wittassek, M., Seiwert, M., Angerer, J., & Kolossa-Gehring, M. (2007). Di-n-butylphthalate and butylbenzylphthalate - Urinary metabolite levels and estimated daily intakes: Pilot study for the German Environmental Survey on children. *Journal of Exposure Science and Environmental Epidemiology*, *17*(4), 378–387. https://doi.org/10.1038/sj.jes.7500526

Koch, Holger M., Bolt, H. M., & Angerer, J. (2004). Di(2-ethylhexyl)phthalate (DEHP) metabolites in human urine and serum after a single oral dose of deuterium-labelled DEHP. *Archives of Toxicology*, *78*(3), 123–130. https://doi.org/10.1007/s00204-003-0522-3

Koch, Holger M, Kolossa-Gehring, M., Schröter-Kermani, C., Angerer, J., & Brüning, T. (2012). Bisphenol A in 24 h urine and plasma samples of the German Environmental Specimen Bank from 1995 to 2009: A retrospective exposure evaluation. *Journal of Exposure Science and Environmental Epidemiology*, *22*(6), 610–616. https://doi.org/10.1038/jes.2012.39

Kojima, H., Iida, M., Katsura, E., Kanetoshi, A., Hori, Y., & Kobayashi, K. (2003). Effects of a diphenyl ether-type herbicide, chlornitrofen, and its amino derivative on androgen and estrogen receptor activities. *Environmental Health Perspectives*, *111*(4), 497–502. https://doi.org/10.1289/ehp.5724

Kojima, H., Katsura, E., Takeuchi, S., Niiyama, K., & Kobayashi, K. (2004). Screening for estrogen and androgen receptor activities in 200 pesticides by in vitro reporter gene assays using Chinese hamster ovary cells. *Environmental Health Perspectives*, *112*(5), 524–531. https://doi.org/10.1289/ehp.6649

Kojima, H., Takeuchi, S., Sanoh, S., Okuda, K., Kitamura, S., Uramaru, N., Sugihara, K., & Yoshinari, K. (2019). Profiling of bisphenol A and eight its analogues on transcriptional activity via human nuclear receptors. *Toxicology*, *413*, 48–55. https://doi.org/10.1016/j.tox.2018.12.001

Kojima, H., Takeuchi, S., Uramaru, N., Sugihara, K., Yoshida, T., & Kitamura, S. (2009). Nuclear hormone receptor activity of polybrominated diphenyl ethers and their hydroxylated and

methoxylated metabolites in transactivation assays using Chinese hamster ovary cells. *Environmental Health Perspectives*, *117*(8), 1210–1218. https://doi.org/10.1289/ehp.0900753

Koo, J. W., Parham, F., Kohn, M. C., Masten, S. A., Brock, J. W., Needham, L. L., & Portier, C. J. (2002). The association between biomarker-based exposure estimates for phthalates and demographic factors in a human reference population. *Environmental Health Perspectives*, *110*(4), 405–410. https://doi.org/10.1289/ehp.02110405

Kortenkamp, A., & Faust, M. (2010). Combined exposures to anti-androgenic chemicals: steps towards cumulative risk assessment. *INTERNATIONAL JOURNAL OF ANDROLOGY*, *33*(2), 463–472. https://doi.org/10.1111/j.1365-2605.2009.01047.x

Krüger, T., Long, M., & Bonefeld-Jørgensen, E. C. (2008). Plastic components affect the activation of the aryl hydrocarbon and the androgen receptor. *Toxicology*, *246*(2–3), 112–123. https://doi.org/10.1016/j.tox.2007.12.028

LaKind, J. S., Berlin, C. M., Sjödin, A., Turner, W., Wang, R. Y., Needham, L. L., Paul, I. M., Stokes, J. L., Naiman, D. Q., & Patterson, D. G. (2009). Do human milk concentrations of persistent organic chemicals really decline during lactation? Chemical concentrations during lactation and milk/serum partitioning. *Environmental Health Perspectives*, *117*(10), 1625–1631. https://doi.org/10.1289/ehp.0900876

Larsson, K., Ljung Bjorklund, K., Palm, B., Wennberg, M., Kaj, L., Lindh, C. H., Jonsson, B. A. G., & Berglund, M. (2014). Exposure determinants of phthalates, parabens, bisphenol A and triclosan in Swedish mothers and their children. *Environment International*, *73*, 323–333. https://doi.org/10.1016/j.envint.2014.08.014

Lemaire, G., Terouanne, B., Mauvais, P., Michel, S., & Rahmani, R. (2004). Effect of organochlorine pesticides on human androgen receptor activation in vitro. *Toxicology and Applied Pharmacology*, *196*(2), 235–246. https://doi.org/10.1016/j.taap.2003.12.011

Lioy, P. J., Hauser, R., Gennings, C., Koch, H. M., Mirkes, P. E., Schwetz, B. A., & Kortenkamp, A. (2015). Assessment of phthalates/phthalate alternatives in children's toys and childcare articles: Review of the report including conclusions and recommendation of the Chronic Hazard Advisory Panel of the Consumer Product Safety Commission. *Journal of Exposure Science and Environmental Epidemiology*, *25*(4), 343–353. https://doi.org/10.1038/jes.2015.33

Liu, H., Hu, W., Sun, H., Shen, O., Wang, X., Lam, M. H. W., Giesy, J. P., Zhang, X., & Yu, H. (2011). In vitro profiling of endocrine disrupting potency of 2,2′,4,4′-tetrabromodiphenyl ether (BDE47) and related hydroxylated analogs (HO-PBDEs). *Marine Pollution Bulletin*, *63*(5–12), 287–296. https://doi.org/10.1016/j.marpolbul.2011.04.019

Llop, S., Ballester, F., Estarlich, M., Ibarluzea, J., Manrique, A., Rebagliato, M., Esplugues, A., & 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/10.1016/j.scitotenv.2008.07.070

Llop, S., Murcia, M., Alvarez-Pedrerol, M., Grimalt, J. O., Santa-Marina, L., Julvez, J., Goñi-Irigoyen, F., Espada, M., Ballester, F., Rebagliato, M., & Lopez-Espinosa, M. J. (2017). Association between exposure to organochlorine compounds and maternal thyroid status: Role of the iodothyronine deiodinase 1 gene. *Environment International*, *104*, 83–90. https://doi.org/10.1016/j.envint.2016.12.013

Lopez-Espinosa, M. J., Costa, O., Vizcaino, E., Murcia, M., Fernandez-Somoano, A., Iñiguez, C., Llop, S., Grimalt, J. O., Ballester, F., & Tardon, A. (2015). Prenatal Exposure to Polybrominated Flame Retardants and Fetal Growth in the INMA Cohort (Spain). *Environmental Science and Technology*, *49*(16), 10108–10116. https://doi.org/10.1021/acs.est.5b01793

Ma, R., Cotton, B., Lichtensteiger, W., & Schlumpf, M. (2003). UV filters with antagonistic action at androgen receptors in the MDA-kb2 cell transcriptional-activation assay. *Toxicological Sciences*, *74*(1), 43–50. https://doi.org/10.1093/toxsci/kfg102

Mannetje, A. t., Coakley, J., Mueller, J. F., Harden, F., Toms, L. M., & Douwes, J. (2012). Partitioning of persistent organic pollutants (POPs) between human serum and breast milk: A literature review. *Chemosphere*, *89*(8), 911–918. https://doi.org/10.1016/j.chemosphere.2012.06.049

Manzano-Salgado, C. B., Granum, B., Lopez-Espinosa, M. J., Ballester, F., Iñiguez, C., Gascón, M., Martínez, D., Guxens, M., Basterretxea, M., Zabaleta, C., Schettgen, T., Sunyer, J., Vrijheid, M., & Casas, M. (2019). Prenatal exposure to perfluoroalkyl substances, immune-related outcomes, and lung function in children from a Spanish birth cohort study. *International Journal of Hygiene and Environmental Health*, *222*(6), 945–954. https://doi.org/10.1016/j.ijheh.2019.06.005

Marchitti, S. A., Fenton, S. E., Mendola, P., Kenneke, J. F., & Hines, E. P. (2017). Polybrominated diphenyl ethers in human milk and serum from the U.S. EPA MAMA study: Modeled predictions of infant exposure and considerations for risk assessment. *Environmental Health Perspectives*, *125*(4), 706–713. https://doi.org/10.1289/EHP332

Mariscal-Arcas, M., Lopez-Martinez, C., Granada, A., Olea, N., Lorenzo-Tovar, M. L., & Olea-Serrano, F. (2010). Organochlorine pesticides in umbilical cord blood serum of women from Southern Spain and adherence to the Mediterranean diet. *Food and Chemical Toxicology*, *48*(5), 1311–1315. https://doi.org/10.1016/j.fct.2010.02.029

Mes, J., Doyle, J. A., Adams, B. R., Davies, D. J., & Turton, D. (1984). Polychlorinated biphenyls and organochlorine pesticides in milk and blood of Canadian women during lactation. *Archives of Environmental Contamination and Toxicology*, *13*(2), 217–223. https://doi.org/10.1007/BF01055879

Messerlian, C., Mustieles, V., Minguez-Alarcon, L., Ford, J. B., Calafat, A. M., Souter, I., Williams, P. L., & Hauser, R. (2018). Preconception and prenatal urinary concentrations of phenols and birth size of singleton infants born to mothers and fathers from the Environment and Reproductive Health (EARTH) study. *Environment International*, *114*(November 2017), 60–68. https://doi.org/10.1016/j.envint.2018.02.017

Metzdorff, S. B., Dalgaard, M., Christiansen, S., Axelstad, M., Hass, U., Kiersgaard, M. K., Scholze, M., Kortenkamp, A., & Vinggaard, A. M. (2007). Dysgenesis and histological changes of genitals and perturbations of gene expression in male rats after in utero exposure to antiandrogen mixtures. *Toxicological Sciences*, *98*(1), 87–98. https://doi.org/10.1093/toxsci/kfm079

Metzler, M., Pfeiffer, E., & Hildebrand, A. A. (2010). Zearalenone and its metabolites as endocrine disrupting chemicals. *World Mycotoxin Journal*, *3*(4), 385–401. https://doi.org/10.3920/WMJ2010.1244

Mínguez-Alarcón, L., Chiu, Y. H., Messerlian, C., Williams, P. L., Sabatini, M. E., Toth, T. L., Ford, J. B., Calafat, A. M., & Hauser, R. (2016). Urinary paraben concentrations and in vitro fertilization

outcomes among women from a fertility clinic. *Fertility and Sterility*, *105*(3), 714–721. https://doi.org/10.1016/j.fertnstert.2015.11.021

Mirocha, C. J. (1981). Comparative Metabolism of Zearalenone Transmission. *Food and Cosmetics Toxicology*, *19*, 25–30.

Misaki, K., Suzuki, G., Tue, N. M., Takahashi, S., Someya, M., Takigami, H., Tajima, Y., Yamada, T. K., Amano, M., Isobe, T., & Tanabe, S. (2015). Toxic Identification and Evaluation of Androgen Receptor Antagonistic Activities in Acid-Treated Liver Extracts of High-Trophic Level Wild Animals from Japan. *Environmental Science and Technology*, *49*(19), 11840–11848. https://doi.org/10.1021/acs.est.5b02288

Molina-Molina, J. M., Amaya, E., Grimaldi, M., Sáenz, J. M., Real, M., Fernández, M. F., Balaguer, P., & Olea, N. (2013). In vitro study on the agonistic and antagonistic activities of bisphenol-S and other bisphenol-A congeners and derivatives via nuclear receptors. *Toxicology and Applied Pharmacology*, *272*(1), 127–136. https://doi.org/10.1016/j.taap.2013.05.015

Molina-Molina, J. M., Escande, A., Pillon, A., Gomez, E., Pakdel, F., Cavaillès, V., Olea, N., Aït-Aïssa, S., & Balaguer, P. (2008). Profiling of benzophenone derivatives using fish and human estrogen receptor-specific in vitro bioassays. *Toxicology and Applied Pharmacology*, *232*(3), 384–395. https://doi.org/10.1016/j.taap.2008.07.017

Monteagudo, C., Mariscal-Arcas, M., Heras-Gonzalez, L., Ibañez-Peinado, D., Rivas, A., & Olea-Serrano, F. (2016). Effects of maternal diet and environmental exposure to organochlorine pesticides on newborn weight in Southern Spain. *Chemosphere*, *156*, 135–142. https://doi.org/10.1016/j.chemosphere.2016.04.103

Moos, R. K., Angerer, J., Dierkes, G., Brüning, T., & Koch, H. M. (2016). Metabolism and elimination of methyl, iso- and n-butyl paraben in human urine after single oral dosage. *Archives of Toxicology*, *90*(11), 2699–2709. https://doi.org/10.1007/s00204-015-1636-0

Moos, R. K., Apel, P., Schröter-Kermani, C., Kolossa-Gehring, M., Brüning, T., & Koch, H. M. (2017). Daily intake and hazard index of parabens based upon 24h urine samples of the German environmental specimen bank from 1995 to 2012. *Journal of Exposure Science and Environmental Epidemiology*, *27*(6), 591–600. https://doi.org/10.1038/jes.2016.65

Mørck, T. A., Erdmann, S. E., Long, M., Mathiesen, L., Nielsen, F., Siersma, V. D., Bonefeld-Jørgensen, E. C., & Knudsen, L. E. (2014). PCB concentrations and dioxin-like activity in blood samples from Danish school children and their mothers living in urban and rural areas. *Basic and Clinical Pharmacology and Toxicology*, *115*(1), 134–144. https://doi.org/10.1111/bcpt.12214

Morgan, D. P., Hetzler, H. L., Slach, E. F., & Lin, L. I. (1977). Urinary excretion of paranitrophenol and alkyl phosphates following ingestion of methyl or ethyl parathion by human subjects. *Archives of Environmental Contamination and Toxicology*, *6*(1), 159–173. https://doi.org/10.1007/BF02097758

Motorykin, O., Santiago-delgado, L., Rohlman, D., Schrlau, J. E., Harris, S., Harding, A., Kile, M. L., & Simonich, S. L. M. (2016). *Salmon for Native American Volunteers*. *541*, 170–177. https://doi.org/10.1016/j.scitotenv.2015.01.083.Metabolism

Mukherjee, D., Royce, S. G., Alexander, J. A., Buckley, B., Isukapalli, S. S., Bandera, E. V., Zarbl, H., & Georgopoulos, P. G. (2014). Physiologically-based toxicokinetic modeling of zearalenone and its metabolites: Application to the Jersey girl study. *PLoS ONE*, *9*(12), 1–30. https://doi.org/10.1371/journal.pone.0113632

Myridakis, A., Fthenou, E., Balaska, E., Vakinti, M., Kogevinas, M., & Stephanou, E. G. (2015). Phthalate esters, parabens and bisphenol-A exposure among mothers and their children in Greece (Rhea cohort). *Environment International*, *83*, 1–10. https://doi.org/10.1016/j.envint.2015.05.014

Nakiwala, D., Peyre, H., Heude, B., Bernard, J. Y., Béranger, R., Slama, R., Philippat, C., Charles, M. A., De Agostini, M., Forhan, A., Heude, B., Ducimetière, P., Kaminski, M., Saurel-Cubizolles, M. J., Dargent-Molina, P., Fritel, X., Larroque, B., Lelong, N., Marchand, L., … Botton, J. (2018). In-utero exposure to phenols and phthalates and the intelligence quotient of boys at 5 years. *Environmental Health: A Global Access Science Source*, *17*(1), 1–11. https://doi.org/10.1186/s12940-018-0359-0

NCCT US-EPA. CompTox Chemicals Dashboard. Accessed January 7, 2021 from https://comptox.epa.gov/dashboard/

Nellemann, C., Dalgaard, M., Lam, H. R., & Vinggaard, A. M. (2003). The combined effects of vinclozolin and procymidone do not deviate from expected additivity in vitro and in vivo. *Toxicological Sciences*, *71*(2), 251–262. https://doi.org/10.1093/toxsci/71.2.251

Nøst, T. H., Berg, V., Hanssen, L., Rylander, C., Gaudreau, E., Dumas, P., Breivik, K., & Sandanger, T. M. (2019). Time trends of persistent organic pollutants in 30 year olds sampled in 1986, 1994, 2001 and 2007 in Northern Norway: Measurements, mechanistic modeling and a comparison of study designs. *Environmental Research*, *172*(March), 684–692. https://doi.org/10.1016/j.envres.2019.02.047

Orton, F., Ermler, S., Kugathas, S., Rosivatz, E., Scholze, M., & Kortenkamp, A. (2014). Mixture effects at very low doses with combinations of anti-androgenic pesticides, antioxidants, industrial pollutant and chemicals used in personal care products. *Toxicology and Applied Pharmacology*, *278*(3), 201–208. https://doi.org/10.1016/j.taap.2013.09.008

Orton, F., Rosivatz, E., Scholze, M., & Kortenkamp, A. (2011). Widely used pesticides with previously unknown endocrine activity revealed as in Vitro antiandrogens. *Environmental Health Perspectives*, *119*(6), 794–800. https://doi.org/10.1289/ehp.1002895

Peinado, F. M., Lendínez, I., Sotelo, R., Iribarne-Durán, L. M., Fernández-Parra, J., Vela-Soria, F., Olea, N., Fernández, M. F., Freire, C., León, J., Pérez-Cabrera, B., Ocón-Hernández, O., & Artacho-Cordón, F. (2020). Association of urinary levels of bisphenols A, F, and S with endometriosis risk: Preliminary results of the EndEA study. *International Journal of Environmental Research and Public Health*, *17*(4). https://doi.org/10.3390/ijerph17041194

Pěnčíková, K., Svržková, L., Strapáčová, S., Neča, J., Bartoňková, I., Dvořák, Z., Hýžďalová, M., Pivnička, J., Pálková, L., Lehmler, H. J., Li, X., Vondráček, J., & Machala, M. (2018). In vitro profiling of toxic effects of prominent environmental lower-chlorinated PCB congeners linked with endocrine disruption and tumor promotion. *Environmental Pollution*, *237*, 473–486. https://doi.org/10.1016/j.envpol.2018.02.067

Philippat, C., Heude, B., Botton, J., Alfaidy, N., Calafat, A. M., Slama, R., Annesi-Maesano, I., Bernard, J., Botton, J., Charles, M. A., Dargent-Molina, P., De Lauzon-Guillian, B., Ducimetière, P., De Agostini, M., Foliguet, B., Forhan, A., Fritel, X., Germa, A., Goua, V., … Thiebaugeorge, O. (2019). Prenatal exposure to select phthalates and phenols and associations with fetal and placental weight among male births in the EDEN Cohort (France). *Environmental Health Perspectives*, *127*(1), 1–8. https://doi.org/10.1289/EHP3523

Philippat, C., Mortamais, M., Chevrier, C., Petit, C., Calafat, A. M., Ye, X., Silva, M. J., Brambilla, C., Pin, I., Charles, M. A., Cordier, S., & Slama, R. (2012). Exposure to phthalates and phenols during pregnancy and offspring size at birth. *Environmental Health Perspectives*, *120*(3), 464–470. https://doi.org/10.1289/ehp.1103634

Philippat, C., Nakiwala, D., Calafat, A. M., Botton, J., De Agostini, M., Heude, B., Slama, R., Annesi-Maesano, I., Bernard, J., Botton, J., Charles, M. A., Dargent-Molina, P., De Lauzon-Guillain, B., Ducimetière, P., De Agostini, M., Foliguet, B., Forhan, A., Fritel, X., Germa, A., … Thiebaugeorge, O. (2017). Prenatal exposure to nonpersistent endocrine disruptors and behavior in boys at 3 and 5 years. *Environmental Health Perspectives*, *125*(9), 1–9. https://doi.org/10.1289/EHP1314

Philips, E. M., Jaddoe, V. W. V, Asimakopoulos, A. G., Kannan, K., Steegers, E. A. P., Santos, S., & Trasande, L. (2018). Bisphenol and phthalate concentrations and its determinants among pregnant women in a population-based cohort in the Netherlands, 2004–5. *Environmental Research*, *161*(November 2017), 562–572. https://doi.org/10.1016/j.envres.2017.11.051

Philips, E. M., Santos, S., Steegers, E. A. P., Asimakopoulos, A. G., Kannan, K., Trasande, L., & Jaddoe, V. W. V. (2020). Maternal bisphenol and phthalate urine concentrations and weight gain during pregnancy. *Environment International*, *135*(December 2019), 105342. https://doi.org/10.1016/j.envint.2019.105342

Polanska, K., Dettbarn, G., Jurewicz, J., Sobala, W., Magnus, P., Seidel, A., & Hanke, W. (2014). Effect of prenatal polycyclic aromatic hydrocarbons exposure on birth outcomes: The polish mother and child cohort study. *BioMed Research International*, *2014*. https://doi.org/10.1155/2014/408939

Polańska, K., Hanke, W., Sobala, W., Brzeźnicki, S., & Ligocka, D. (2011). Predictors of environmental exposure to polycyclic aromatic hydrocarbons among pregnant women - Prospective cohort study in Poland. *International Journal of Occupational Medicine and Environmental Health*, *24*(1), 8–17. https://doi.org/10.2478/s13382-011-0003-5

Pollack, A. Z., Mumford, S. L., Krall, J. R., Carmichael, A. E., Sjaarda, L. A., Perkins, N. J., Kannan, K., & Schisterman, E. F. (2018). Exposure to bisphenol A, chlorophenols, benzophenones, and parabens in relation to reproductive hormones in healthy women: A chemical mixture approach. *Environment International*, *120*(April), 137–144. https://doi.org/10.1016/j.envint.2018.07.028

Pollack, A. Z., Perkins, N. J., Sjaarda, L., Mumford, S. L., Kannan, K., Philippat, C., Wactawski-Wende, J., & Schisterman, E. F. (2016). Variability and exposure classification of urinary phenol and paraben metabolite concentrations in reproductive-aged women. *Environmental Research*, *151*, 513–520. https://doi.org/10.1016/j.envres.2016.08.016

Pop, A., Drugan, T., Gutleb, A. C., Lupu, D., Cherfan, J., Loghin, F., & Kiss, B. (2016). Individual and combined in vitro (anti)androgenic effects of certain food additives and cosmetic preservatives. *Toxicology in Vitro*, *32*, 269–277. https://doi.org/10.1016/j.tiv.2016.01.012

Proença, S., Escher, B. I., Fischer, F. C., Fisher, C., Grégoire, S., Hewitt, N. J., Nicol, B., Paini, A., & Kramer, N. I. (2021). Effective exposure of chemicals in in vitro cell systems: A review of chemical distribution models. *Toxicology in Vitro*, *73*(March), 105133. https://doi.org/10.1016/j.tiv.2021.105133

Pulkrabova, J., Stupak, M., Svarcova, A., Rossner, P., Rossnerova, A., Ambroz, A., Sram, R., & Hajslova, J. (2016). Relationship between atmospheric pollution in the residential area and

concentrations of polycyclic aromatic hydrocarbons (PAHs) in human breast milk. *Science of the Total Environment*, *562*, 640–647. https://doi.org/10.1016/j.scitotenv.2016.04.013

Punt, A., Aartse, A., Bovee, T. F. H., Gerssen, A., van Leeuwen, S. P. J., Hoogenboom, R. L. A. P., & Peijnenburg, A. A. C. M. (2019). Quantitative in vitro-to-in vivo extrapolation (QIVIVE) of estrogenic and anti-androgenic potencies of BPA and BADGE analogues. *Archives of Toxicology*, *93*(7), 1941–1953. https://doi.org/10.1007/s00204-019-02479-6

Rajapakse, N., Silva, E., & Kortenkamp, A. (2002). Combining xenoestrogens at levels below individual no-observed-effect concentrations dramatically enhances steroid hormone action. *Environmental Health Perspectives*, *110*(9), 917–921. https://doi.org/10.1289/ehp.02110917

Ratelle, M., Coté, J., & Bouchard, M. (2015). Time profiles and toxicokinetic parameters of key biomarkers of exposure to cypermethrin in orally exposed volunteers compared with previously available kinetic data following permethrin exposure. *Journal of Applied Toxicology*, *35*(12), 1586–1593. https://doi.org/10.1002/jat.3124

Rochester, J. R., & Bolden, A. L. (2015). Bisphenol S and F: A systematic review and comparison of the hormonal activity of bisphenol a substitutes. *Environmental Health Perspectives*, *123*(7), 643–650. https://doi.org/10.1289/ehp.1408989

Roelofs, M. J. E., Temming, A. R., Piersma, A. H., van den Berg, M., & van Duursen, M. B. M. (2014). Conazole fungicides inhibit Leydig cell testosterone secretion and androgen receptor activation in vitro. *Toxicology Reports*, *1*, 271–283. https://doi.org/10.1016/j.toxrep.2014.05.006

Rosenmai, A. K., Dybdahl, M., Pedersen, M., van Vugt-Lussenburg, B. M. A., Wedebye, E. B., Taxvig, C., & Vinggaard, A. M. (2014). Are structural analogues to bisphenol a safe alternatives? *Toxicological Sciences*, *139*(1), 35–47. https://doi.org/10.1093/toxsci/kfu030

Rostkowski, P., Horwood, J., Shears, J. A., Lange, A., Oladapo, F. O., Besselink, H. T., Tyler, C. R., & Hill, E. M. (2011). Bioassay-directed identification of novel antiandrogenic compounds in bile of fish exposed to wastewater effluents. *Environmental Science and Technology*, *45*(24), 10660–10667. https://doi.org/10.1021/es202966c

Ruel, M. V. M., Bos, A. F., Soechitram, S. D., Meijer, L., Sauer, P. J. J., & Berghuis, S. A. (2019). Prenatal exposure to organohalogen compounds and children's mental and motor development at 18 and 30 months of age. *NeuroToxicology*, *72*(September 2018), 6–14. https://doi.org/10.1016/j.neuro.2019.01.003

Safe Work Australia. (2020). *Report on PARATHION*. https://engage.swa.gov.au/65429/widgets/323765/documents/191562

Sakhi, A. K., Sabaredzovic, A., Papadopoulou, E., Cequier, E., & Thomsen, C. (2018). Levels, variability and determinants of environmental phenols in pairs of Norwegian mothers and children. *Environment International*, *114*, 242–251. https://doi.org/10.1016/j.envint.2018.02.037

Sanchis, Y., Coscollà, C., & Yusà, V. (2019). Analysis of four parabens and bisphenols A, F, S in urine, using dilute and shoot and liquid chromatography coupled to mass spectrometry. *Talanta*, *202*(November 2018), 42–50. https://doi.org/10.1016/j.talanta.2019.04.048

Sandborgh-Englund, G., Adolfsson-Erici, M., Odham, G., & Ekstrand, J. (2006). Pharmacokinetics of triclosan following oral ingestion in humans. *Journal of Toxicology and Environmental Health - Part A: Current Issues*, *69*(20), 1861–1873. https://doi.org/10.1080/15287390600631706

Satoh, K., Ohyama, K., Aoki, N., Iida, M., & Nagai, F. (2004). Study on anti-androgenic effects of bisphenol a diglycidyl ether (BADGE), bisphenol F diglycidyl ether (BFDGE) and their derivatives using cells stably transfected with human androgen receptor, AR-EcoScreen. *Food and Chemical Toxicology*, *42*(6), 983–993. https://doi.org/10.1016/j.fct.2004.02.011

Satoh, Kanako, Nonaka, R., Ohyama, K. I., & Nagai, F. (2005). Androgenic and antiandrogenic effects of alkylphenols and parabens assessed using the reporter gene assay with stably transfected CHO-K1 cells (AR-EcoScreen system). *Journal of Health Science*, *51*(5), 557–568. https://doi.org/10.1248/jhs.51.557

SCCS (Scientific Committee on Consumer Safety). (2013). *Opinion on parabens*. https://doi.org/10.2772/66369

Scholze, M., Taxvig, C., Kortenkamp, A., Boberg, J., Christiansen, S., Svingen, T., Lauschke, K., Frandsen, H., Ermler, S., Hermann, S. S., Pedersen, M., Lykkeberg, A. K., Axelstad, M., & Vinggaard, A. M. (2020). Quantitative in vitro to in vivo extrapolation (QIVIVE) for predicting reduced anogenital distance produced by anti-androgenic pesticides in a rodent model for male reproductive disorders. *Environmental Health Perspectives*, *128*(11), 1–17. https://doi.org/10.1289/EHP6774

Schrader, T. J., & Cooke, G. M. (2000). Examination of selected food additives and organochlorine food contaminants for androgenic activity in vitro. *Toxicological Sciences*, *53*(2), 278–288. https://doi.org/10.1093/toxsci/53.2.278

Schreurs, R. H. M. M., Sonneveld, E., Jansen, J. H. J., Seinen, W., & van der Burg, B. (2005). Interaction of polycyclic musks and UV filters with the estrogen receptor (ER), androgen receptor (AR), and progesterone receptor (PR) in reporter gene bioassays. *Toxicological Sciences*, *83*(2), 264–272. https://doi.org/10.1093/toxsci/kfi035

Shen, H., Main, K. M., Virtanen, H. E., Damggard, I. N., Haavisto, A. M., Kaleva, M., Boisen, K. A., Schmidt, I. M., Chellakooty, M., Skakkebaek, N. E., Toppari, J., & Schramm, K. W. (2007). From mother to child: Investigation of prenatal and postnatal exposure to persistent bioaccumulating toxicants using breast milk and placenta biomonitoring. *Chemosphere*, *67*(9), 256–262. https://doi.org/10.1016/j.chemosphere.2006.05.106

Shen, O., Du, G., Sun, H., Wu, W., Jiang, Y., Song, L., & Wang, X. (2009). Comparison of in vitro hormone activities of selected phthalates using reporter gene assays. *Toxicology Letters*, *191*(1), 9–14. https://doi.org/10.1016/j.toxlet.2009.07.019

Shin, M. Y., Shin, C., Choi, J. W., Lee, J., Lee, S., & Kim, S. (2019). Pharmacokinetic profile of propyl paraben in humans after oral administration. *Environment International*, *130*(June), 104917. https://doi.org/10.1016/j.envint.2019.104917

Silva, E., Rajapakse, N., & Kortenkamp, A. (2002). Something from “nothing” - Eight weak estrogenic chemicals combined at concentrations below NOECs produce significant mixture effects. *Environmental Science and Technology*, *36*(8), 1751–1756. https://doi.org/10.1021/es0101227

Skogheim, T. S., Villanger, G. D., Weyde, K. V. F., Engel, S. M., Surén, P., Øie, M. G., Skogan, A. H., Biele, G., Zeiner, P., Øvergaard, K. R., Haug, L. S., Sabaredzovic, A., & Aase, H. (2020). Prenatal exposure to perfluoroalkyl substances and associations with symptoms of attention-deficit/hyperactivity disorder and cognitive functions in preschool children. *International Journal of Hygiene and Environmental Health*, *223*(1), 80–92. https://doi.org/10.1016/j.ijheh.2019.10.003

Suzuki, G., Tue, N. M., Malarvannan, G., Sudaryanto, A., Takahashi, S., Tanabe, S., Sakai, S. I., Brouwer, A., Uramaru, N., Kitamura, S., & Takigami, H. (2013). Similarities in the endocrine-disrupting potencies of indoor dust and flame retardants by using human osteosarcoma (U2OS) cell-based reporter gene assays. *Environmental Science and Technology*, *47*(6), 2898–2908. https://doi.org/10.1021/es304691a

Takeuchi, S., Iida, M., Kobayashi, S., Jin, K., Matsuda, T., & Kojima, H. (2005). Differential effects of phthalate esters on transcriptional activities via human estrogen receptors α and β, and androgen receptor. *Toxicology*, *210*(2–3), 223–233. https://doi.org/10.1016/j.tox.2005.02.002

Thayer, K. A., Doerge, D. R., Hunt, D., Schurman, S. H., Twaddle, N. C., Churchwell, M. I., Garantziotis, S., Kissling, G. E., Easterling, M. R., Bucher, J. R., & Birnbaum, L. S. (2015). Pharmacokinetics of bisphenol A in humans following a single oral administration. *Environment International*, *83*, 107–115. https://doi.org/10.1016/j.envint.2015.06.008

Timchalk, C., Selim, S., Sangha, G., & Bartels, M. J. (1998). The pharmacokinetics and metabolism of 14C/13C-labeled ortho-phenylphenol formation following dermal application to human volunteers. *Human and Experimental Toxicology*, *17*(8), 411–417. https://doi.org/10.1191/096032798678909115

Todaka, T., Hirakawa, H., Kajiwara, J., Hori, T., Tobiishi, K., Yasutake, D., Onozuka, D., Sasaki, S., Miyashita, C., Yoshioka, E., Yuasa, M., Kishi, R., Iida, T., & Furue, M. (2010). Relationship between the concentrations of polychlorinated dibenzo-p-dioxins, polychlorinated dibenzofurans, and polychlorinated biphenyls in maternal blood and those in breast milk. *Chemosphere*, *78*(2), 185–192. https://doi.org/10.1016/j.chemosphere.2009.09.047

Todaka, T., Hirakawa, H., Kajiwara, J., Onozuka, D., Sasaki, S., Miyashita, C., Yoshioka, E., Yuasa, M., Kishi, R., Iida, T., Uchi, H., & Furue, M. (2011). Concentrations of polychlorinated dibenzo-p-dioxins, polychlorinated dibenzofurans, and polychlorinated biphenyls in blood and breast milk collected from pregnant women in Sapporo City, Japan. *Chemosphere*, *85*(11), 1694–1700. https://doi.org/10.1016/j.chemosphere.2011.09.014

Toriba, A., Kitaoka, H., Dills, R. L., Mizukami, S., Tanabe, K., Takeuchi, N., Ueno, M., Kameda, T., Tang, N., Hayakawa, K., & Simpson, C. D. (2007). Identification and quantification of 1-nitropyrene metabolites in human urine as a proposed biomarker for exposure to diesel exhaust. *Chemical Research in Toxicology*, *20*(7), 999–1007. https://doi.org/10.1021/tx700015q

Tsang, H. L., Wu, S., Leung, C. K. M., Tao, S., & Wong, M. H. (2011). Body burden of POPs of Hong Kong residents, based on human milk, maternal and cord serum. *Environment International*, *37*(1), 142–151. https://doi.org/10.1016/j.envint.2010.08.010

Vafeiadi, M., Roumeliotaki, T., Chalkiadaki, G., Rantakokko, P., Kiviranta, H., Fthenou, E., Kyrtopoulos, S. A., Kogevinas, M., & Chatzi, L. (2017). Persistent organic pollutants in early pregnancy and risk of gestational diabetes mellitus. *Environment International*, *98*(226285), 89–95. https://doi.org/10.1016/j.envint.2016.10.005

Valvi, D., Oulhote, Y., Weihe, P., Dalgård, C., Bjerve, K. S., Steuerwald, U., & Grandjean, P. (2017). Gestational diabetes and offspring birth size at elevated environmental pollutant exposures. *Environment International*, *107*(February), 205–215. https://doi.org/10.1016/j.envint.2017.07.016

Viau, C., Carrier, G., Vyskočil, A., & Dodd, C. (1995). Urinary excretion kinetics of 1-hydroxypyrene in volunteers exposed to pyrene by the oral and dermal route. *Science of the Total Environment*, *163*(1–3), 179–186. https://doi.org/10.1016/0048-9697(95)04494-L

Vinggaard, A. M., Hnida, C., & Larsen, J. C. (2000). Environmental polycyclic aromatic hydrocarbons affect androgen receptor activation in vitro. *Toxicology*, *145*(2–3), 173–183. https://doi.org/10.1016/S0300-483X(00)00144-X

Vinggaard, A. M., Nellemann, C., Dalgaard, M., Jørgensen, E. B., & Andersen, H. R. (2002). Antiandrogenic effects in vitro and in vivo of the fungicide prochloraz. *Toxicological Sciences*, *69*(2), 344–353. https://doi.org/10.1093/toxsci/69.2.344

Vinggaard, A. M., Niemelä, J., Wedebye, E. B., & Jensen, G. E. (2008). Screening of 397 chemicals and development of a quantitative structure-activity relationship model for androgen receptor antagonism. *Chemical Research in Toxicology*, *21*(4), 813–823. https://doi.org/10.1021/tx7002382

Wang, S., Rijk, J. C. W., Besselink, H. T., Houtman, R., Peijnenburg, A. A. C. M., Brouwer, A., Rietjens, I. M. C. M., & Bovee, T. F. H. (2014). Extending an in vitro panel for estrogenicity testing: The added value of bioassays for measuring antiandrogenic activities and effects on steroidogenesis. *Toxicological Sciences*, *141*(1), 78–89. https://doi.org/10.1093/toxsci/kfu103

WHO. (1989). Aldrin and Dieldrin. *Environmental Health Criteria 91. Geneva: World Health Organization*.

Wikström, S., Lin, P. I., Lindh, C. H., Shu, H., & Bornehag, C. G. (2020). Maternal serum levels of perfluoroalkyl substances in early pregnancy and offspring birth weight. *Pediatric Research*, *87*(6), 1093–1099. https://doi.org/10.1038/s41390-019-0720-1

Williams, R. L., Bernard, C. E., & Krieger, R. I. (2003). Human exposure to indoor residential cyfluthrin residues during a structured activity program. *Journal of Exposure Analysis and Environmental Epidemiology*, *13*(2), 112–119. https://doi.org/10.1038/sj.jea.7500257

Wilson, A., & Hall, L. (1986). Pharmacokinetic Study of Alachlor Distribution and Elimination in the Long-Evans Rat: Part I: Absorption, Distribution and Excretion: Lab Project Number: EHL 84027: ML-84-096: 1156. *Unpublished Study*.

Wittsiepe, J., Fürst, P., Schrey, P., Lemm, F., Kraft, M., Eberwein, G., Winneke, G., & Wilhelm, M. (2007). PCDD/F and dioxin-like PCB in human blood and milk from German mothers. *Chemosphere*, *67*(9). https://doi.org/10.1016/j.chemosphere.2006.05.118

Wolff, M. S. (1983). Occupationally derived chemicals in breast milk. *American Journal of Industrial Medicine*, *4*(1–2), 259–281. https://doi.org/10.1002/ajim.4700040120

Woollen, B. H., Marsh, J. R., Laird, W. J. D., & Lesser, J. E. (1992). The metabolism of cypermethrin in man: Differences in urinary metabolite profiles following oral and dermal administration. *Xenobiotica*, *22*(8), 983–991. https://doi.org/10.3109/00498259209049904

Xu, L. C., Liu, L., Ren, X. M., Zhang, M. R., Cong, N., Xu, A. Q., & Shao, J. H. (2008). Evaluation of androgen receptor transcriptional activities of some pesticides in vitro. *Toxicology*, *243*(1–2), 59–65. https://doi.org/10.1016/j.tox.2007.09.028

Xu, L. C., Sun, H., Chen, J. F., Bian, Q., Qian, J., Song, L., & Wang, X. R. (2005). Evaluation of androgen receptor transcriptional activities of bisphenol A, octylphenol and nonylphenol in vitro. *Toxicology*, *216*(2–3), 197–203. https://doi.org/10.1016/j.tox.2005.08.006

Xu, L. C., Sun, H., Chen, J. F., Bian, Q., Song, L., & Wang, X. R. (2006). Androgen receptor activities of p,p′-DDE, fenvalerate and phoxim detected by androgen receptor reporter gene assay. *Toxicology Letters*, *160*(2), 151–157. https://doi.org/10.1016/j.toxlet.2005.06.016

Ye, X., Pierik, F. H., Angerer, J., Meltzer, H. M., Jaddoe, V. W. V, Tiemeier, H., Hoppin, J. A., & Longnecker, M. P. (2009). Levels of metabolites of organophosphate pesticides, phthalates, and bisphenol A in pooled urine specimens from pregnant women participating in the Norwegian Mother and Child Cohort Study (MoBa). *International Journal of Hygiene and Environmental Health*, *212*(5), 481–491. https://doi.org/10.1016/j.ijheh.2009.03.004

Ye, X., Pierik, F. H., Hauser, R., Duty, S., Angerer, J., Park, M. M., Burdorf, A., Hofman, A., Jaddoe, V. W. V., Mackenbach, J. P., Steegers, E. A. P., Tiemeier, H., & Longnecker, M. P. (2008). Urinary metabolite concentrations of organophosphorous pesticides, bisphenol A, and phthalates among pregnant women in Rotterdam, the Netherlands: The Generation R study. *Environmental Research*, *108*(2), 260–267. https://doi.org/10.1016/j.envres.2008.07.014

Zeman, F. A., Boudet, C., Tack, K., Floch Barneaud, A., Brochot, C., Péry, A. R. R., Oleko, A., & Vandentorren, S. (2013). Exposure assessment of phthalates in French pregnant women: Results of the ELFE pilot study. *International Journal of Hygiene and Environmental Health*, *216*(3), 271–279. https://doi.org/10.1016/j.ijheh.2012.12.005

Zhang, F., Hu, W., Yu, H., Sun, H., Shen, O., Wang, X., Liu, H., Lam, M. H. W., Giesy, J. P., & Zhang, X. (2011). Endocrine disruption effects of 2,2′,4,4′,6- pentabromodiphenylether (BDE100) in reporter gene assays. *Journal of Environmental Monitoring*, *13*(4), 850–854. https://doi.org/10.1039/c0em00654h

Zwart, N., Andringa, D., de Leeuw, W. J., Kojima, H., Iida, M., Houtman, C. J., de Boer, J., Kool, J., Lamoree, M. H., & Hamers, T. (2017). Improved androgen specificity of AR-EcoScreen by CRISPR based glucocorticoid receptor knockout. *Toxicology in Vitro*, *45*(August), 1–9. https://doi.org/10.1016/j.tiv.2017.08.004

8 Case study 3: A mixture risk assessment for developmental neurotoxicity with a focus on declines in IQ

Mousumi Chatterjee, Andreas Kortenkamp, Brunel University London

8.1 Introduction

Some chemicals such as inorganic arsenic, PCBs, methyl mercury and lead are recognized developmental neurotoxicants. Exposure to these substances early in fetal life can cause irreversible brain injury at doses considerably lower than those required to affect adult brains. Developmental neurotoxicity, assessed in terms of declines in IQ scores, is regarded as the critical toxicity for these pollutants and has been used to derive HBGV.

Significant proportions of the general population, and in particular expectant mothers, currently experience exposures around the HBGV of some developmental neurotoxicants. While this gives reason for concern, the combined impact of multiple developmental neurotoxicants on brain development has not been assessed. This case study is intended to fill this gap.

Our specific aims were to investigate the extent of risks associated with combined exposures of developmental neurotoxicants and to identify pollutants that drive these risks.

Developmental neurotoxicity encompasses a wide spectrum of disorders, including autism, attention deficit hyperactivity disorder, cerebral palsy, mental retardation and declines in cognitive ability. The available evidence suggests that these disorders can be triggered at exposures different from those associated with declines in cognitive ability, measured as IQ scores. Comparisons of RQs derived by consideration of exposures associated with several different developmental disorders might confound the MRA. To avoid such confounding, we have focused our assessment on declines in IQ scores and sought to identify exposures associated with such effects.

We have based our MRA on combined exposures experienced by the general population in Europe. Because lasting damage from developmental neurotoxicants can arise from exposures during pregnancy or in neonatal life, we gave particular attention to exposures of expectant mothers.

8.2 Grouping criteria and cumulative assessment group

Exposures of expectant mothers to lead, methyl mercury and non-dioxin-like PCBs are associated with declines in IQ scores in their children. The HBGV used in Europe for lead and methyl mercury are based on IQ declines as the critical toxicity. We therefore included lead, methyl mercury and PCBs in our assessment group. To gain a more comprehensive view of other developmental neurotoxicants also associated with IQ declines, we consulted the overview paper by (Grandjean and Landrigan, 2006) and added inorganic arsenic, fluoride, manganese, organophosphate pesticides and polybrominated diphenyl ethers.

It can be expected that a variety of molecular modes of action lead to declines in cognitive ability. These include interference with the thyroid hormone system or direct toxicity to developing brain cells. In many cases, the precise mode of action is unknown. For this reason, mode of action considerations could not play a part in our grouping efforts.

Where epidemiological studies showed associations with IQ declines, but did not permit the quantifications of tolerable exposures, we sought supporting evidence from animal studies.

Thus, the CAG for this case study includes:

- Inorganic arsenic
- Fluoride
- Methyl mercury
- Lead
- PCBs
- Manganese
- Organophosphate pesticides
- PBDEs

8.3 Hazard characterization

For all the substances included in the CAG, we collated quantitative dose estimates for declines in IQ scores, **Reference Doses for developmental neurotoxicity,** from here on referred to as RfDs. Declines in IQ scores may not always equate with the critical toxicity used to derive Health-based Guidance Values (HBGV) for single chemical risk assessments (i.e. the toxicity that appears first at doses > 0). In some cases, declines in IQ scores can become evident at doses higher than those associated with the critical toxicity. Unless declines in IQ scores represent the critical toxicity of a substance, as with e.g. lead or methyl mercury, **the RfDs used in this case study are therefore not applicable to single chemical risk assessments**.

As much as possible, we retrieved RfDs from existing evaluations of competent authorities. In some cases, however, it was necessary to conduct separate reviews to derive the respective RfDs de novo. To relate all RfDs to the same effect magnitude, we sought to collect data about exposures associated with IQ losses by 1 point.

We applied the principles of a tiered assessment which we defined in the previous Deliverable D 15.4. Because we selected RfDs related to a specific common effect, declines in IQ scores, it was possible to skip Tier 1 of the workflow and to commence the assessment in Tier 2 of the hazard assessment arm of the MRA. An overview of the RfDs used in this case study is shown in Table 8.1 below.

Table 8.1: RfDs for declines in IQ scores used in the case study

Chemical	Reference dose RfD	Dose unit	Endpoint for RfD	Derived from animal studies?	Derived from human studies?	Comments
Arsenic, inorganic	0.42	µg/kg d	IQ loss		x	
Fluoride	6.9	µg/kg d	IQ loss		x	
Methyl mercury	0.1	µg/kg d	IQ loss		x	
Lead	0.54	µg/kg d	IQ loss		x	
Organophosphate metabolites	200	nmol/g creatinine	IQ loss		x	derived by read-across from BDE 47
PCB	15	ng/kg d	IQ loss		x	
BDE 28	68.8	ng/kg d				read-across from BDE 47
BDE 47	68.8	ng/kg d	mice, locomotion	x		
BDE 99	1.68	ng/kg d	mice, total activity	x		
BDE 100	1.68	ng/kg d				read-across from BDE 99
BDE 153	3.84	ng/kg d	mice, total activity	x		
BDE 154	3.84	ng/kg d				read-across from BDE 153
BDE 183	3.84	ng/kg d				read-across from BDE 153
BDE 196	17000	ng/kg d				read-across from BDE 209
BDE 197	17000	ng/kg d				read-across from BDE 209
BDE 203	17000	ng/kg d				read-across from BDE 209
BDE 206	17000	ng/kg d				read-across from BDE 209
BDE 207	17000	ng/kg d				read-across from BDE 209
BDE 208	17000	ng/kg d				read-across from BDE 209
BDE 209	17000	ng/kg d	mice, total activity	x		

8.3.1 Inorganic arsenic

We adopted the values used by Tsuji *et al.*, (2015) in their systematic review of arsenic developmental neurotoxicity and risk assessment. Tsuji et al. evaluated several epidemiological studies of associations between inorganic arsenic exposure and IQ scores and rated the study by Hamadani *et al.*, (2011) as most suitable for quantitative risk assessments. Hamadani et al. observed a decrease in cognitive ability by 2.6 IQ points in girls for every 100 µg/L increase in speciated urinary arsenic levels. This was related to contemporaneous arsenic exposures; a window of vulnerability for inorganic arsenic and developmental neurotoxicity is poorly defined. Conversion to an IQ loss by 1 point is associated with an increase by 38.5 µg/L speciated urinary arsenic levels. By application of a toxicokinetic model, Tsuji et al. (2015) estimated that such urinary arsenic levels result from daily intakes of between 1.1 and 1.47 µg/kg d. Since these exposures represent a LOAEL, we applied an assessment factor of 3 to extrapolate to a NOAEL of 0.36 to 0.49 µg/kg d. We selected the midpoint of this range (0.42 µg/kg d) as RfD.

8.3.2 Fluoride

Grandjean (2019) reviewed recent prospective epidemiological studies of associations between fluoride exposure and developmental neurotoxicity. By application of benchmark dose modelling to the findings of these recent studies, it is estimated that maternal urinary fluoride levels of between 0.1 and 0.2 mg/L are associated with IQ losses by 1 point in children. Assuming 24 h urine volumes of between 0.8 and 2.0 L, such urinary fluoride levels are associated with a daily maternal fluoride excretion of 0.1 – 0.4 mg/d. Rugg-Gunn *et al.,* (2011) have recorded the relationship between total fluoride intake and daily urinary fluoride excretion. Based on 8 studies among adults with a total of 269 data pairs (Fig 3 in Rugg-Gunn 2011) it can be estimated that a daily excretion of 0.1 to 0.4 mg fluoride is to be expected with daily intakes of 0.2 to 0.8 mg. Assuming a body weight of 70 kg, this converts to intakes of 2.8 and 11 µg/kg d. We adopted the average of these values as RfD (6.9 µg/kg d).

8.3.3 Methyl mercury

Evidence of declines in cognitive ability after maternal methyl mercury exposure during pregnancy comes from three main epidemiological cohorts, those in the Faroe Islands, New Zealand and the Seychelles. Reviewing data from these three cohorts, Rice *et al.*, (2003) conducted benchmark dose modelling on the Faroe Islands data and estimated that maternal mercury levels of between 4 and 25 ppm are associated with IQ losses by 1 point in their children. Toxicokinetic modelling shows that these hair levels result from maternal daily methyl mercury intakes of between 0.447 and 1.9 µg/kg d. By application of an assessment factor of 10 (to account for differences in maternal toxicokinetics), Rice et al. estimated a daily intake of 0.1 µg/kg d as tolerable. However, as pointed out by Groth (2017), this is likely to underestimate risks as several more recent studies have documented declines in IQ scores at maternal mercury hair levels around 12 to 15 ppm. Nevertheless, we adopted 0.1 µg/kg d as RfD for our case study.

8.3.4 Lead

We followed the considerations of EFSA's CONTAM panel (EFSA 2010). Based on the study by Lanphear *et al.*, (2005) the Panel estimated that a blood lead level in children of 12 µg/L is associated with an IQ loss of 1 point. By toxicokinetic modelling EFSA converted this blood lead level into a daily intake of 0.5 µg/kg d. Taking account of a fetal/maternal blood lead ratio of 0.9, this is equivalent to a daily intake of 0.54 µg/kg d by expectant mothers. We adopted this value as the RfD for our case study.

8.3.5 Organophosphates

Several epidemiological studies, analyzed by Engel *et al.*, (2016), have shown associations between maternal exposure to organophosphate pesticides and declines in cognition in their offspring. Exposure to organophosphates is commonly assessed in terms of urinary dialkylphosphates (DAP). DAPs are non-specific metabolites which can be formed during the break-down of several different organophosphate pesticides. Jusko *et al.* (2019) observed declines in the non-verbal IQ of 6-year-old children with rising total DAP urinary levels in pregnancy urine, but only with urine samples from early and late pregnancy. The association was non-linear, with more pronounced declines at low exposures. From Fig 3 in their paper, we estimated that total urinary DAP levels (late pregnancy) of 200 nmol/g creatinine are associated with a 1-point IQ loss. We selected this value as the RfD.

8.3.6 PCBs

From the study of IQ loss in children of PCB-exposed mothers by Jacobson *et al.* (2002), a benchmark concentration of 0.63-0.71 µg/g lipid in mother's milk is associated with a benchmark response of 5 % in terms of full-scale IQ loss (benchmark dose, lower limit, see Table 8.3 in Jacobson et al.). This value applies to all PCBs. If adipose tissue constitutes 20 % of an adult's body weight, these PCB lipid levels correspond to PCB body burdens of between 126 and 142 µg/kg body weight. Based on a half-life of 5000 days (9.5 years) and an absorbed fraction of 95 %, the daily PCB maternal intakes that will give rise to such body burdens at steady state can be estimated as 26-30 ng/kg d. Considering that the benchmark concentrations given by Jacobson et al. do not correspond to IQ losses of 1 point, we lower these values and chose 15 ng/kg d as the RfD (assessment factor of 2).

8.3.7 Polybrominated Diphenylethers (PBDEs)

We adopted the congener-specific values for BDE 47, 99, 153 and 209 which EFSA (2011) used for margin of exposure considerations related to developmental neurotoxicity. These values are derived from rodent studies. EFSA regarded the available data for other congeners as too unreliable to establish similar values. To approximate the potency of other prevalent BDE congeners, we used the read-across approach by (Martin *et al.*, 2017). The RfD for BDE congeners are shown in Table 8.1.

8.3.8 Manganese

Several studies have shown associations between environmental manganese exposures and declines in cognitive ability in children; however, results have been mixed depending on which biomarker was used to estimate exposures. The recent systematic review and meta-analysis of cross-sectional studies by Liu *et al.* (2020) showed that elevated manganese in hair and drinking water, but not blood or teeth, was associated with poorer cognitive performance. The meta-analysis of hair manganese levels suggests that 10-fold increases result in a 2.5-point loss in full scale IQ tests. The mean hair manganese levels reported in many studies exceeded 0.55 µg/g hair, however, we could not retrieve information that would have allowed us to convert hair levels into estimated daily intakes. A RfD for manganese could therefore not be estimated.

8.4 Exposure assessment

For the substances considered together in this MRA, we could not locate studies that documented measurements of multiple chemicals in one and the same sample. We therefore collated single chemical exposure data from different sources and made the unspoken assumption that the general population is exposed to all members of the CAG simultaneously, most of the time. We

further assumed that basing the assessment on median exposure levels for multiple substances gives a reasonable approximation of the cumulative exposure patterns at a population level. To construct a worst-case scenario, we also investigated cumulative exposures with all substances at higher exposures.

Table 8.2: Median and high exposures for the chemicals in the DNT CAG

Chemical	Median	High exposures, often 95th percentile	Dose unit	Comments
Arsenic, inorganic	0.11	0.54	µg/kg d	EFSA 2021, median lower bound, P95 upper bound in children, Tab 11, p 24
Fluoride	10	35	µg/kg d	Guth 2020, Europe, mean in low F areas, mean in high F areas
Methyl mercury	0.03	0.16	µg/kg d	EFSA 2012, median middle bound, P95 middle bound, Tab 11, p 46
Lead	0.83	1.64	µg/kg d	EFSA 2010, women, median of average and high consumers, Tab 29, p 61
Organophosphate metabolites	316	810	nmol/g creatinine	Ye 2008, total dialkylphosphates, urinary levels, Tab 3
PCB, non-dioxin like	11	23	ng/kg d	EFSA 2005, adults, average Italy, PCB 28, 52, 101, 138, 153, 180; Tab 10, p43, P95 Annex 1, Tab 2, p 40
BDE 28	0.17	0.4	ng/kg d	Martin 2017, low: Level 2, scenario 4; high: Level 2, scenaro 2
BDE 47	0.76	6.12	ng/kg d	Martin 2017, low: Level 2, scenario 4; high: Level 2, scenaro 2
BDE 99	0.41	1.16	ng/kg d	Martin 2017, low: Level 2, scenario 4; high: Level 2, scenaro 2
BDE 100	0.32	2.39	ng/kg d	Martin 2017, low: Level 2, scenario 4; high: Level 2, scenaro 2
BDE 153	0.29	0.76	ng/kg d	Martin 2017, low: Level 2, scenario 4; high: Level 2, scenaro 2
BDE 154	0.28	0,87	ng/kg d	Martin 2017, low: Level 2, scenario 4; high: Level 2, scenaro 2
BDE 183	0.26	0.84	ng/kg d	Martin 2017, low: Level 2, scenario 4; high: Level 2, scenaro 2
BDE 196	0.01	0.01	ng/kg d	Martin 2017, low: Level 2, scenario 4; high: Level 2, scenaro 2
BDE 197	0.02	0.02	ng/kg d	Martin 2017, low: Level 2, scenario 4; high: Level 2, scenaro 2
BDE 203	0.01	0.01	ng/kg d	Martin 2017, low: Level 2, scenario 4; high: Level 2, scenaro 2
BDE 206	0.04	0.04	ng/kg d	Martin 2017, low: Level 2, scenario 4; high: Level 2, scenaro 2
BDE 207	0.04	0.04	ng/kg d	Martin 2017, low: Level 2, scenario 4; high: Level 2, scenaro 2
BDE 208	0.29	0.29	ng/kg d	Martin 2017, low: Level 2, scenario 4; high: Level 2, scenaro 2
BDE 209	2.48	4.25	ng/kg d	Martin 2017, low: Level 2, scenario 4; high: Level 2, scenaro 2

As much as possible we relied on exposure data from Human Biomonitoring studies. Where this was not possible, we retrieved exposure data derived from pollutant occurrence data and food consumption data. A compilation of the exposure estimates we used for this case study is given in Table 8.2.

8.4.1 Inorganic arsenic

The RfD for IQ losses due to inorganic arsenic exposures is related to contemporaneous exposures of children. We therefore used the exposures estimated for children in Europe by EFSA (2021) and selected the median lower bound intake estimate of 0.11 µg/kg d and the 95th percentile upper bound estimate of 0.54 µg/kg d.

8.4.2 Fluoride

Comprehensive quantitative information about total fluoride intakes (drinking water and diet) in Europe is difficult to locate. We based our assessment on the values reported by Guth *et al.* (2020). In low fluoride areas in Europe, the total fluoride intake (drinking water and food) is estimated to be in the range between 5 and 14 µg/kg d. This rises to 30 to 40 µg/kg d in areas with high fluoride drinking water levels. These values compare well with estimates for Sweden (6 – 14 µg/kg d in low fluoride areas, 30 – 63 in high fluoride areas), France (total intake via diet and drinking water 29 µg/kg d) and the UK (total dietary intake, excluding drinking water 17 µg/kg d). For our assessment, we chose the midpoints of the European average ranges for low (10 µg/kg d) and high (35 µg/kg d) fluoride areas.

8.4.3 Methyl mercury

EFSA (2012) estimated methyl mercury intakes for adults in the European Union of 0.03 and 0.16 µg/kg d (median, medium bound and 95th percentile, medium bound, respectively). We based our MRA on these estimates.

8.4.4 Lead

We chose the lead intake estimated by EFSA (2010) for women of reproductive age who consume an average diet (0.83 µg/kg d). For consumers of food commodities more highly polluted with lead, the daily intake can rise to 1.64 µg/kg d which we adopted for our worst-case assessment.

8.4.5 Organophosphates

We selected the total DAP urinary levels reported by Ye *et al.* (2008). They measured a median DAP of 316 nmol/g creatinine. The 95th percentile was 810 nmol/g creatinine.

8.4.6 PCBs

We selected the daily intakes of non-dioxin-like PCBs given for adults in Italy (EFSA 2005). The daily intake of PCBs 28, 52, 101, 138, 153 and 180 was estimated as 11 ng/kg d (median). At the 95th percentile, this rose to 23 ng/kg d. The values compare well with average European intakes of 10-45 ng/kg d (EFSA 2005, Tab 10, p 43).

8.4.7 Polybrominated Diphenylethers (PBDEs)

Earlier, we conducted a congener-specific MRA for BDEs (Martin et al. 2017) which we based on the intake estimates for Europe given in EFSA (2011). For this case study, we selected as average intakes the values described in Martin et al. as “Level 2, scenario 2” which includes intakes via dust, but does not consider diets rich in fish. For higher exposures, we chose the values in “Level 2, scenario 2”. This scenario takes account of consumers of fish.

8.4.8 Manganese

We were unable to locate studies of manganese hair levels of the general population.

8.5 Mixture risk assessment

We adhered to the workflow defined for the hazard assessment arm of the MRA in the previous Deliverable D15.4 but left out Tier 1 and commenced the assessment in Tier 2 of the hazard assessment arm of the MRA. The RfDs we derived for declines in IQ scores, although mainly based on human epidemiological data, also included data from animal studies (PBDEs). For this reason, it was difficult to apply the PODI approach. We therefore used the HI method.

For the exposure assessment arm, we conducted the assessment in two Tiers. In the first Tier, we used high exposure estimates, often at the 95th percentile. We complemented this analysis by using estimates closer to median exposures in Tier 2.

8.5.1 Tier 1 assessment: High exposures

The outcome of the MRA based on high exposures is shown in Table 8.3. The RQs summed up to a value of 21.4. The RQs for inorganic arsenic, fluoride, manganese, methyl mercury, lead, organophosphates, PCBs and BDE 100 individually exceeded the value of 1. The sum of RQs for these substances amounted to 20.

Table 8.3: Risk quotients and sum of risk quotients for Tier 1 of the MRA, high exposures

Tier 1: High exposures					
Chemical	**Exposure**	**Unit**	**Reference dose**	**Unit**	**Risk Quotient**
Arsenic, inorganic	0.54	µg/kg d	0.42	µg/kg d	1.285714
Fluoride	35	µg/kg d	6.9	µg/kg d	5.072464
Manganese					2
Methyl mercury	0.16	µg/kg d	0.1	µg/kg d	1.6
Lead	1.64	µg/kg d	0.54	µg/kg d	3.037037
Organophosphate metabolites	810	nmol/g creatinine	200	nmol/g creatinine	4.05
PCB	23	ng/kg d	15	ng/kg d	1.533333
BDE 28	0.4	ng/kg d	68.8	ng/kg d	0.005814
BDE 47	6.12	ng/kg d	68.8	ng/kg d	0.088953
BDE 99	1.16	ng/kg d	1.68	ng/kg d	0.690476
BDE 100	2.39	ng/kg d	1.68	ng/kg d	1.422619
BDE 153	0.76	ng/kg d	3.84	ng/kg d	0.197917
BDE 154	0.87	ng/kg d	3.84	ng/kg d	0.226563
BDE 183	0.84	ng/kg d	3.84	ng/kg d	0.21875
BDE 196	0.01	ng/kg d	17000	ng/kg d	5.88E-07
BDE 197	0.02	ng/kg d	17000	ng/kg d	1.18E-06
BDE 203	0.01	ng/kg d	17000	ng/kg d	5.88E-07
BDE 206	0.04	ng/kg d	17000	ng/kg d	2.35E-06
BDE 207	0.04	ng/kg d	17000	ng/kg d	2.35E-06
BDE 208	0.29	ng/kg d	17000	ng/kg d	1.71E-05
BDE 209	4.25	ng/kg d	17000	ng/kg d	0.00025
Sum of RQ					**21.42991**

However, it was impossible to relate hair manganese levels to daily intakes, and a RQ could therefore not be calculated. To nevertheless include manganese exposures in our assessment, we took account of recent evaluations that raised concerns about manganese exposures in excess of tolerable levels especially with the use of infant formulae (Mitchell *et al.*, 2020). We arbitrarily assigned a value of 2 for manganese.

If all RQs that exceed 1 dropped to 1 or below, the HI would sum up to 9.4.

8.5.2 Tier 2 assessment: Median exposures

Based on median exposures, we obtained a value of 7.5 as the sum of risk quotients (Table 8.4). In this scenario, the RQs of fluoride, lead and organophosphates exceeded 1. For reasons already discussed, we set the RQ for manganese to 1.

The RQs for these four substances summed up to 5.56 or 74 % of the HI.

If the RQs for fluoride, lead and organophosphates individually were reduced to 1, the HI would come to 5.95.

Table 8.4: Risk quotients and sum of risk quotients for Tier 2 of the MRA, median exposures

Tier 2: Median exposures					
Chemical	**Exposure**	**Unit**	**Reference dose**	**Unit**	**Risk Quotient**
Arsenic, inorganic	0.11	µg/kg d	0.42	µg/kg d	0.261905
Fluoride	10	µg/kg d	6.9	µg/kg d	1.449275
Manganese					1
Methyl mercury	0.03	µg/kg d	0.1	µg/kg d	0.3
Lead	0.83	µg/kg d	0.54	µg/kg d	1.537037
Organophosphate metabolites	316	nmol/g creatinine	200	nmol/g creatinine	1.58
PCB	11	ng/kg d	15	ng/kg d	0.733333
BDE 28	0.17	ng/kg d	68.8	ng/kg d	0.002471
BDE 47	0.76	ng/kg d	68.8	ng/kg d	0.011047
BDE 99	0.41	ng/kg d	1.68	ng/kg d	0.244048
BDE 100	0.32	ng/kg d	1.68	ng/kg d	0.190476
BDE 153	0.29	ng/kg d	3.84	ng/kg d	0.075521
BDE 154	0.28	ng/kg d	3.84	ng/kg d	0.072917
BDE 183	0.26	ng/kg d	3.84	ng/kg d	0.067708
BDE 196	0.01	ng/kg d	17000	ng/kg d	5.88E-07
BDE 197	0.02	ng/kg d	17000	ng/kg d	1.18E-06
BDE 203	0.01	ng/kg d	17000	ng/kg d	5.88E-07
BDE 206	0.04	ng/kg d	17000	ng/kg d	2.35E-06
BDE 207	0.04	ng/kg d	17000	ng/kg d	2.35E-06
BDE 208	0.29	ng/kg d	17000	ng/kg d	1.71E-05
BDE 209	2.48	ng/kg d	17000	ng/kg d	0.000146
Sum of RQ					**7.525908**

8.5.3 Drivers of mixture risks

We arrayed all RQs in ascending order and identified as drivers those chemicals that increased the HI beyond 1. As shown in Figure 8.1 for the Tier 1 MRA, the RQs of all BDE congeners taken together reached 1, while the RQs of BDE 100, inorganic arsenic, PCBs, methyl mercury, manganese, lead, organophosphates and fluoride collectively took the HI above 1. In this high exposure scenario, these chemicals are drivers of mixture risks.

Based on median exposures, the number of drivers decreased somewhat, and only included methyl mercury, PCBs, manganese, fluoride, lead and organophosphates (Figure 8.2).

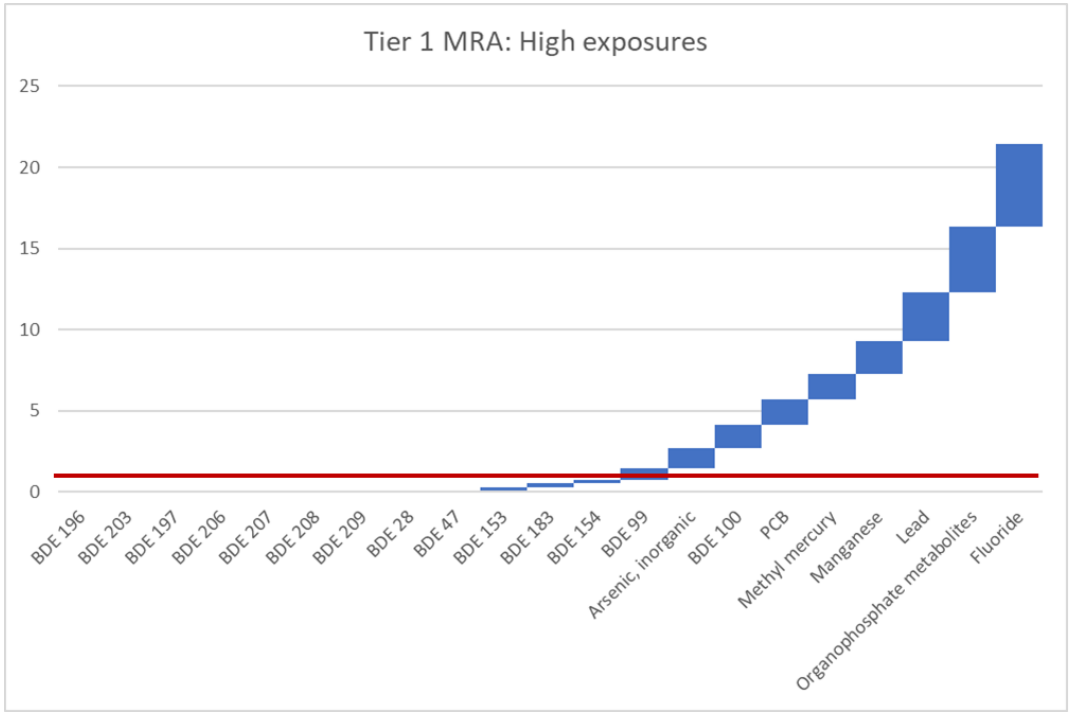


Figure 8.1: Drivers of mixture risk at Tier 1 of the MRA (exposures at the 95th percentile). The risk quotients of all chemicals in the CAG were arrayed in ascending order. The graph shows cumulative HIs. The red horizontal line depicts HI = 1; the substances above that line drive combined exposures above HI = 1.

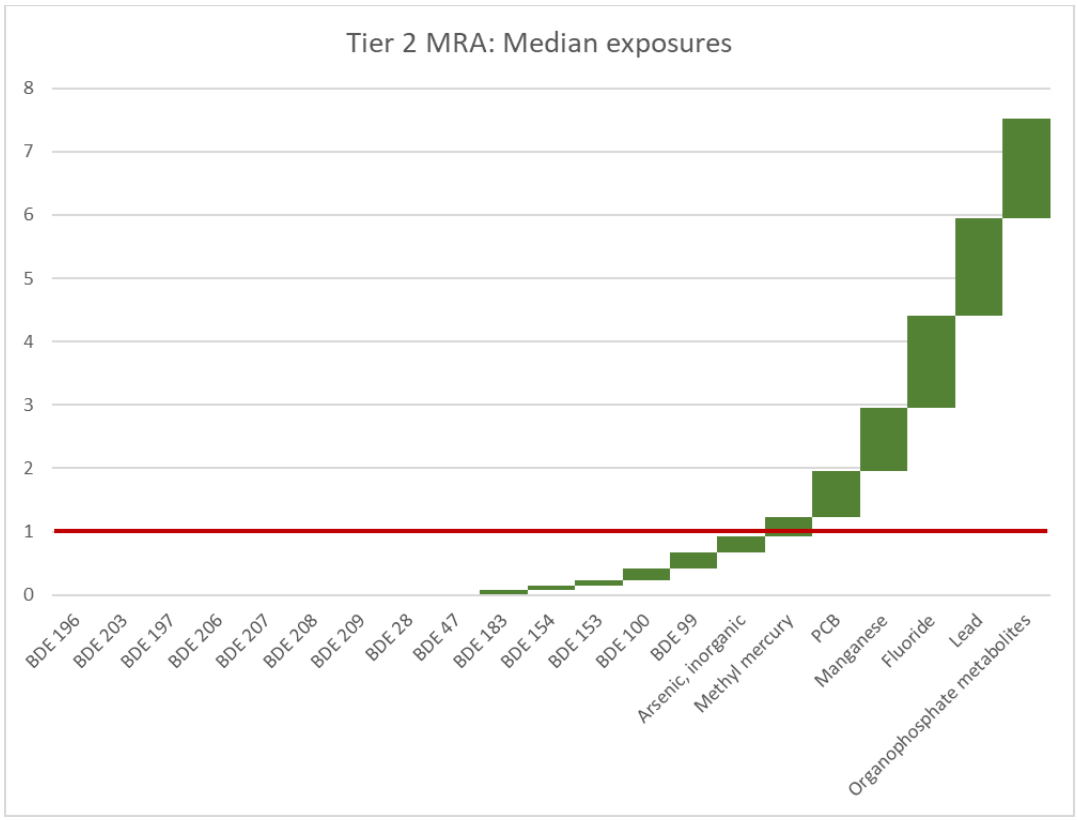


Figure 8.2: Drivers of mixture risks at Tier 2 of the MRA (median exposures)

8.5.4 Interpretation of results and uncertainty considerations

HIs larger than 1 must be interpreted as fold-exceedances of combined acceptable exposures. Since position and gradient of the underlying dose-response relationships are not known in all cases, it is difficult to infer the magnitude of risks from these data. However, as in single chemical risk assessment, exceedances of acceptable exposure indicate the need for risk management measures.

There are some uncertainties associated with our evaluation which stem from both the exposure assessment and hazard characterization components of our analysis.

8.5.4.1 Uncertainties in the hazard characterization step

Except for PBDEs, the RfDs for all substances in the CAG are based on epidemiological studies of associations with IQ loss. In many cases, such as lead and methyl mercury, this evidence forms the basis for HBGV. The RfD for fluoride is from recent, high quality prospective studies and the RfD for PCBs from the authoritative study by Jacobson et al. (2002). While associations of organophosphates with impairment of cognitive abilities are well recognized, there are some uncertainties regarding the quantification of effects in the low exposure range. Although all epidemiological studies measured IQ loss in children at various ages, there are differences in the details of the IQ tests used which may compromise the comparability of our RfDs. The extent of uncertainty introduced in this way is difficult to quantify.

The RfD we used for methyl mercury is likely an underestimation of tolerable exposures, considering that recent epidemiological studies uncovered poorer cognitive performance at lower maternal mercury hair levels.

From the study by Jacobson et al. (2002) it is difficult to delineate the contribution of PCBs from additional risks conferred by PCDD/F.

The RfDs for BDE congeners are rather uncertain. While there is good evidence of detrimental effects of PBDEs on cognition in children exposed in fetal life, the available epidemiological studies make it difficult to relate these neurodevelopmental effects to congener-specific exposure estimates.

Apart from the chemicals considered here, there are indications that bisphenol A, cadmium, phthalates and perfluorinated chemicals are associated with neurodevelopmental effects after exposures in fetal life. An evaluation of the quality of these studies, with de novo derivation of quantitative estimates of risks at low exposures could not be achieved with the resources available in this project. The omission of these substances from the MRA may have led to underestimations of risks.

8.5.4.2 Uncertainties in the exposure assessment arm of the MRA

A source of considerable uncertainty is the lack of data about co-exposures to the chemicals in our CAG. We therefore had to rely on median exposures to each individual chemical to approximate the patterns experienced on a population level. The aggregation of high exposures, often at the 95th percentile, only serves to demarcate the upper limits of high combined exposures. We regard it as unlikely that such high exposures are experienced by individual subjects.

8.6 Conclusions

Combined exposures to developmental neurotoxicants that can be evaluated as tolerable on a chemical-by-chemical basis are exceeded by rather large margins in Europe. Based on median or average exposures, we estimate 7.5-fold exceedances of combined acceptable levels. It can be

assumed that a fraction of the population experiences combined risks above these levels. This should be addressed in several subgroup analyses, e.g. of combined risks in high fluoride areas.

As seen with chemicals contributing to declining semen quality, the issue is a mixed constellation of combined risks from chemicals that comply with their single HBGV and those that already on their own exceed their individual RfDs, signaling the need for risk management measures for single chemicals. But even if the RQs of all the substances considered here stayed below 1, the sum of RQ would still exceed 1, highlighting the need for dedicated mixture risk mitigation strategies.

Our findings can provide solid guidance for future Human Biomonitoring studies fit for purpose to support mixture risk assessments. We propose that methods for the simultaneous monitoring of PBDEs, arsenic, PCBs, methyl mercury, fluoride, lead and organophosphate metabolites (DAPs) should be developed and implemented. The use of such biomarkers in future epidemiological studies could greatly improve the health impact assessment of combined exposures to developmental neurotoxicants.

8.7 References

EFSA (2005) Opinion of the Scientific Panel on contaminants in the food chain on a request from the Commission related to the presence of non-dioxin-like PCBs in feed and food, *EFSA Journal*, 284, 1-137

EFSA (2010) 'Scientific Opinion on Lead in Food', *EFSA Journal*, 8(4). doi: 10.2903/j.efsa.2010.1570.

EFSA (2011) 'Scientific Opinion on Polybrominated Diphenyl Ethers (PBDEs) in Food', *EFSA Journal*, 9(5), pp. 1–274. doi: 10.2903/j.efsa.2011.2156.

EFSA (2012) 'Scientific Opinion on the risk for public health related to the presence of mercury and methylmercury in food', *EFSA Journal*, 10(12). doi: 10.2903/j.efsa.2012.2985.

EFSA (2021) 'Chronic dietary exposure to inorganic arsenic', *EFSA Journal*, 19(1), pp. 1–50. doi: 10.2903/j.efsa.2021.6380.

Engel, S. M. *et al.* (2016) 'Prenatal organophosphorus pesticide exposure and child neurodevelopment at 24 months: An analysis of four birth cohorts', *Environmental Health Perspectives*, 124(6), pp. 822–830. doi: 10.1289/ehp.1409474.

Grandjean, P. (2019) 'Developmental fluoride neurotoxicity: An updated review', *Environmental Health: A Global Access Science Source*. Environmental Health, 18(1), pp. 1–17. doi: 10.1186/s12940-019-0551-x.

Grandjean, P. and Landrigan, P. (2006) 'Developmental neurotoxicity of industrial chemicals', *Lancet*, 368(9553), pp. 2167–2178. doi: 10.1016/S0140-6736(06)69665-7.

Groth, E. (2017) 'Scientific foundations of fish-consumption advice for pregnant women: Epidemiological evidence, benefit-risk modeling, and an integrated approach', *Environmental Research*, 152, pp. 386–406. doi: 10.1016/j.envres.2016.07.022.

Guth, S. *et al.* (2020) *Toxicity of fluoride: critical evaluation of evidence for human developmental neurotoxicity in epidemiological studies, animal experiments and in vitro analyses*, *Archives of Toxicology*. Springer Berlin Heidelberg. doi: 10.1007/s00204-020-02725-2.

Hamadani, J. D. *et al.* (2011) 'Critical windows of exposure for arsenic-associated impairment of cognitive function in pre-school girls and boys: A population-based cohort study', *International Journal of Epidemiology*, 40(6), pp. 1593–1604. doi: 10.1093/ije/dyr176.

Jacobson, J. L. *et al.* (2002) 'A benchmark dose analysis of prenatal exposure to polychlorinated biphenyls', *Environmental Health Perspectives*, 110(4), pp. 393–398. doi: 10.1289/ehp.02110393.

Jusko, T.A. *et al.* (2019) 'Organophosphate pesticide metabolite concentrations in urine during pregnancy and offspring attention-deficit hyperactivity disorder and autistic traits', *Environment International*, 131(January), pp. 1–11. doi: 10.1016/j.envint.2019.105002.

Lanphear, B. P. *et al.* (2005) 'Low-level environmental lead exposure and children's intellectual function: An international pooled analysis', *Environmental Health Perspectives*, 113(7), pp. 894–899. doi: 10.1289/ehp.7688.

Liu, W. *et al.* (2020) 'Biomarkers of environmental manganese exposure and associations with childhood neurodevelopment: A systematic review and meta-analysis', *Environmental Health: A Global Access Science Source*. Environmental Health, 19(1), pp. 1–22. doi: 10.1186/s12940-020-00659-x.

Martin, O. V. *et al.* (2017) 'A human mixture risk assessment for neurodevelopmental toxicity associated with polybrominated diphenyl ethers used as flame retardants', *Environmental Health Perspectives*, 125(8). doi: 10.1289/EHP826.

Mitchell, E. J. *et al.* (2020) 'Estimating daily intakes of manganese due to breast milk, infant formulas, or young child nutritional beverages in the United States and France: Comparison to sufficiency and toxicity thresholds', *Journal of Trace Elements in Medicine and Biology*. Elsevier GmbH, 62. doi: 10.1016/j.jtemb.2020.126607.

Rice, D. C., Schoeny, R. and Mahaffey, K. (2003) 'Methods and rationale for derivation of a reference dose for methylmercury by the U.S. EPA', *Risk Analysis*, 23(1), pp. 107–115. doi: 10.1111/1539-6924.00294.

Rugg-Gunn, A. J., Villa, A. E. and Buzalaf, M. R. A. (2011) 'Contemporary biological markers of exposure to fluoride', *Monographs in Oral Science*, 22(June), pp. 37–51. doi: 10.1159/000325137.

Tsuji, J. S. *et al.* (2015) 'Low-level arsenic exposure and developmental neurotoxicity in children: A systematic review and risk assessment', *Toxicology*. Elsevier Ireland Ltd, 337, pp. 91–107. doi: 10.1016/j.tox.2015.09.002.

Ye, X. *et al.* (2008) 'Urinary metabolite concentrations of organophosphorous pesticides, bisphenol A, and phthalates among pregnant women in Rotterdam, the Netherlands: The Generation R study', *Environmental Research*, 108(2), pp. 260–267. doi: 10.1016/j.envres.2008.07.014.

9 Case study 4: Occupational exposure to hexavalent chromium, nickel and PAHs and lung cancer

Ana Tavares, Susana Viegas, Henriqueta Louro, Thomas Goen, Tiina Santonen, Mirjam Luijten, Maria João Silva

9.1 Grouping criteria

Occupational co-exposure to polycyclic aromatic hydrocarbons (PAHs), hexavalent chromium (Cr(VI)), nickel (Ni) and other metals occurs in several industrial settings, e.g. aluminium production, iron and steel founding industries and coke oven plants (IARC, 2012; Wang et al., 2015). A study in Finland demonstrated that Cr, Ni, benzo(a)pyrene (BaP) and other PAHs were among the main substances to which welders and flame cutters were exposed to, between 2007 and 2009 (EU-OSHA, 2014). Moreover, exposure to Cr(VI) and Ni is frequent in industrial activities, such as welding operations (Pesch et al., 2018; Weiss et al., 2013). Other activities, such as painting have been also reported to render high exposure to mixtures of Cr and PAHs, as both substances are common paints components (IARC, 2012).

The above-mentioned likelihood of human co-exposure, mainly through inhalation, and the similarity of the main associated heath outcome, i.e., lung cancer, promoted the development of a case study on Cr(VI), Ni and PAHs and lung cancer. In this case study, a mixture risk assessment (MRA) exercise was conducted based on external exposure and HBM data gathered from the literature on occupational exposure to these substances. In a second phase real data from the occupational study run under WP 8, Task 8.5, will be used for MRA.

9.2 Hazard characterization

Cr(VI), Ni and PAHs (e.g. benzo(*a*)pyrene) are lung carcinogens (IARC, group 1) (IARC, 2012, 2014). Occupational exposure to Cr(VI) has been associated with lung cancer and with several other forms of cancer (nasal and sinus, trachea, and bronchus) (CDC/NIOSH, 2013). Similarly, Ni exposure causes lung fibrosis and cancer apart from immunological sensitization, epithelial dysplasia and asthma (Annangi et al., 2016; Barchowsky et al., 2003; Grimsrud et al., 2002). Long-term exposure to PAHs may cause an increased risk of lung cancer, as well.

At cellular and molecular levels Cr(VI) and Ni are able to induce oxidative stress, DNA strand breaks and chromosomal alterations, such as chromosome breaks and micronuclei (MN) (Annangi et al., 2016; IARC, 2018). Both Cr(VI) and Ni can also interfere with DNA repair mechanisms, enhancing the genotoxic effect of other agents (Annangi et al., 2016; Das et al., 2019). Likewise, PAHs such as benzo(*a*)pyrene (BaP), can cause DNA adducts, and chromosomal alterations, reflected in the induction of chromosomal breaks, sister chromatic exchanges (SCEs) and MN (Kim et al., 2013; EC, 2016). Some *in vitro* and *in vivo* studies have shown that combinations of Cr, Ni and PAHs can lead to adverse effects that are higher than those observed for each single substance (Feng et al., 2003; Peng et al., 2015; Sánchez-Martín et al., 2015).

Current regulatory practices are usually based on single chemical substances (Kortenkamp and Faust, 2018), without integrating the possibility of co-exposures or aggregated exposures. The failure to account for the combined effects of these substances can lead to a risk underestimation.

9.3 Exposure assessment retrieved from the literature

Pubmed® database was searched between May and June 2020, using combinations of key words and/or Medical Subject Headings (MeSH) terms (e.g.: “Human Biomonitoring”; “PAHs AND Ni AND Cr(VI)”) to find articles presenting HBM measurements. Only journal articles published between 2000 and 2020, with abstracts, and written in English, Spanish, French or Portuguese were considered. The following inclusion criteria were established:

1. Studies presenting data on biomarkers of exposure to chromium, nickel and/or PAHs
2. HBM performed within occupational settings
3. Studies performed in EU countries

Firstly, titles and abstracts were screened for inclusion/exclusion, and the selected articles proceeded to full text analysis. In a second approach, references of the retrieved articles were searched for more relevant articles. Overall, 356 articles were retrieved from the Pubmed search. After removing duplicates, applying selection criteria, and conducting a second search approach throughout articles references, a total of 24 articles were selected. A total of 270 documents were excluded for not meeting the selection criteria. Data on external (occurrence data in food, air, water) and internal exposure monitoring (Human Biomonitoring, HBM) from each selected study were collected. Biomarkers of exposure in urine were presented in 21 out of 24 articles, while 5 articles referred biomarkers in blood (blood, plasma and/or erythrocytes), 5 articles included biomarkers in exhaled breath condensate (EBC), and only 1 article presented measurements in hair and saliva. Most studies (n=15) measured both urinary chromium (U-Cr) and urinary nickel (U-Ni), and only two studies measured U-Cr, U-Ni and urinary 1-hydroxypyrene (U-1-OHP), the most commonly measure biomarker for PAHs exposure.

9.4 Mixture risk assessment

Using the values reported in the previous literature search, the Sum of the Risk Quotients (SRQ) was calculated whenever possible, using Equation 1, to infer on the risk that the mixture represents to human health, in each study. Risk Quotient (RQ) is defined as the ratio between the exposure level (EL), i.e., the concentration of a substance/metabolite in a biological specimen, and the acceptable level (AL) or reference doses (RfD) — EL_i/AL_i.

$$SRQ = \sum_{i=1}^{n} \frac{EL_i}{AL_i}$$

(Equation 1)

Table 9.1 presents the Reference Values for Cr, Ni and BaP in air used for RQ and the Sum of Risk Quotients (SRQ) calculated in this study. In order to contribute to a reduction in the risk of cancer, health-based reference values derived from dose-response curves for the three substances were chosen, assuming an excess of lung cancer risk of 4:1000 population. Acknowledging the necessity of coherence between the reference values chosen for air measurements and the corresponding values in urine (HBM) to enable a comparison of the obtained results, calculations using equation 1 were also performed with HBM data using the exposure equivalents in urine as limit values (Table 9.1). In this case, due to the lack of HBM guidance values related to risks, the sum of RQ was designated as “Background Exposure Exceedance Score” (BEES).

Table 9.1: Reference Values used for risk assessment calculations

Substances	Reference values in air [a]	Exposure equivalents in urine
Cr (VI)	1 µg/m^3 [c]	0.4 µg/L [b]
Nickel [d]	6 µg/m^3	0.9 µg/L [b]
Benzo(a)pyrene	0.07 µg/m^3	2.16 µg/L [e] (for 1-OHP)

[a] Limit values established in the Technical Rules for Hazardous Substances, AGS, 2021.
[b] Calculated based on the List of MAK and BAT Values, 2020.
[c] Also applicable for exposure to welding fumes.
[d] Nickel metal, nickel oxide, nickel carbonate, nickel sulphide, sulfidic ores.
[e] Calculated based on RAC note from ECHA, 2018.

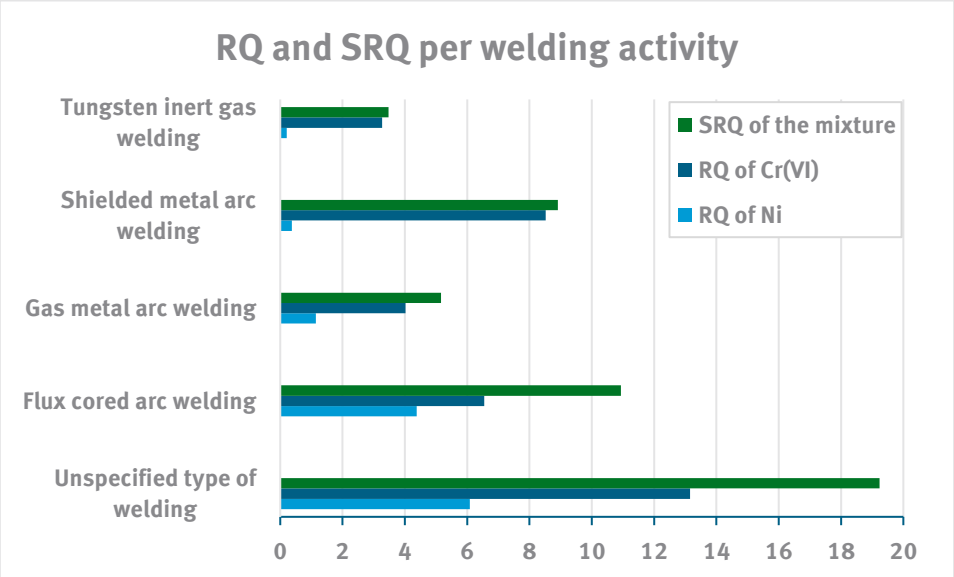
A value of SRQ or BEES below 1, indicates low concern after workers exposure to the mixture. Conversely, if SRQ or BEES >1, it indicates that it exceeds the risk level considered to be acceptable, and the exposure raises concern in terms of its impact on human health.

Among the studies reporting the external exposure levels, all studies conducted in welding exposure scenarios measured both U-Cr and U-Ni and no study could be identified concerning the 3 substances. The SRQ values obtained ranged from 0.12 to 31.7, with and extreme value of 252.8 for settings of Gas Metal Arc Welding with massive or flux cored wire of stainless steel (high exposure group). Of the 14 studies considered, only 4 yielded values of SRQ<1.

Conversely, among the 16 studies reporting exposure levels of U-Cr and U-Ni, all welding activities resulted in BEES>1 based on the HBM data. All studies reported levels of U-Cr above the reference values (RQ>1), and most resulted in a RQ>1 for U-Ni levels.

When comparing the results using air monitoring or HBM, there are differences regarding SRQ and BEES results. While some studies revealed SRQ<1 when considering air measurements, BEES>1 were observed in all studies when considering HBM data on welding activities. Figure 9.1 depicts the mean RQ, SRQ and BEES values obtained from each welding activity.

a)



b)

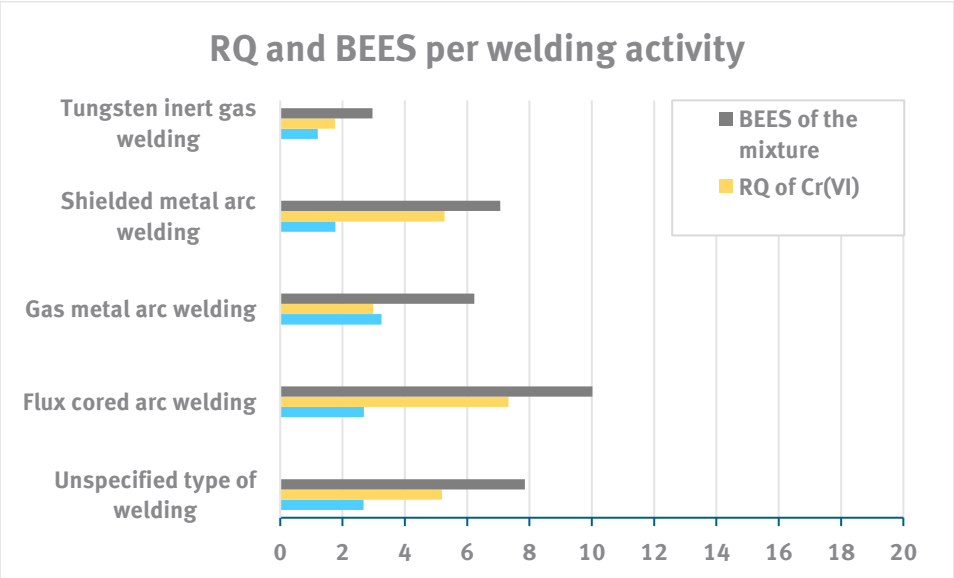


Figure 9.1: Mixture risk assessment outcomes per welding activity based on exposure levels of Cr and Ni: a) measured in air (mean of RQ and SRQ obtained from the analyzed studies); and b) measured in workers' urine (mean of RQ and BEES obtained from the same studies).

Only two studies retrieved from the literature, presented biomarkers of exposure in urine for the three substances, Cr, Ni, and PAHs. These studies were conducted in hazard waste incinerator settings, but they did not present air measurements. Calculations for these studies showed a value of BEES>1 for all exposure scenarios, even among workers performing activities considered of low exposure or non-exposed, such as laboratory and administrative activities, respectively. Figure 9.2 depicts the mean RQ and BEES values obtained from the analyzed studies for each activity conducted within the hazard waste incinerators context.

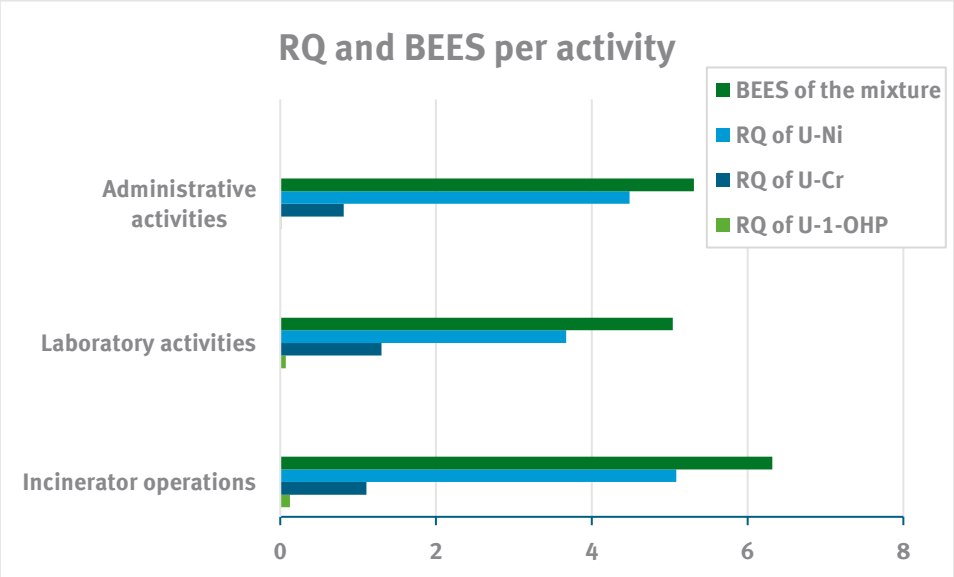


Figure 9.2: Mixture risk assessment outcomes per activity performed in hazard waste incinerators based on exposure levels of Cr, Ni, and 1-OHP measured in workers' urine (mean of RQ and BEES obtained from the analyzed studies).

9.5 Lessons learned and recommendations

- In this study, HBM data showed lower health risks from co-exposure to mixtures than air levels.

Both the risk quotient calculations and their sum differed when based on air measurements or when based on HBM data. The lower RQ and BEES obtained using HBM data may be related, for instance, with workers using respiratory protection equipment, or to uncertainties related to the air-urine correlations at the low exposure levels reported in some studies. Nevertheless, HBM data reflects occupational exposure across all routes of exposure and from all sources, including exposure to environmental pollutants. This is particularly relevant in case of metals and PAHs since other exposure routes, besides inhalation, have an important role in the overall exposure, being revealed by the HBM measurements.

- The risks from occupational co-exposure to chromium (Cr), nickel (Ni) and PAHs may become underestimated if considered separately.

This is evidenced by significantly higher SRQ and BEES in many scenarios when compared to the RQs calculated for the Cr and Ni separately. In some scenarios, e.g., gas metal arc welding (HBM data) RQ>1 were obtained for the single metals and both metals seemed to contribute to the combined risk significantly. In other ones, e.g., tungsten inert gas welding (air measurements) risk seemed to be driven by Cr. Ideally, these findings should be translated into risk mitigation actions, in order to reduce workers' exposure and to protect their health.

- The selection of reference values for conducting MRA based on HBM data is a critical point due to the limited existence of reference values with regulatory acceptance.

This is particularly relevant since we are dealing with different regulatory frameworks (REACH and OSH regulation) that defined guidance values based on single substances.

- Effect biomarkers rarely included in HBM occupational studies.

The usefulness of incorporating effect biomarkers in occupational HBM studies, with the goal of capturing early biological effects from possible additive or interactive effects is difficult to demonstrate due to their limited used in the studies analyzed. Nevertheless, the use of effect biomarkers in occupational health studies where carcinogenic mixtures exposure occurs should be recommended, in order to increase the possibility of detecting early signs of disease associated to worker groups with higher risk.

- Methodologies used in occupational studies (exposure and risk assessment) would benefit from harmonization and standardization.

To provide results that can support regulatory actions, methodologies used in occupational studies (exposure and risk assessment), for instance which work shifts are assessed and how data is reported in technical and scientific reports, need harmonization to enable appropriate comparison of obtained results and to evaluate the efficacy of the risk management measures already in place (technical and regulatory).

9.6 References

Annangi B, Bonassi S, Marcos R, Hernández A. Biomonitoring of humans exposed to arsenic, chromium, nickel, vanadium, and complex mixtures of metals by using the micronucleus test in lymphocytes. Mutat Res - Rev Mutat Res. 2016;770(2015):140–61. doi:10.1016/j.mrrev.2016.03.003

Ausschuss für Gefahrstoffe (AGS). Technische Regeln für Gefahrstoffe. 2021. Available at: www.baua.de/ags

Barchowsky A, O'Hara KA. Metal-induced cell signaling and gene activation in lung diseases. Free Radic Biol Med. 2003;34(9):1130–5. doi:10.1016/S0891-5849(03)00059-5

CDC/NIOSH. Criteria for a Recommended Standard. Occupational exposure to hexavalent chromium. DHHS (NIOSH) Publication No. 2013–128. 2013.

Das KK, Reddy RC, Bagoji IB, Das S, Bagali S, Mullur L, et al. Primary concept of nickel toxicity - An overview. J Basic Clin Physiol Pharmacol. 2019;30(2):141–52. doi:10.1515/jbcpp-2017-0171

Deutsche Forschungsgemeinschaft (DFG). List of MAK and BAT Values Permanent Senate Commission for the Investigation of Health Hazards of Chemical Compounds in the Work Area. 2020. Report 56.

EU-OSHA. Lißner L, Hamburg K. Exposure to carcinogens and work-related cancer: A review of assessment methods European Risk Observatory Report European Agency for Safety and Health at Work European Agency for Safety and Health at Work Exposure to carcinogens and work-related cancer: A. 2014. 1831. doi:10.2802/33336

European Chemicals Agency (ECHA). Note on reference dose-response relationship for the carcinogenicity of pitch, coal tar, high temperature and on PBT and vPvB properties. 2018.

European Commission. SCOEL/REC/404 Polycyclic Aromatic Hydrocarbon mixtures containing benzo[a]pyrene (PAH). Recommendation from the Scientific Committee on Occupational Exposure Limits. Luxembourg: Publications Office of the European Union; 2016. doi:10.2767/977505

Feng Z, Hu W, Rom WN, Costa M, Tang MS. Chromium(VI) exposure enhances polycyclic aromatic hydrocarbon-DNA binding at the p53 gene in human lung cells. Carcinogenesis. 2003;24(4):771–8. doi:10.1093/carcin/bgg012

Grimsrud TK, Berge SR, Haldorsen T, Andersen A. Exposure to different forms of nickel and risk of lung cancer. Am J Epidemiol. 2002;156(12):1123–32. doi:10.1093/aje/kwf165

IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Chemical agents and related occupations. IARC Monogr Eval Carcinog Risks Hum. 2012;100(Pt F):9–562.

IARC. Arsenic, Metals, Fibres and Dusts: Nickel and Nickel Compounds. In: IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. 2012. p. 169–218.

IARC. Benzo(a)pyrene. In: Encyclopedia of Toxicology: Third Edition. 2014. p. 423–8. doi:10.1016/B978-0-12-386454-3.00250-5

IARC. Chromium (VI) compounds. In: IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. 2018. p. 147–67.

Kim KH, Jahan SA, Kabir E, Brown RJC. A review of airborne polycyclic aromatic hydrocarbons (PAHs) and their human health effects. Environ Int. 2013;60:71–80. doi:10.1016/j.envint.2013.07.019

Kortenkamp A, Faust M. Regulate to reduce chemical mixture risk. Science (80-). 2018 Jul;361(6399):224–6. doi:10.1126/science.aat9219

Peng C, Muthusamy S, Xia Q, Lal V, Denison MS, Ng JC. Micronucleus formation by single and mixed heavy metals/loids and PAH compounds in HepG2 cells. Mutagenesis. 2015;30(5):593–602. doi:10.1093/mutage/gev021

Pesch B, Lehnert M, Weiss T, Kendzia B, Menne E, Lotz A, et al. Exposure to hexavalent chromium in welders: Results of the WELDOX II field study. Ann Work Expo Heal. 2018;62(3):351–61. doi:10.1093/annweh/wxy004

Sánchez-Martín FJ, Fan Y, Carreira V, Ovesen JL, Vonhandorf A, Xia Y, et al. Long-term coexposure to hexavalent chromium and B[a]P causes tissue-specific differential biological effects in liver and gastrointestinal tract of mice. Toxicol Sci. 2015;146(1):52–64. doi:10.1093/toxsci/kfv070

Wang T, Feng W, Kuang D, Deng Q, Zhang W, Wang S, et al. The effects of heavy metals and their interactions with polycyclic aromatic hydrocarbons on the oxidative stress among coke-oven workers. Environ Res. 2015;140:405–13. doi:10.1016/j.envres.2015.04.013

Weiss T, Pesch B, Lotz A, Gutwinski E, Van Gelder R, Punkenburg E, et al. Levels and predictors of airborne and internal exposure to chromium and nickel among welders-Results of the WELDOX study. Int J Hyg Environ Health. 2013 Mar;216(2):175–83. doi:10.1016/j.ijheh.2012.07.003

10 Case study 5: Hazard index assessment of the combined chronic dietary exposure to heavy metals in the general population

Annick van den Brand, Gerda van Donkersgoed, Marcel J.B. Mengelers, Mirjam Luijten (RIVM)

10.1 Grouping approach

Within HBM4EU, four metals were identified as priority metals for the exposure of the general population, being arsenic (As), cadmium (Cd), mercury (Hg) and lead (Pb). In humans, some of these metals exert a neurodevelopmental effect but all of them exert a nephrotoxic effect. This nephrotoxic effect can be specified to an adverse effect on the glomerulus and the proximal tubule of the kidney. To gain more insight into the mode of action of this nephrotoxic effect, an adverse outcome pathway has been developed for these metals in this work package (Schillemans et al., 2021 *submitted*). The grouping of these metals for a mixture risk assessment was based on a shared effect on the kidney, with a focus on the inorganic forms of arsenic and mercury. The population under investigation is the general population, chronically exposed to these metals through contaminated food.

10.2 Hazard assessment

Two hazard scenarios were considered in this assessment. The first scenario considered all metals equipotent. To assess this scenario, the cumulative exposure of the metals was based on the toxicological reference value (TRV) of the index compound cadmium, being 0.36 µg/kg bw/day (which equals a TWI (tolerable weekly intake) of 2.5 µg/kg bw/w) assuming dose additivity (EFSA, 2009a).

The second scenario did not consider the metals equipotent. In this scenario, the TRVs for the metals were used to derive relative potency factors, with cadmium being the index compound (see Table 10.1). The (provisional TWI for inorganic mercury as established by EFSA (2012) and JECFA (2011) is 4 µg/kg bw/w. This corresponds to 0.57 µg/kg bw/day. Compared to cadmium, this results in a relative potency factor of 0.63. For lead, no TRV has been established by EFSA, since there is no evidence for a threshold for critical lead-induced effects. However, a point of departure (POD) of 0.63 µg/kg bw/day ($BMDL_{10}$) was derived based on the effects of lead on chronic kidney disease. Applying a default uncertainty factor of 3.16, for the uncertainty related to toxicodynamics, the hazard value that can be derived for lead is 0.20 µg/kg bw/day. This results in a relative potency factor for lead of 1.79 compared to cadmium. For inorganic arsenic, no TRV has been derived by EFSA or JECFA, nor was a POD for nephrotoxic effects identified. As it is known that arsenic also targets the kidney (Schillemans et al., 2021 *submitted*), a relative potency factor for inorganic arsenic was derived based on internal reference values (see occupational PODI section). Based on the difference between these reference values of cadmium (4 µg/L) and arsenic (7 µg/L) in urine, a relative potency factor of 0.57 was obtained. This factor was used to derive the hazard value for inorganic arsenic by multiplying it with the hazard value of cadmium, resulting in a hazard value for inorganic arsenic of 0.63 µg/kg bw/day.

Table 10.1: Overview of the available health-based guidance values and related points of departure (POD) and the derived hazard values and relative potency factors (RPFs) used for the cumulation of the metals in this case study.

	TRV (µg/kg bw/week)	TRV (µg/kg bw/week)	POD (µg/kg bw/week)	Species	Uncertainty factor	Hazard value (µg/kg bw/week)	RPF	Source
Cadmium	2.5	0.36	1.4 (BMDL05)	Human	* 3.9	0.36	1	EFSA, 2009a
Inorganic mercury	4	0.57	60 (BMDL10)	Rat	100 (10 x 10)	0.57	0.63	JECFA, 2011; EFSA, 2012
Lead	-	-	0.63 (BMDL10)	Human	*** 3.16*	0.20	1.79	EFSA, 2010
Inorganic arsenic	No TRV/POD considering nephrotoxic effects available. Based on potency difference internal TRV urine cadmium (4 ug/L) -arsenic (7 ug/L) (factor 1.75)					0.63	0.57	See PODI section

* chemical specific adjustment factor to account for human variability in urinary cadmium within each dose-subgroup.

** chemical specific adjustment factor for lead is not possible based on the data reported in EFSA (2010); a default of 3.16 was used.

10.3 Exposure assessment

10.3.1 Occurrence data

The occurrence data used in this assessment were obtained from the EFSA Data Warehouse from 13 European countries in the circle of trust and all combined into one occurrence data database. In addition, national occurrence data were extracted from the EFSA Data Warehouse by Portugal and added to the combined occurrence data from the other countries. Data related to the param codes containing cadmium, lead, mercury, and arsenic and between the years 2014-2018 were extracted. Data were cleaned by removing targeted strategy samples, duplicate or otherwise aberrant samples (e.g. one sample with an inorganic mercury concentration of 23 kg/kg) and samples with invalid concentration units. Samples coded as ‘lead and derivatives’, ‘cadmium and derivatives’ and ‘mercury and derivatives’ were removed because additional information related to the specific compound analyzed was missing. Samples coded as ‘arsenic and derivatives’ or ‘arsenic’ were translated to ‘total arsenic’ samples, in accordance to the methods of EFSA (EFSA, 2021). In addition, all ‘methylmercury’, ‘organic mercury’ and ‘organic arsenic’ entries were removed as these were not relevant for this assessment.

Subsequently, the ‘total mercury’ entries were translated to ‘inorganic mercury’ according to the fractions indicated by EFSA in their exposure assessment (EFSA, 2012). Inorganic mercury was considered 20 % of the total mercury reported in fish samples and 50 % in shell fish samples. All other samples that reported total mercury were considered 100 % inorganic mercury. In addition, the ‘total arsenic’ entries were translated to ‘inorganic arsenic’ according to the median fractions reported by EFSA (EFSA, 2021). First, for the samples which reported both total arsenic and inorganic arsenic, the total arsenic samples were removed. Then, the fraction of inorganic arsenic in the total arsenic entries was assigned according to the average fractions reported recently by EFSA (Table 10 in EFSA, 2021). No average fraction of inorganic arsenic to total arsenic was given for fish, in contrast to other seafood. Total arsenic levels in fish were however not assigned a fraction of inorganic arsenic in this assessment, as EFSA also reported that 94 % of the fish

sample were reported below the detection or quantification limits of inorganic arsenic, while the same samples reported quantified levels for total arsenic. In addition, samphire ("*zeekraal*") samples from the Netherlands analyzed for arsenic were originally coded as 'leafy vegetables'. 'Leafy vegetables' were therefore removed from the occurrence data and recoded as 'sea weeds'.

After cleaning the data, a total of 108,226 samples were used for the exposure assessment. Most of this data was provided for lead with 40 % of the total samples, followed by cadmium with 38 %.

10.3.2 Consumption data

The food consumption data that were used in this exposure assessment were obtained from 10 countries: the Netherlands (2012-2016, 1-79 years old), Belgium (2004, 14-105 years old), Cyprus (2003, 11-15 years old), Czech Republic (2004, 4-64 years old), Denmark (2003-2008, 4-75 years old), Spain (2011, 18-71 years old), France (2005-2012, 3-79 years old), Greece (1988-2003, 3-94 years old), Slovenia (2008, 18-65 years old) and the United Kingdom (2000-2001, 19-64 years old). More information regarding the food consumption data can be found in Crépet et al. (2019).

The consumption data were classified as FoodEx1 or FoodEx2. In order to obtain uniformity, all data were translated to FoodEx1. Consent to use more recent food consumption datasets was obtained from several European countries. The requests for these data have been made to EFSA by the end of 2020 and an affirmative response from EFSA was received on March 31st 2021. The transfer of the updated consumption data by EFSA is currently pending (April 2021) but the exposure assessments will be updated with these consumption data upon further reporting (publication in a scientific journal, expected by the end of 2021).

10.3.3 Other data

Refinement of the concentrations of the metals in consumed foods can be obtained by correcting for the effects of dilution and food processing that can lead to changes in the concentrations of the metals in the food (e.g. due to the dilution of the concentration by adding water to coffee or degradation of the product as a result of heating processes). Therefore, several dilution factors following EFSA (2021 and 2012) were applied to the products such as coffee and porridge. Regarding the consumption of tea, we accounted for the addition of water (possible additional exposure) to the tea leaves by translating the consumed tea to a fraction tea leaves and a fraction drinking water. In addition, when no concentration data was available for a consumed sub product, we translated the food product to a higher hierarchic FoodEx level. Furthermore, the metals are often considered stable in food and are resistant to the effects that are generally encountered during the processing of food. Such processing factors were not identified or applicable and also not reported or omitted by EFSA in their exposure assessments of these metals.

10.3.4 Exposure assessment

The cumulative chronic exposure assessment was performed using the Monte Carlo Risk Assessment tool v9.1 (MCRA, https://mcra.rivm.nl). The exposure was assessed probabilistically using the observed individual means (OIM) model.

The food consumption as well as the occurrence data were both codified according to the FoodEx1 classification system. The dietary exposure to the metals was therefore assessed by multiplying the daily consumption for each food at the individual level with the average occurrence of the metals in this food. The exposure assessment was performed by assigning the reported samples below the limit of quantification (LOQ) or detection (LOD) with a value of zero (lower bound (LB) scenario). The summed daily exposure of each individual to every metal was combined according to two scenarios: 1) assuming equipotency and 2) assuming non-equipotency. In the 1st scenario,

the individual's exposure to all metals was combined by summing the calculated exposure to every metal. In the 2nd scenario, the individual's exposure to all metals was combined by multiplying with the RPF of each metal to the index compound (see section hazard assessment) before the calculated exposure was summed. Subsequently, the total dietary exposure was expressed in µg/kg bw/d by dividing the exposure to the individual's body weight.

10.4 Mixture risk assessment

The cumulative risk assessment of the metals was performed following the hazard index (HI) approach. Here, the cumulative exposure in the two scenarios (see section exposure assessment), was divided by the TRV of an index compound, in this case cadmium. It was decided to present the individual and cumulative HI of the metals for the Dutch population as an example because the Dutch consumption data is the most recent.

10.4.1 Equipotency

Assuming the metals are equipotent, the median cumulative HI in the Dutch population is 4.2 and the P95 12.7 (Figure 10.1). The 'European' median HI varies between 1.2 and 14.1 with a median of 4.1 for all the countries included in this exposure assessment. The risk driver in the median cumulative HIs is lead, followed by cadmium, inorganic arsenic and mercury.

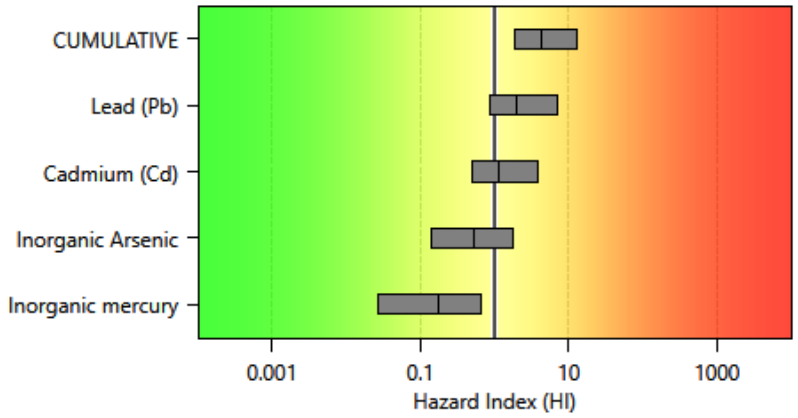


Figure 10.1: Hazard index (HI) of the individual metals as well as the cumulative HI considering equipotency. Bars show variability of the HI (range P5-P95) in the Dutch population.[1]

Using the probabilistic approach to combine the exposure of the metals also shows that simply adding the individual P95 dietary exposure distributions of the separate metals would overestimate the cumulative exposure in the Dutch population (Table 10.2). This was observed for the other HI assessments as well.

Table 10.2: Hazard index (HI) of the cumulative exposure to the metals and the HIs of the individual exposures in the Dutch population in the equipotent scenario

HI	P50	P95
Cumulative	4.2	12.7
lead	2.04	6.96
cadmium	1.15	3.93
arsenic	0.54	1.8
mercury	0.18	0.65

[1] A HI slightly above 1 may be related to a higher risk and refinement of the assessment is required, hence the yellow/orange color. A clear risk based on the HI >10 is indicated by the red color, no risk based on the HI < 0.1 is indicated by the green color.

10.4.2 Non-equipotency

Assuming the scenario that the metals are not equipotent, the cumulative HI in the Dutch population is 5.5 and the P95 is 18.2 (Figure 10.2). The 'European' median HI varies between 1.5 and 18.3 with a median of 5.3 for all the countries included in this exposure assessment. The risk driver in the median cumulative HIs is lead, followed by cadmium, inorganic arsenic and mercury. This is not surprising considering that lead was the risk driver in the equipotent scenario and the potency of lead in this scenario was considered higher than that of the other metals. The median HI for lead increased from 2.04 in the equipotent scenario to 3.67 in the non-equipotent scenario. In contrast, the median HI for mercury decreased from 0.18 to 0.11.

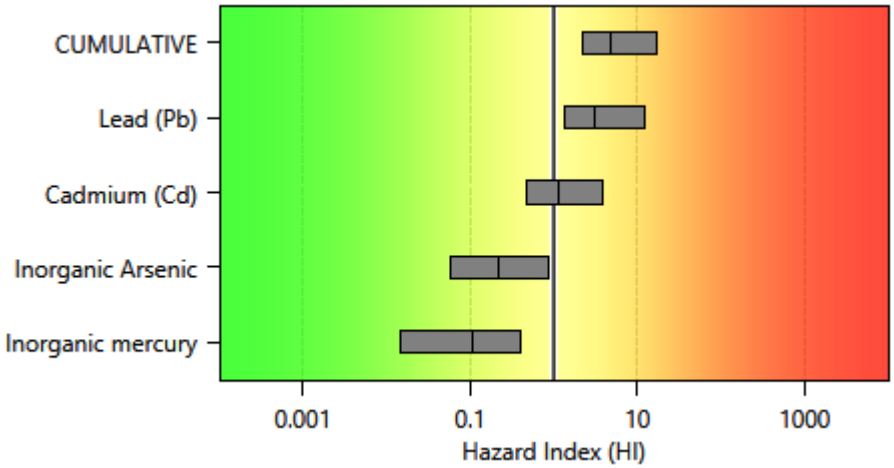


Figure 10.2: Hazard index (HI) of the individual metals as well as the cumulative HI considering non-equipotency. Bars show variability of the HI (range P5-P95) in the Dutch population[2]

[2] See footnote 1 for explanation

10.5 Uncertainty considerations

The HI approach in general may overestimate the cumulative exposure as different uncertainty factors have been applied to every POD to derive a TRV. A point of departure approach where one general uncertainty assessment factor is applied, might result in a more refined risk assessment. However, this may not be feasible in every scenario, considering that the PODs might be based on studies that vary with respect to species, time and dose duration.

Related to the point addressed above are the uncertain RPFs of the metals. Ideally, the potency of the metals is to be assessed simultaneously. This work is currently ongoing in several in vitro experiments specifically related to the AOP of these metals and the endpoint of nephrotoxicity. This work will be updated when RPFs for these metals have been derived. Meanwhile, the different PODs of the metals used for calculating the PODI in case of occupational exposure vary a factor of 3 (from 5 to 15 µg/L) and the RPFs in this study also vary a factor of 3 (from 0.57 to 1.76).

In these dietary exposure calculations, we replaced all samples that were analyzed below the LOQ and LOD with zero. When the LOQs and LODs would have been more sensitive, these samples might have been quantitatively analyzed and as such not replaced by zero.

For a few countries, a very large exposure distribution for arsenic was observed, resulting in a relatively high P95 individual exposure of arsenic and consequently a high P95 in the cumulative exposure. This large distribution may be explained by limited available occurrence data on inorganic arsenic. When a certain compound is consumed relatively often, but only few concentration data are available, one outlier in the data can already greatly affect the highest exposures. This was for example also the case for arsenic where in one of the countries, the food driving the exposure was ‘meat-based meals’. For this food category, only nine samples were analyzed for arsenic, with a minimum reported concentration of 0.04 mg/kg and a maximum concentration of 11 mg/kg. This one maximum value greatly increased the average concentration of inorganic arsenic in this product category. In addition, the differences in the ages and consumption patterns of the populations may explain the large differences in the exposure between all countries. These were not accounted for in these exposure calculations.

In addition, also ‘leafy vegetables’ were initially reported to contribute very highly to the dietary exposure of arsenic. We were able to check the reported occurrence data of ‘leaf vegetables’ from the Netherlands and found that these were originally coded as samphire (“zeekraal”). These aberrantly coded samples from the Netherlands were therefore corrected to ‘sea weeds’. We were however not able to check for possible other aberrantly coded samples from other countries, yet no samples in the database indicated the need for this.

We used the fractions as reported by EFSA to assign the inorganic arsenic or mercury content of the samples in which total arsenic or mercury was analyzed. There are however uncertainties related to these fractions. EFSA for example reported significant uncertainty in predicting the inorganic arsenic fraction from total arsenic levels in fish and seafood, because the relative proportion of inorganic arsenic tends to decrease as the total arsenic content increases. This also appears to vary depending on the type of fish and seafood (EFSA, 2009b; EFSA, 2021). Estimating the occurrence of inorganic arsenic or mercury based on fractions of total arsenic or mercury introduces an uncertainty to the occurrence data of arsenic and mercury in which the total arsenic or mercury content was analyzed rather than the inorganic arsenic content.

10.6 Lessons learned

Hazard

The toxicological references values on which the relative potencies of the metals were based, are based on nephrotoxic effects observed in different species. Preferably, these values are all derived based on human data, or within a similar experiment.

Exposure

To facilitate an accurate dietary exposure assessment, there is a great need for metal speciation in the analysis. Total mercury and total arsenic concentrations are often analyzed, yet the fraction of the inorganic metal varies greatly between products.

Risk

The median HI varies greatly between European countries. These are likely based on differences related to differences in consumption which may also be related to differences in the age groups of the populations. Using the probabilistic approach to combine the exposure of the metals shows that simply adding the individual dietary exposure distributions of the separate metals would overestimate the cumulative exposure in the highest exposure percentiles.

10.7 References

Crépet, A, Vanacker M, Sprong C, de Boer W, Blaznik U, Kennedy M, Anagnostopoulos C, Christodoulou DL, Ruprich J, Rehurkova I, Domingo JL, Hamborg Jensen B, Metruccio F, Moretto A, Jacxsens L, Spanoghe P, Senaeve D, van der Voet H, 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, pp 291-306 https://doi.org/10.1016/j.ijheh.2018.12.002

EFSA 2009a Scientific Opinion of the Panel on Contaminants in the Food Chain on a request from the European Commission on cadmium in food. The EFSA Journal (2009) 980, 1-139.

EFSA Panel on Contaminants in the Food Chain (CONTAM) 2009b Scientific Opinion on Arsenic in Food. EFSA Journal 2009; 7 (10) 1351. doi:10.2903/j.efsa.2009.1351

EFSA Panel on Contaminants in the Food Chain (CONTAM) 2010 Scientific Opinion on Lead in Food. EFSA Journal 2010; 8 (4) 1570. doi:10.2903/j.efsa.2010.1570

EFSA Panel on Contaminants in the Food Chain (CONTAM) 2012 Scientific Opinion on the risk for public health related to the presence of mercury and methylmercury in food. EFSA Journal (2012) 10 (12) 2985. doi:10.2903/j.efsa.2012.2985

EFSA Panel on Contaminants in the Food Chain (CONTAM) 2021 Scientific report on the chronic dietary exposure to inorganic arsenic. EFSA Journal 2021;19 (1) 6380. https://doi.org/10.2903/j.efsa.2021.6380

JECFA 2011 Seventy-second report of the Joint FAO/WHO Expert Committee on Food Additives – Evaluation of certain contaminants in food. WHO Technical report series no 959

Schillemans T, Åkesson A, Mengelers M, Luijten M. 2021 A common Adverse Outcome Pathway for toxic metals inducing nephrotoxicity to aid human health risk assessment of chemical mixtures. *submitted*

11 Case study 6: Risk assessment of the occupational exposure to a mixture of four toxic metals

RIVM: Marcel Mengelers, Annick van den Brand, Mirjam Luijten, FIOH: Juha Tuovila, Tiina Santonen

11.1 Grouping criteria

11.1.1 Effect-based grouping

In the HBM4EU project four toxic metals have been identified as priority metals for the exposure of the general population as well as workers: arsenic (As), cadmium (Cd), mercury (Hg) and lead (Pb). In humans, some of these metals have a neurodevelopmental effect but all of them exert a nephrotoxic effect. This nephrotoxic effect can be specified to an adverse effect on the glomerulus and the proximal tubule of the kidney. To gain more insight in the mode of action of this nephrotoxic effect, an adverse outcome pathway has been developed for these metals in this work package (Schillemans et al., 2021). In short, the grouping of these metals has been effect based.

11.1.2 Selection of population

Data was collected from the LIMs register of the Finnish Institute of Occupational Health between 1.1.2010 and 31.12.2019. First, we identified sectors and workplaces with exposure to the four metals of interest. Finally, the data selection was focused on employees who were working in the waste management sector. The dataset includes only adult workers, 73 males and 1 female, aged between 20-65 years. It comprises 873 results on blood cadmium levels, blood lead levels, urinary cadmium, urinary inorganic arsenic (As^{3+} and As^{5+}), blood inorganic mercury and urinary total mercury. The dataset also included information on exposure length (years), smoking and analysis date.

11.2 Hazard characterization

11.2.1 Hazard identification

The four toxic metals investigated in this case study (As, Cd, Hg and Pb) are usually referred to as heavy metals although the classification ‘heavy’ is disputed. Heavy metals are generally defined as metals with relatively high densities, atomic weights, or atomic numbers. The criteria used, and whether metalloids are included, vary depending on the author and context.

Arsenic is a metalloid. Arsenic in the environment and food occurs in both organic and inorganic compounds in its trivalent or pentavalent state (As^{3+} or As^{5+}). The nephrotoxic effect is related to inorganic As (iAs). Absorbed arsenic, irrespective of its form, is widely distributed in the human body. Inorganic arsenic can be metabolized in humans to organic compounds like monomethyl and dimethyl arsenide (MMA and DMA). The main route of excretion of all forms is through the kidneys. Absorbed arsenic is excreted in the urine within a few days. The elimination half-life of As in humans is approximately 40 to 60 hours.

Cadmium is usually present in food and in the human body as different salts of bivalent Cd (Cd^{2+}). The elimination profile of Cd in human blood shows a biphasic decrease. The distribution half-life of Cd in blood is about 100 days and the elimination half-life varies from 7-16 years. This slow elimination is reflected in the retention times in various tissues. The longest retentions have been reported in muscles, kidney cortex and liver and may vary from 10 to 40 years.

Mercury is present in the environment and food as inorganic or organic compound. The nephrotoxic effect is related to inorganic Hg (iHg). After absorption, Hg is present in the human

body in bivalent (Hg^{2+}) forms. The elimination half-life of Hg in humans varies from a few days to months. The organs with the longest retention times are the brain, liver, kidneys and testicles.
Lead is present in the environment, food and the human body as different salts of bivalent Pb (Pb^{2+}). The elimination half-life of Pb in humans varies from one month (blood and soft tissues) to years (bones).
Daily exposure of Cd, Hg and Pb during working days in combination with their slow elimination ensures that these metals will accumulate in the body and that at steady-state the time of sampling (blood or urine) is not critical. Although the accumulation of As is probably less, prolonged exposure will probably also lead to little fluctuation in the blood or urine level during the (working) day.

11.2.2 Biomarkers of exposure

Exposure to arsenic is typically measured in urine. Measurement of total arsenic in urine is an unspecific method, which is very much affected by the dietary exposure of organic arsenic compounds. Since occupational exposure is to inorganic arsenic, it is important to perform speciation of urinary arsenic (U-As) and measure separately inorganic arsenic (As^{3+}, As^{5+}), and arsenic species, like DMA and MMA, which are the metabolites of inorganic arsenic but for which it is also possible to be exposed directly via the food (ECHA, 2017). FIOH analyses and reports separately inorganic arsenic (As^{3+}, As^{5+}). These results are used in this report. LOD and LOQ used currently at FIOH for inorganic arsenic are 0.15 and 0.63 µg/L.

Cadmium can be measured either in urine or in blood. Blood cadmium (B-Cd) levels are considered to reflect primarily the recent exposure to cadmium whereas urinary cadmium (U-Cd) gives information on the cumulative kidney burden of cadmium. U-Cd levels have been linked directly with the nephrotoxicity of cadmium whereas correlation between B-Cd and health effects is indirectly based on the correlation between U-Cd and B-Cd (SCOEL, 2017). Since FIOH data on waste workers included mostly B-Cd measurements, it was decided to use B-Cd for the combined risk assessment instead of U-Cd. LOD and LOQ for B-Cd is 0.019 and 0.028 µg/L.

Blood lead (B-Pb) measurement is commonly used as a biomarker for a cumulative exposure to lead. It has been associated directly with the adverse health effects of lead. It is possible to measure lead in urine. Urinary lead (U-Pb) has been mainly used in the past for the assessment of exposure to organic lead compounds (tetraethyl- and tetramethyl lead) but it has also been suggested to indicate the recent exposure to inorganic lead. Since lead health effects are mainly related to the cumulative body burden of lead, B-Pb is the best biomarker to assess health risk arising from lead exposure (ECHA, 2019). Correlation between urinary lead levels and health effects have not been established. Using modern analytical methods, detectable levels of B-Pb can be measured in the general population without additional occupational exposure. LOD and LOQ for B-Pb at FIOH is 0.42 and 0.63 µg/L, respectively.

Also mercury can be measured in blood and urine. Measurement of total mercury in blood (B-Hg) indicates exposure to both inorganic and organic mercury species. Since in occupational health, exposure is exclusively related to inorganic mercury (i-Hg) and intake of organic mercury via food can significantly impact the B-Hg levels, it is important to perform speciation of mercury if it is measured in blood. FIOH routinely measures inorganic and organic mercury separately in blood, LOQ for inorganic blood mercury is currently 0.97 µg/L and LOD is 0.67 µg/L. Since inorganic mercury is excreted in urine more readily than organic mercury, urinary mercury is commonly used as a biomarker for inorganic mercury (SCOEL, 2007). The difference to B-Hg-i measurement is that there is more variation in U-Hg levels during the day and the sampling time is more critical

whereas in the case of B-Hg-i the sample can be taken at any time of the day. In this study, we have used B-Hg-i data.

11.2.3 Toxicological reference values

Several toxicological reference values (TRVs) or Point of Departures (PODs) have been determined in blood and/or urine by various international organizations. For practical reasons, as discussed in section 3.1, we have restricted our mixture risk assessment to the following matrix-metal combinations: arsenic in urine (U-As), cadmium in blood (B-Cd), lead in blood (B-Pb) and mercury in blood (B-Hg). An overview of the various available TRVs or PODs for these matrix-metal combinations is shown in Table 11.1.

Table 11.1: Critical effects and TRVs / PODs of four toxic metals in blood or urine

Compound	Matrix	Critical effect	Type of TRV or POD	Value	Population	Reference
Arsenic (As)	urine	tubular damage	BMDL	102 µg/g creatinine [1]	GenPop	Hong et al. (2004)
	urine	tubular damage	-	23 µg/L [2]	Workers	Foa et al. (1987)
	urine	tubular damage	-	7 µg/L [3]	Workers	-
Cadmium (Cd)	blood	renal dysfunction	HBM-GV	5 µg/L	Workers	Lamkarkach et al. (2021)
	blood	renal dysfunction	BEI	5 µg/L	Workers	WorkSafe (2020)
Lead (Pb)	blood	decreased GFR	BMDL	15 µg/L	GenPop	EFSA (2010)
Inorganic Mercury (i-Hg)	blood	nephrotoxicity	BLV	10 µg/L	Workers	SCOEL (2007)

BEI = Biological Exposure Index ; BLV = Biological Limit Value; BMDL = lower limit of Bench Mark Dose; HBM-GV = Human Biomonitoring-Guidance Value; GenPop = General population
[1] Value for total (inorganic + organic) arsenic
[2] Value for inorganic arsenic including MMA and DMA
[3] Value for inorganic arsenic

In case of As, no TRV in blood was found. We had to derive our own POD in urine because two different TRVs were found depending on the type of inorganic and/or organic As involved. In the study of Hong et al. (2004) urine samples of residents exposed to arsenic by contaminated air and food were analyzed. Hong et al. suggest that, due to low fish consumption in the contaminated area, exposure to organic arsenic was low but due to the applied analytical method (atomic absorption spectroscopy) this cannot be excluded. If in this study the investigated population (n=122) was not or hardly exposed to organic arsenic, the BMDL of 102 µg/g creatinine could be based on the sum of inorganic As and the metabolites MMA and DMA. It has been suggested by Hayes et al. (2010) and Nordberg et al. (2015) that urinary excretion of arsenic consists of approximately 20% inorganic arsenic, 25% MMA and 55% DMA. In that case, a BMDL for iAs including MMA and DMA, could refer to a BMDL of approximately 20 µg iAs/g creatinine and thereby support the finding of Foa et al. (1987). The paper of Foa et al. describes the exposure of industry workers (n=17) to As_2O_3. The results indicated a significant increase in urinary As concentration in exposed workers when compared with a reference population. After comparing the exposed subjects with a matched reference group a borderline significant increase of retinol binding protein (biomarker of tubular damage) was observed. The urinary inorganic As (iAs) levels

(As^{3+} plus As^{5+}) were 23 ± 13 µg/L. It is debatable if this urinary level should be treated as a NOAEL or a LOAEL. Considering the large coefficient of variation (56.5 %) an additional uncertainty factor of 3.2 (= √10, for intraspecies variability in toxicodynamics) is justified to estimate a POD for urinary iAs of approximately 7 µg/L.

For cadmium, within the HBM4EU project, an HBM-GV for workers of 5 µg/liter blood has been proposed by Lamkarkach et al. (2021). This value was based on a Biological Exposure Index (BEI) of 5 µg/L recommended by the American Committee of Governmental Industrial Hygienists (ACGIH) in 2016. This committee based its recommendation on data reporting a dose-effect relationship between B-Cd and the prevalence of tubular proteinuria in 440 workers.

The TRV of B-Pb was the only TRV that was based on a BMD analysis of human data (EFSA, 2010). Nephrotoxicity was analyzed for the prevalence of chronic kidney disease (CKD) based on a reduction in the glomerular filtration rate (GFR) to values below 60 mL/min. For the prevalence of CKD, a BMR of 10 % was chosen resulting in the calculation of a BMDL10 of 15 µg/L.

The TRV of inorganic mercury in blood (10 µg/L) was derived by SCOEL in 2007 and was based on a study of Roels et al. (1985). This B-Hg concentration was treated as a NOAEL for renal toxicity which was indicated by elevated tubular effect biomarker levels in the urine.

11.3 Exposure assessment

11.3.1 Internal exposure

A summary of the occupational exposure levels to the four metals for E-waste workers is presented in Table 11.2. The highest exposure levels were found for lead in blood, followed by cadmium in blood, i-As in urine and i-Hg in blood.

Table 11.2: Exposure of E-waste workers to four selected metals

	U-As	B-Cd	B-Hg	B-Pb
N > LOD	141	317	51	296
N > LOQ	74	317	3	296
P05 (µg/L)	0.31 (=LOQ/2)	0.10	0.48 (=LOQ/2)	6.22
P50 (µg/L)	0.64	0.53	0.48 (=LOQ/2)	12.8
P95 (µg/L)	2.12	4.54	0.48 (=LOQ/2)	43.4

Data from FIOH's earlier study on the exposure of occupationally non-exposed working aged population (non-smokers) to B-Cd and B-Pb is used for comparison. This data is based on the FIOH sample collection during 2010-2011 and includes office workers from Helsinki, Tampere and Kuopio regions. For U-As and B-Hg this data was not available, but for U-As preliminary data from the respective study from year 2019-2020 sample collections is available. In this dataset both P50 and P05 for U-As-I were below the LOQ. It should be emphasized that the numbers are based on preliminary data only. These exposure levels are considered to reflect background exposure of the Finnish population from environmental and dietary sources.

Table 11.3: Toxic metal levels of non-occupationally exposed Finnish population

	U-As (n=151)	B-Cd (n=118) #	B-Hg-i	B-Pb (n=118)
N > LOD	140	118	n.a.	118
N > LOQ	24	118	n.a.	118
P05	0.31 (=LOQ/2)	0.29	n.a.	5.56
P50	0.31 (=LOQ/2)	0.16	n.a.	10.31
P95	1	0.54	n.a.	26.45

n. a. - not available, # data includes only non-smokers, for a small number of smokers (n=18), levels were 1.2 (P95), 0.8 (P50), 0.23 (P05) µg/L.

11.3.2 Exposure scenarios

In case of B-Hg, more than 90 % of the samples are below LOQ but above LOD. We used three scenarios to deal with these 'missing' values. An *upper* bound scenario: all samples below LOQ will receive a concentration equal to LOQ. A *medium* bound scenario: all samples below LOQ will be assigned to ½ LOQ. A *lower* bound scenario: all samples below LOQ will be set at LOD. No samples below LOQ will be set at zero.

It is proposed to follow a tiered approach in the estimation of the PODI. The lowest tier could be a worst-case scenario where the ratio is determined of the upper bound concentrations. If the overall PODI is below 1, using an overall UF, then the combined risk is low or negligible. However, if the overall PODI is below 1 using no UF, but above 1 with UF, then the overall UF needs further investigation in tier 2. If the overall PODI is above 1, using no UF, then tier 3 is entered. In this tier the effect of the medium bound and lower bound concentrations on the PODI needs to be investigated. In short:

- Tier 1: PODI x UF < 1
- Tier 2: PODI < 1 and PODI x UF > 1
- Tier 3: PODI > 1

The contributions of the different metals to the combined PODI will also be assessed.

11.4 Mixture risk assessment

11.4.1 Approach

In the Finnish database, the individual internal concentrations of the heavy metals are available and it was decided to determine the distribution of the individual PODIs (based on dose addition). The individual PODI will be calculated as follows:

$$individual\ PODI = \frac{iEL_{As}}{POD_{As}} + \frac{iEL_{Cd}}{POD_{Cd}} + \frac{iEL_{Hg}}{POD_{Hg}} + \frac{iEL_{Pb}}{POD_{Pb}}$$

iEL_{As} / POD_{As} = *individual Exposure Level of (inorganic) Arsenic / POD of (inorganic) Arsenic*

Subsequently, for a group of 51 individuals the distribution of all iPODIs was determined for three different exposure scenarios, as mentioned in section 3.2 (upper, medium and lower bound). In this way, the calculated PODI will not be overestimated by avoiding the addition of a high exposure (P95) level of each metal separately. In addition, the contribution of each toxic metal to the PODI was by assessing the fraction of each metal to the individual PODI.

Due to time constraints, we also calculated the risk quotients (EL/POD) for a group of people (n=118-151) that have been non-occupationally exposed to the metals. This group serves as a population with 'background' exposure to these metals. In the near future we aim to calculate the iPODI of these non-occupationally exposed people.

11.4.2 PODIs related to exposure scenarios

In Table 11.4, the median (P50) and the 95th percentile (P95) of the iPODI for the mixture of U-As, B-Cd, B-Hg and B-Pb are given for the three different exposure scenarios. In the upper, medium and lower bound scenario all samples below LOQ were assigned a concentration equal to LOQ, ½ LOQ or LOD.

Table 11.4: The median (P50) and the 95th percentile (P95) of the iPODI for the mixture of U-As, B-Cd, B-Hg and B-Pb for the three lower, medium and higher bound exposure scenario

Percentile	Exposure scenarios		
	Lower bound	Medium bound	Upper bound
P50	1.18	1.18	1.19
P95	3.07	3.08	3.09

Additionally, the average contribution of the individual metals to the iPODIs was shown in Figure 11.1. It is clear that blood lead levels are the mainly driving the iPODIs, followed by urinary arsenic levels, blood cadmium levels and to a lesser extent by blood mercury levels.

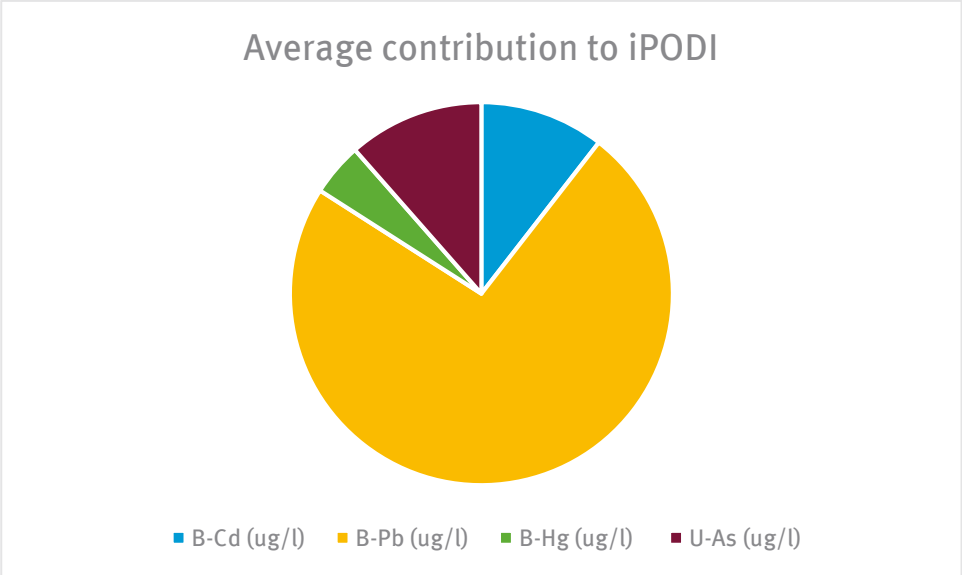


Figure 11.1: The average contribution of each metal to the iPODI of all individuals

In Table 11.5, the P50 and P95 of the risk quotients (EL/POD) are given for a group of people (n=118-151) non-occupationally exposed to the arsenic, cadmium and lead. Note that this does not include the data on B-Hg.

Table 11.5: The (absolute and relative) risk quotients (EL/POD) for a group of people (n=118-151) non-occupationally exposed to the four toxic metals

Percentile	EL / POD							
	U-iAs		B-Cd		B-iHg		B-Pb	
	abs	%	abs	%	abs	%	abs	%
P50	0.04	5 %	0.032	4 %	n.a	n.a	0.69	91 %
P95	0.14	7 %	0.11	5 %	n.a	n.a	1.76	88 %

11.4.3 Conclusions

The median as well as the P95 of the PODI is above 1 for all exposure scenarios (Table 11.2). Therefore, there is no need to use the tiered approach as proposed in section 3.2. Due to the fact the TRV for each matrix-metal combination is a deterministic value (in contrast to the exposure), an overall uncertainty factor could be used to express e.g. a common uncertainty related to an interindividual variation in nephrotoxicity.

As can be seen in Figure 11.1, B-Pb has the largest contribution to the PODI, followed by B-Cd, U-As and B-Hg.

The contribution of the different matrix-metal combinations for the non-occupationally exposed people is, in descending order, B-Pb, U-iAs and B-Cd (Table 11.3). The data on B-iHg is not available yet but it is very likely to have only a minor contribution to the overall risk like in the case of exposed E-waste workers.

11.5 Uncertainty considerations

11.5.1 Hazard

Within a mixture risk assessment, it is assumed that the adverse effect of the investigated substances is additive. Within this WP we have already developed an adverse outcome pathway for As, Cd, Hg and Pb and it was shown that the adverse outcome for these metals at organism level was proximal tubular damage (Schillemans et al., 2021). All applied TRVs were based on nephrotoxicity as critical effect in humans, and more specifically, on measurements of proximal tubule and/or glomerular effect biomarkers.

The TRVs of the four metals are based on measurements in populations of different sizes. In case of Pb the data were taken from the general population: approximately 15,000 adults aged at least 20 years were sampled from the NHANES (1999-2006) study. For As, Cd and Hg the populations consisted of adult workers and group sizes were respectively 17, 440 and 185. Clearly, the TRV of As is based on a very small population size but this TRV could be supported by the study of Hong et al. (2004), as has been described in section 2.3, that consisted of a study population of 122 people.

Apart from Pb, the TRVs for the other metals are based on a NOAEL or LOAEL and in the near future we would like to quantify the uncertainties related to these TRVs by means of a tool called APROBA-Plus (Bokkers et al., 2017). This tool can be used to convert a deterministic TRV to a probabilistic TRV, while addressing the contribution of the uncertainty factors. Subsequently, we will decide to what extent an overall uncertainty factor is needed for this PODI.

11.5.2 Exposure

The population selected for this case study represents workers in waste management sector who have been monitored for these four metals as part of their health surveillance. The data may not be representative of the waste workers exposure to these metals in general since it depends on how well occupational health care has recognized a need for monitoring these metals in this group of workers. It may include workers involved in many different tasks with variable exposure. No specific information on the specific tasks performed by the workers is available. The data involves both smokers and non-smokers. Differentiation between smokers and non-smokers is possible but has not been done yet. Smoking *per se* increases exposure to e.g. cadmium but at workplaces one significant route of exposure is also through the contact of contaminated hands to mouth due to the smoking habit. Regarding exposure levels sufficient quantifiable concentrations for B-Cd and B-Pb were available but only half of the concentrations of U-As were above LOQ and just a few for B-Hg. The data from control population represents non-smokers.

11.6 Lessons learned

11.6.1 General

General lessons learned while working on this case study are:

- Analogous to the risk assessment of a single substance, mixture risk assessment benefits from detailed problem formulation. Additionally, inclusion and/or exclusion criteria for substances (similarity in specified effect), (test) methods (in animals or humans, *in vitro* or *in vivo*) and exposure (matrix type) and hazard data (single values or confidence intervals) are to be included.
- Use of methodologies for mixture risk assessment may involve the use of toxicological reference values (TRVs) such as the TDI. Guidance on the selection of appropriate TRVs, for example when a TRV is a NOAEL/LOAEL or when a TRV is not available for a compound in the assessment group, will greatly facilitate implementation of mixture risk assessment.
- Identification of important risk drivers point out which hazard should receive attention in terms of risk mitigation.

11.6.2 Case specific

Lessons learned that are specific for this case study are:

- Use of individual exposure data for mixture risk assessment provides opportunities for refinement of the assessment, which prevents an overestimation of the risk specifically in the higher exposure percentiles.
- The use of HBM data for mixture risk assessment involved data obtained from different matrices (blood and/or urine). Converting or combining data from one matrix into another introduces another source of uncertainty to the assessment, but is unavoidable when no appropriate TRVs are available for that matrix.
- The current derivation of the PODI is semi-probabilistic: a distribution of the exposure levels is combined with a deterministic value for the POD. It would be worthwhile to investigate if the POD could be expressed in terms of a confidence interval (like a BMD output) and thereby quantify the uncertainties in the PODI.

11.7 References

ACGIH (2016). American Conference of Governmental Industrial Hygienists. Cadmium and Inorganic Compounds. Chemical Substances. Documentation of the Threshold Limit Values and Biological Exposure Indices, 7th Edition.

Bokkers B, Mengelers MJ, Bakker MI, Chiu WA, Slob W (2017). APROBA-Plus: A probabilistic tool to evaluate and express uncertainty in hazard characterization and exposure assessment of substances. Food and Chemical Toxicology 110: 408–417.

ECHA (2017). Committee for Risk Assessment (RAC). Opinion on arsenic acid and its inorganic salts.

ECHA (2019). ECHA Scientific report for evaluation of limit values for lead and its compounds at the Workplace.

EFSA (2010). Panel on Contaminants in the Food Chain (CONTAM). Scientific Opinion on Lead in Food. EFSA Journal 2010; 8 (4) 1570. doi:10.2903/j.efsa.2010.1570

Foà V, Colombia A, Maroni M, Barbieri F, Franchini I, Mutti A, De Rosa E, Bartolucci GB (1987). Study of kidney function of workers with chronic low level exposure to inorganic arsenic. In: Monitoring Human Toxicity, Foà V, Emett EA, Maroni M, Colombi A (eds), Ellis Horwood, Chichester, 362-367.

Hays SM, Aylward LL, Gagné M, Nong A, Krishnan K (2010). Biomonitoring Equivalents for inorganic arsenic. Regulatory Toxicology and Pharmacology 58: 1–9.

Hong F, Jin T, Zhang A (2004). Risk assessment on renal dysfunction caused by co-exposure to arsenic and cadmium using benchmark dose calculation in a Chinese population. BioMetals 17: 573–580.

Lamkarkach F, Ougier E, Garnier R, Viau C, Kolossa-Gehring M, Lange R, Apel P (2021). Human Biomonitoring initiative (HBM4EU): Human Biomonitoring guidance values (HBM-GVs) derived for cadmium and its compounds. Environment International 147; 106337. https://doi.org/10.1016/j.envint.2020.106337

Nordberg GF, Fowler BA, Nordberg M (2015). Handbook on the Toxicology of Metals: Fourth Edition. Volume 1: General considerations. & Volume 2: Specific metals.

Roels H, Gennart JP , Lauwerys R, Buchet JP, Malchaire J and Bernard A (1985). Surveillance of workers exposed to mercury vapour: Validation of a previously proposed biological threshold limit value for mercury concentration in urine. Am J Ind Med 7, 45-71.

SCOEL (2007). Recommendation from the Scientific Committee on Occupational Exposure Limits for elemental mercury and inorganic divalent mercury compounds.

SCOEL (2017). Opinion from the Scientific Committee on Occupational Exposure Limits. Cadmium and its inorganic compounds.

Schillemans T, Åkesson A, Mengelers M, Luijten M (2021). A common Adverse Outcome Pathway for toxic metals inducing nephrotoxicity to aid human health risk assessment of chemical mixtures. Submitted

Worksafe (2020). New Zealand Government. Biological Exposure Index (BEI) review. Cadmium and cadmium compounds.

12 Case study 7: Relying on repeated biospecimens to reduce the effects of classical-type exposure measurement error in studies linking the exposome to health

Inserm: Lydiane Agier, Rémy Slama, ISGlobal Barcelona: Xavier Basagaña

12.1 Introduction

The exposome concept was defined as encompassing the totality of environmental exposures from the prenatal period onwards (Wild, 2012). Exposome studies typically rely on biomarkers to assess chemical exposure from multiple sources (Calafat et al., 2006; Schisterman and Albert, 2012); often, a single biospecimen is collected in each subject (LaKind et al., 2019; Perrier et al., 2016). For chemicals with a short biological half-life, such as bisphenol A or phthalates, within-subject biomarker concentrations strongly vary over time. Whatever the accuracy of the measurement technique used, this will induce measurement error, mainly for the least persistent compounds (those with the lowest intraclass coefficient of correlation, or ICC), for which a spot biospecimen will provide a poor estimate of exposure over long time windows (Casas et al., 2018; Perrier et al., 2016; Preau et al., 2010; Vernet et al., 2019; Wielgomas, 2013; Ye et al., 2011).

A realistic assumption is that chemicals exposure is measured by biomarkers with an independent additive error, such that the within-subject average of repeated measurements is an unbiased estimate of exposure. This type of measurement error corresponds to classical- type measurement error (Carroll et al., 2006). In a single exposure context, classical-type exposure measurement error can strongly impact the estimation of exposure-health relationships (Armstrong, 1998; Brakenhoff et al., 2018a; Carroll et al., 2006; de Klerk et al., 1989; Jurek et al., 2005; Perrier et al., 2016): naïve models not accounting for measurement error provide regression estimates that are attenuated, and have decreased statistical power compared to the ideal situation without measurement error (Carroll et al., 2006). Several statistical techniques have been developed to limit estimation bias when classical type exposure measurement error is present. Regression calibration (RC) (Carroll et al., 2006) is often used, and provides approximately consistent estimates (Buzas et al., 2005). It requires either a (within-subject) repeated assessment of exposures, at least in a subset of the population, or an unbiased estimate of exposures' ICC.

An alternative, also relying on the collection of repeated biospecimens, is to pool samples within-subject and assess the exposures biomarker in the pooled sample, corresponding to the so-called within-subject biospecimens pooling approach (Perrier et al., 2016; Vernet et al., 2019). In a single exposure context, within-subject biospecimens pooling, possibly followed by a posteriori disattenuation of coefficients estimates, is efficient to limit attenuation bias in the exposure-health association and to increase power (Perrier et al., 2016; Vernet et al., 2019).

The impact of classical-type exposure measurement error on bias and power_may be compounded in an exposome context. Indeed, in exposome epidemiological studies, efficiently identifying the exposures affecting health (i.e. with good sensitivity, or power; and low false detection proportion, or FDP) is generally a challenge, in particular when exposures show some correlation, which is a realistic assumption (Agier et al., 2016). Additionally, exposome studies can simultaneously consider several biomarker-based exposures with differential measurement error (i.e. of different amplitudes across exposures) (Slama and Vrijheid, 2015); typically, persistent pollutants such as PCBs, for which a spot biospecimen provides a rather good exposure proxy, and non persistent chemicals, such as organophosphate pesticides or bisphenol A, with ICCs based on spot biospecimens in the 0.1-0.5 range (Casas et al., 2018), can both be included in the

same study (Agier et al., 2020, 2019). This differential measurement error in exposome studies is likely to further affect the efficiency to identify exposures affecting health, beyond the performance loss due to the consideration of a large number of correlated compounds. Indeed, when applying a variable selection model in an exposome study relying on spot biospecimens, the chances of a given exposure being selected might be lower for compounds with a low ICC. In addition, in a multivariable situation, regression estimates can be biased in any direction as a consequence of classical-type measurement error (Brakenhoff et al., 2018b; Carroll et al., 2006).

Up to date, three simulation studies have estimated the performances of specific variable selection methods in assessing exposure-outcome associations in a multi-exposure context with classical-type exposure measurement error. Guangning et al. tested several two-step approaches (applying a measurement error correction method and using the resulting exposure proxies in a variable selection method) in a 2-exposures context (Guangning, 2014). Vasquez et al. applied a corrected least absolute shrinkage and selection operator (LASSO) procedure to 100 uncorrelated exposures (Vasquez et al., 2019). Brown et al. compared LASSO to a novel variable selection method in a context where 10 exposures affected the outcome, were not correlated to the other 90 exposures, and jointly explained over 90 % of the outcome variability (Brown et al., 2019), which is much higher than the outcome variability that a few environmental exposures are expected to explain.

To our knowledge, no study assessed the impact of exposure measurement error on exposure-outcome relationship in a realistic exposome context, i.e. in a context with many exposures that are all correlated (Tamayo-Uria et al., 2019), with a small global explanatory power of exposures on the outcome. Moreover, the efficiency of the recently proposed within-subject biospecimens pooling approach (Perrier et al., 2016) has never been estimated in a multi-exposure context.

12.2 Methods

12.2.1 Overview of the simulation model

We simulated an exposome study with p=1, 10 or 237 error-prone exposures in a fictitious population of 1 200 subjects (the approximate population size of the HELIX exposome project, which we mainly refer to for the design of this simulation study (Agier et al., 2020; Vrijheid et al., 2014)). We assumed that k=0, 1, 3 or 10 exposures linearly influenced a health outcome $\mathbf{Y}$, which is an assumption of rare features realistic for ‘omics data, including the exposome (Donoho and Jin, 2008). We considered that all exposures suffered classical-type measurement error, with varying levels of error which were quantified from their intra-class correlation coefficient (ICC). Exposure-health associations were assessed using 2-step approaches. We first applied either RC or within-subject biospecimens pooling, assuming that 2, 5 or 10 error-contaminated biospecimens (collecting 10 biospecimens per subject is now logistically feasible (Lyon-Caen et al., 2019)) were available per subject (we did not cover a greater range of values for the number of biospecimens per subject because our aim is to illustrate that repeated biospecimens collection is a solution, without necessarily trying to identify an optimum number of biospecimens per subject).

With both these methods, we obtained a set of transformed exposures over which, in the second step, we applied a variable selection method (Deletion/Substitution/Addition algorithm, DSA, (Sinisi and van der Laan, 2004)) to identify exposures associated with $\mathbf{Y}$. The performances of the RC and pooling approaches were quantified through their sensitivity, bias for causal exposures, and false discovery proportion, which were averaged over 200 datasets (or simulation runs). Performances were compared to those obtained when directly applying DSA either on the error-free exposures (which corresponds to the ideal approach); or on a dataset including a single error-contaminated measure of each exposure per individual, without applying any measurement error correction

(named hereafter the *naive approach,* what epidemiologists ignoring measurement error would do, and which is expected to correspond to the least efficient approach) or after applying RC using *a priori* known ICC values for all exposures.

12.2.2 Generation of exposures

We generated a matrix $\boldsymbol{T} = (T_i)_{\mathbf{i=1,\ldots,p}}$ of *p* error-free exposures from a multivariate normal distribution with covariance matrix $\boldsymbol{\Sigma}$ (i.e., $\boldsymbol{T} \sim \mathcal{N}(0, \boldsymbol{\Sigma})$), in a fictitious population of N=1 200 subjects. $\boldsymbol{\Sigma}$ was centered-reduced such that all exposures were generated with an average of 0 and a variance of 1, and could easily be compared. For *p*=237 exposures, $\boldsymbol{\Sigma}$ was obtained from environmental factors measured in the INfancia y Medio Ambiente mother-child, or INMA cohort (one of the HELIX project cohorts) (Guxens et al., 2012). For *p*=10 exposures, $\boldsymbol{\Sigma}$ was generated so as to include a mixture of low and high pairwise correlations (in order to cover the range of correlation levels observed for p=237). For *p*=1 exposure, its variance was also set to 1 ($\boldsymbol{\Sigma}$=1).

Classical-type error assumes that the exposures are measured with independent additive errors. We thus generated, for each exposure T_i, 10 error-contaminated biospecimens per individual, denoted $(W_{ij})_{\mathrm{j=1,\ldots,10}}$, that were assumed to be collected at random time points within the exposure window of interest, using the same equation as in (Perrier et al., 2016):

$$W_{ij} = T_i + U_{ij} \qquad \text{Eq. (1)}$$

where $U_{ij} \sim \mathcal{N}(0,\ \sigma^2_{U_i})$, with $\sigma^2_{U_i} = \sigma^2_{T_i}\left(\frac{1}{ICC_i} - 1\right)$, $\sigma^2_{T_i}$ being the variance and ICC_i the ICC of exposure T_i. We generated ICC_i in the [0.15, 1.00] range, from a balanced mixture of two normal distributions $\mathcal{N}(0.95, 0.15)$ and $\mathcal{N}(0.50, 0.20)$.

12.2.3 Generation of the health outcome

The health outcome Y was generated as a function of the error-free exposures according to:

$$Y = \sum_{\mathrm{i=1}}^{\mathrm{p}} \beta_{\mathrm{i}} T_i + \epsilon \qquad \text{Eq. (2)}$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$ and where the regression coefficients β_{i} were set to 0, except for the k true predictors (TP) that were assumed to be causally related to the outcome, for which β_{i} was 1. TPs were randomly selected amongst all exposures at each simulation run. The proportion of variance explained by the *k* TPs was set to $\mathrm{R}^2 = 3\,\% \times k$.

12.2.4 Estimation of the exposure-health association

Two methods aiming to correct for exposure measurement error were applied to the simulated data. In the within-subject biospecimens pooling approach (Perrier et al., 2016), all concentrations W_{ij} that are considered available for each exposure T_i are averaged at the individual level with equal weights; this measure is named hereafter $\widehat{\boldsymbol{T}}_{\boldsymbol{pool}} = \overline{\boldsymbol{W}} = (\overline{\boldsymbol{W_i}})_{\boldsymbol{i=1,\ldots,p}}$. In the RC approach (Carroll et al., 2006), in the absence of adjustment factors, proxies of exposures are obtained by predicting $\boldsymbol{T} = (T_i)_{\mathrm{i=1,\ldots,p}}$ from all available error-contaminated measures relying on a multivariate linear regression model:

$$\widehat{\boldsymbol{T}}_{\boldsymbol{RC}} = \hat{\mu}_{\overline{W}} + \frac{\widehat{\boldsymbol{\Sigma}}_{\mathbf{T}}^{\mathbf{t}}}{\widehat{\boldsymbol{\Sigma}}_{\mathbf{T}} + \widehat{\boldsymbol{\Sigma}}_{\mathbf{U}}}(\overline{W} - \hat{\mu}_{\overline{W}}) \qquad \text{Eq. (3)}$$

where $\hat{\mu}_{\overline{W}}$ is the exposure-by-exposure mean of $\overline{\boldsymbol{W}}$, $\widehat{\boldsymbol{\Sigma}}_{\mathbf{T}}$ is the estimated covariance matrix of the error-free exposures $\boldsymbol{T}$, and $\widehat{\boldsymbol{\Sigma}}_{\mathbf{U}}$ is the estimated covariance matrix of measurement error terms, with diagonal coefficients being divided by the average number of repeated biospecimens per

subject for the given exposure. The estimated ICCs (named $(\widehat{ICC}_i)_{i=1,\ldots,p}$) can be derived from $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{U}}$. When no repeated biospecimens are available, one has to assume that the ICCs for the given exposures are known from a previous study.

We separately assessed the associations of $\widehat{\boldsymbol{T}}_{\boldsymbol{pool}}$ and $\widehat{\boldsymbol{T}}_{\boldsymbol{RC}}$ with the health outcome Y using the deletion/substitution/addition (DSA) algorithm (Sinisi and van der Laan, 2004). We selected this sequential variable selection algorithm for it was shown to have equivalent or better performances in comparison with other dimension reduction or variable selection approaches such as elastic net or exposome-wide association study in a simulation study based on a similar approach and similar hypotheses but without exposure measurement error (Agier et al., 2016). DSA model search process starts with the intercept model, and, at each iteration, considers removing a term, replacing a term by another, or adding a term to the current model. The final model is selected by minimizing the prediction root mean squared error using 5-fold cross-validated data. We did not allow polynomial or interaction terms, and limited the maximum model size to 40 covariates, a number never reached in our simulations.

12.2.5 Statistical performance assessment

For each simulation run and scenario, we assessed the performances of each approach by calculating the sensitivity, defined as the proportion of TP that were selected by the method; and the false discovery proportion (FDP), defined as the proportion of selected variables that were not TP. When no variable was selected, the FDP was given a value of 0. We also computed the mean absolute bias for TP, which measures the accuracy of the estimated coefficients by comparing the coefficient value β_i that was used to generate the outcome *Y* with its estimated value $\hat{\beta}_i$ obtained by DSA, i.e.:

$$\frac{1}{k}\sum_{i=1}^{p} |\beta_i - \hat{\beta}_i| \times \beta_i \qquad \text{Eq. (4)}$$

(note that β_i only takes values 0 or 1).

Performances were averaged by scenario and measurement error correction method. We further investigated the effect of the ICC level by stratifying exposures by ICC decile and estimating these criteria values within each decile group.

Additional scenarios

Unbalanced designs: In order to limit the number of biospecimens to be collected, we considered: (i) collecting repeated biospecimens in a subset of the study participants only, i.e. in 10 %, 40 %, or 70 % randomly selected subjects within our population; or (ii) collecting or assaying repeated biospecimens only for exposures known to have a high within-subject variability, i.e. exposures with an ICC lower than 0.4, 0.6 or 0.8 (in this case, RC requires an *a priori* known ICC value for each exposure without repeated biospecimens).

Varying correlation levels: Three additional situations were considered with varying exposures correlation structures: one in which the off-diagonal elements of the covariance matrix of the true exposures $\boldsymbol{\Sigma}$ were divided by two (labelled $\boldsymbol{\Sigma}^{-}$); one in which they were multiplied by two (labelled $\boldsymbol{\Sigma}^{+}$) and one with a diagonal correlation structure ($\boldsymbol{\Sigma}^{0}$). When needed, we computed the closest positive semi-definite matrix of the correlation structure before generating the exposures (Agier et al., 2016). These scenarios were tested in a *p*=237 exposures context only, with 100 simulation runs. Analyses were performed using the R software (www.r-project.org).

12.3 Main findings

12.3.1 Low dimension exposure-health studies

In an exposure-health study assuming that 10 exposures were assessed without measurement error and that a single TP affected the health outcome, sensitivity was 100 %, mean absolute bias for TP was 0.15, and FDP was 18 %. In the presence of uncorrected measurement error, sensitivity decreased to 94 %, bias increased to 0.42, and FDP increased to 31 %. Applying measurement error correction methods improved the performances compared to this naive approach, with efficiency raising as the number of repeated biospecimens increased and reaching, with both methods, values that were similar to the case without exposure measurement error (Table 12.1). Similar results were obtained when we considered 3 or 10 TP, or in a single exposure-health study.

Table 12.1: Performances in identifying exposure-health associations, scenarios considering either 1, 10, or 237 exposures in total

Exposure assessment	Measurement error correction	1-exposure study [1]		10-exposures study			237-exposures study		
		Sensitivity	MAB[2]	Sensitivity	FDP	MAB[2]	Sensitivity	FDP	MAB [2]
Error-free exposures	None	1.00	0.13	1.00	0.18	0.15	0.80	0.24	0.30
Error-contaminated biospecimens									
1 biospecimen	None (naive approach)	0.98	0.41	0.94	0.31	0.42	0.49	0.40	0.63
	RC (ICCs provided)	0.98	0.18	0.90	0.24	0.26	0.56	0.41	0.52
2 biospecimens	Pooling	0.99	0.29	0.98	0.24	0.31	0.60	0.39	0.53
	RC	0.99	0.15	0.96	0.22	0.19	0.62	0.37	0.47
5 biospecimens	Pooling	1.00	0.18	1.00	0.19	0.21	0.68	0.34	0.43
	RC	1.00	0.14	0.98	0.17	0.17	0.76	0.26	0.34
10 biospecimens	Pooling	1.00	0.15	1.00	0.22	0.18	0.72	0.31	0.38
	RC	1.00	0.13	0.98	0.18	0.17	0.77	0.25	0.33

Sensitivity, FDP and mean absolute bias for true predictors (MAB) are estimated, considering scenarios with the continuous health outcome being affected by a single true predictor that was randomly selected amongst all available exposures at each simulation run. Results are given for DSA directly applied on the error-free exposures and on one error-contaminated measure (naive approach); for DSA after applying RC on one error-contaminated measure with ICCs values being provided; and for DSA after applying RC or the within-subject biospecimens pooling approach on error-contaminated measures with a balanced design with 2, 5 or 10 repeated measures per subject. Average values computed over 200 simulation runs are displayed.

Abbreviations: DSA: Deletion/substitution/addition algorithm; FDP: False discovery proportion; ICC: Intra-class correlation; MAB: Mean absolute bias for true predictors (variables that were not

predictors were ignored); Pooling: Within-subject biospecimens pooling; RC: Regression calibration.

[1] FDP is not defined when there is a single exposure that is a true predictor (which means there is no variable that is not a true predictor).

[2] MAB is estimated among true predictors only, for which the true parameter value β_i was one.

12.3.2 Exposome-health studies

When we considered an exposome context with 237 exposures and a single TP affecting the health outcome, in the absence of measurement error, sensitivity was 80 %, FDP was 24 % and bias for TP was 0.30. In the presence of measurement error, if no repeated biospecimen was available, the naive approach resulted in a strong performance deterioration: sensitivity decreased to 49 %, FDP was 40 % and bias for TP was 0.63. Assuming that external ICCs were available, RC allowed slightly improving sensitivity (56 %) and bias (0.52), while FDP remained unchanged. When repeated biospecimens were available and a balanced design was considered (i.e., an equal number of repeated biospecimens for all subjects and exposures), performances of methods correcting for measurement error varied between those of the naive approach and those obtained in the absence of measurement error (without reaching the latter). RC was marginally more efficient than pooling, with a major improvement being observed when increasing the number of repeated biospecimens from 2 to 5, and almost no difference between the scenarios with 5 and 10 repeated biospecimens. In comparison, pooling displayed a more regular improvement in performances as the number of repeated biospecimens increased (Figure 12.1, Table 12.1). Similar observations were made when we considered 3 or 10 TP, except that performances were slightly decreased for all scenarios (an overall 13 % decrease in sensitivity, 11 % increase in FDP and 0.19 increase in bias when considering 10 compared to 1 TP, Figure 12.1). When we assumed that no exposure affected the health outcome, the chance to mistakenly select an exposure that was not a TP did not increase as a result of exposures measurement error, compared to a situation where exposures were measured without error.

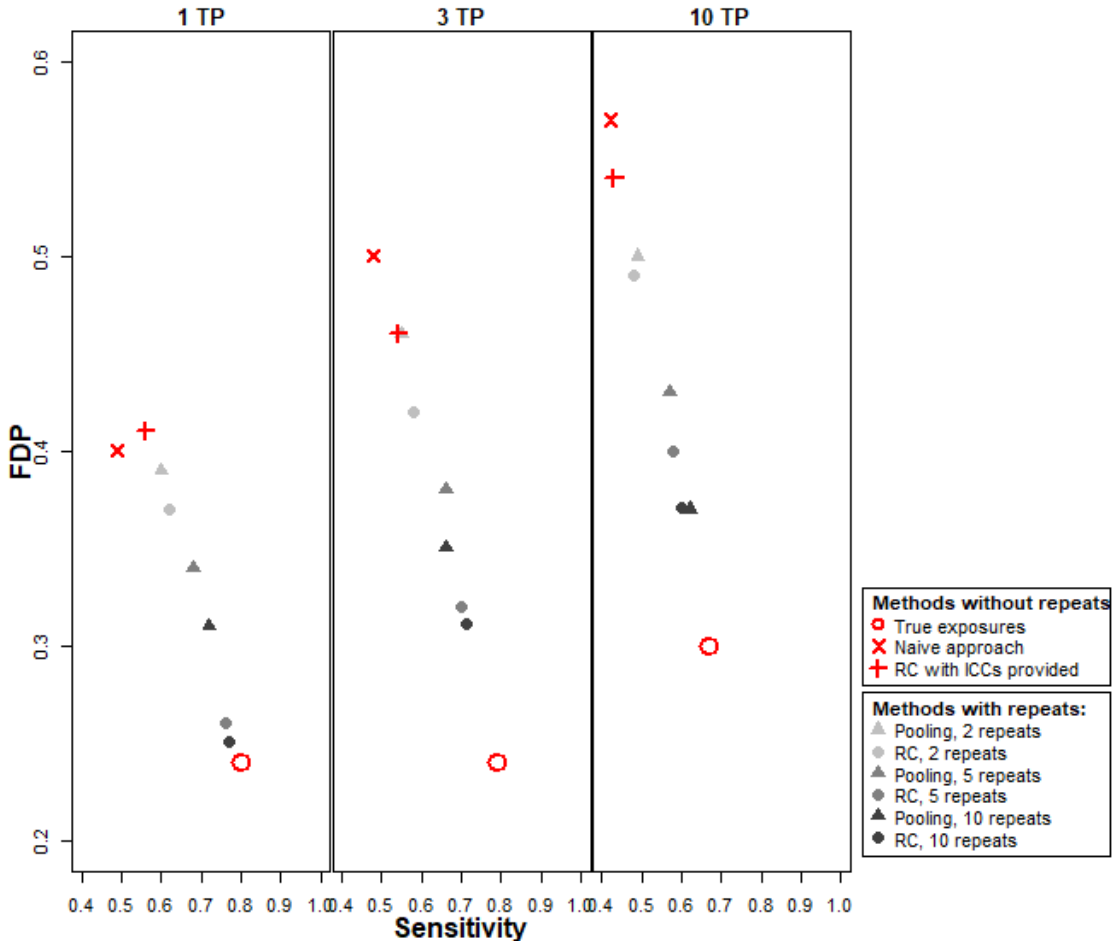


Figure 12.1: Sensitivity and FDP, considering scenarios with the continuous health outcome being affected by 1, 3 or 10 true predictors randomly selected amongst a set of 237 exposures

Results are given for DSA directly applied on the error-free exposures (red circle) and on one error-contaminated measure (naive approach, red cross); for DSA after applying RC on one error-contaminated measure with ICCs values being provided (red + sign); and for DSA after applying RC (triangles) or the within-subject biospecimens pooling approach (circles) on error-contaminated measures with a balanced design with 2,5 or 10 repeated measures per subject. Average values computed over 200 simulation runs are displayed.

DSA: Deletion/substitution/addition algorithm; FDP: False discovery proportion; ICC: Intra-class correlation; Pooling: Within-subject biospecimens pooling; RC: Regression calibration; TP: True predictors.

Plotting performances as a function of the exposures' ICC (Figure 12.2) showed that for persistent exposures (those with an ICC close to 1), sensitivity and bias were close to the ones observed in the absence of measurement error, even when using the naïve approach. The performances of both methods decreased as the exposure variability increased (i.e., when the ICC decreased), with a linear decline in sensitivity and a linear increase in bias (Figure 12.2 A, C). In contrast, the risk of false discovery (FDP, Figure 12.2B) was always greater than the one observed in the absence of measurement error, and varied little as a function of the ICC in the [0.4, 1.0] range: for a given number of repeated biospecimens and a given correction method (including no correction), all exposures with an ICC in the [0.4, 1.0] range had a similar chance to be mistakenly selected when they were not truly associated with the outcome. In comparison, the least persistent exposures

(those with an ICC in the [0.15;0.40] range) had a high risk of being mistakenly selected, even when 5 or 10 repeated biospecimens were collected.

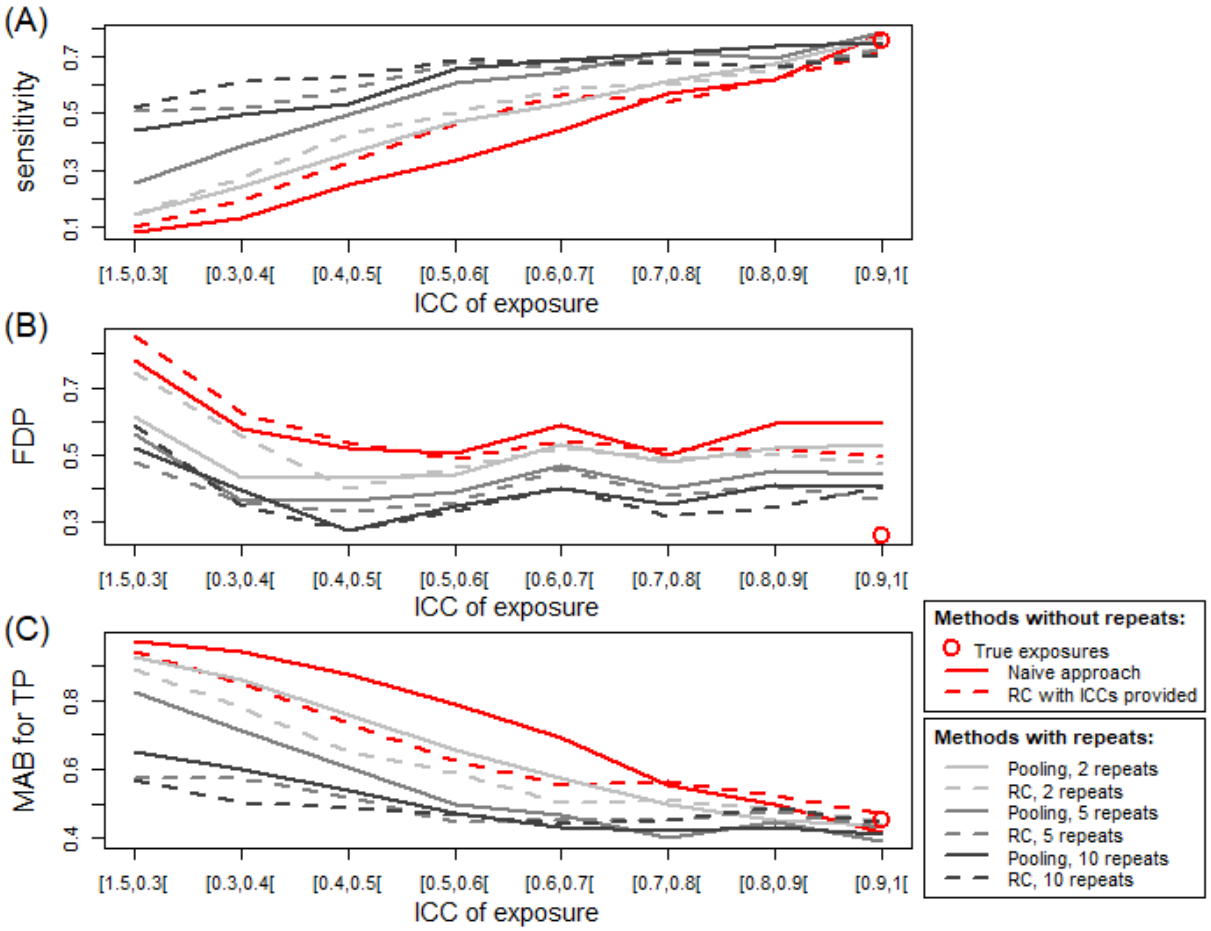


Figure 12.2: Sensitivity (A), FDP (B) and mean absolute bias for true predictors (C) amongst exposures inside each ICC decile, considering scenarios with the continuous health outcome being affected by 1, 3 or 10 true predictors randomly selected amongst a set of 237 exposures.

Results are given for DSA directly applied on error-free exposures (red circle) and on one error-contaminated measure (naive approach, plain thick line); for DSA after applying RC on one error-contaminated measure with ICCs values being provided (dashed thick line); and for DSA after applying RC (plain lines) or the within-subject biospecimens pooling approach (dashed lines) on error-contaminated measures with a balanced design with 2, 5 or 10 repeated measures per subject. Average values per ICC decile computed over all simulation runs (200 simulation runs per scenario, pooled results) are displayed.

DSA: Deletion/substitution/addition algorithm; FDP: False discovery proportion; ICC: Intra-class correlation; MAB: Mean absolute bias (variables that were not predictors were ignored); Pooling: Within-subject biospecimens pooling; RC: Regression calibration; TP: True predictors.

12.3.3 Unbalanced designs

The performances of the methods decreased linearly as the proportion of subjects with repeated biospecimens decreased from 100 % to 40 %; at 40 % of subjects with repeated biospecimens, performances were similar to the ones observed with the naive approach; at 10 % of subjects with

repeated biospecimens, measurement error correction methods displayed even some loss over the naive approach (Figure 12.3). RC performed better than the pooling approach, except when 2 repeated biospecimens were collected for less than half of the subjects. Sampling 5 biospecimens in 10 % of the population, or 2 biospecimens in 40 % of the population, the remaining subjects having a unique exposure measure (which corresponds to the same total number of biospecimens collected and assessed, 1.4 times the number of subjects) made no difference in terms of performance in capturing the exposure-health association.

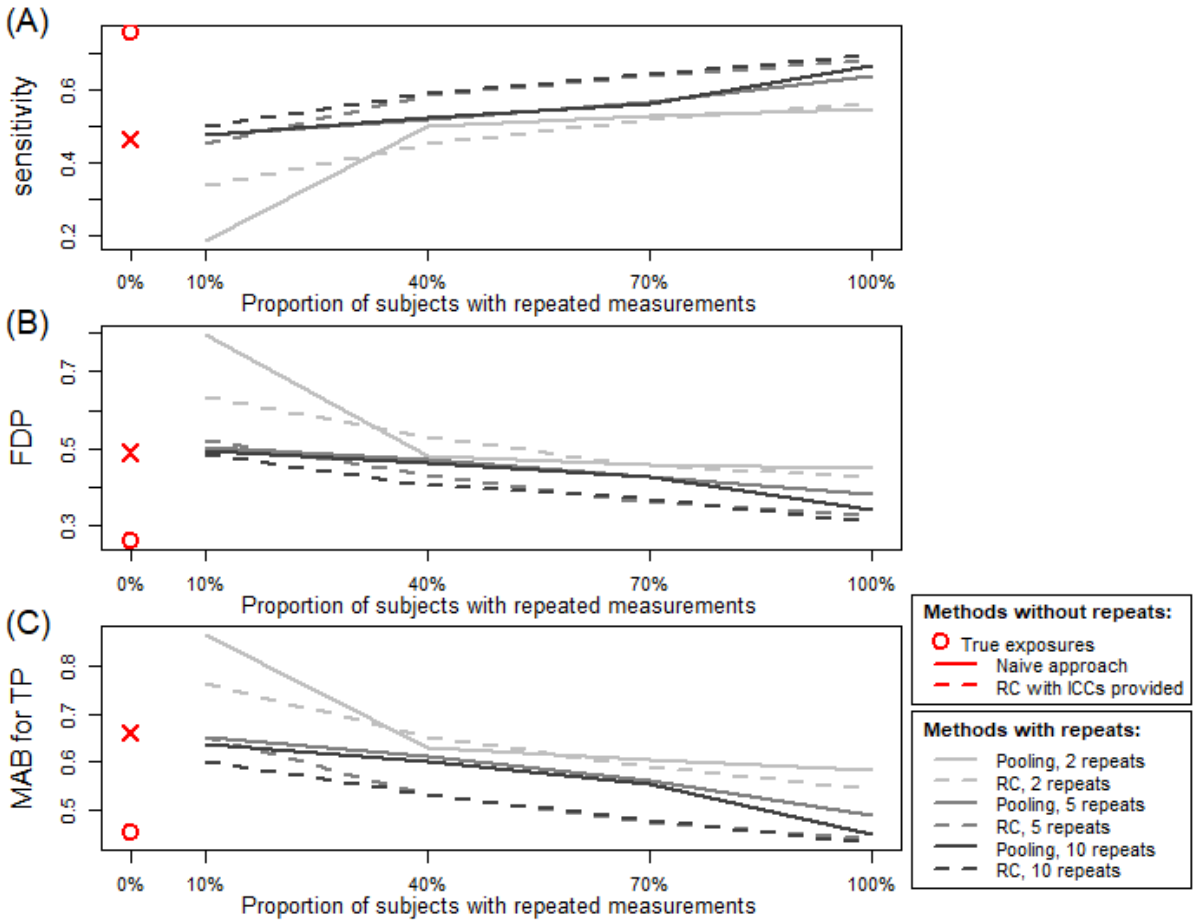


Figure 12.3: Sensitivity (A), FDP (B) and mean absolute bias for true predictors (C) when repeated biospecimens were collected in a subset of subjects, considering scenarios with the continuous health outcome being affected by 1, 3 or 10 true predictors randomly selected amongst a set of 237 exposures.

We considered the proportion of subjects with repeated biospecimens increased from 10 % to 100 %. Results are given for DSA directly applied on error-free exposures (red circle) and on one error-contaminated measure (naive approach, x symbol); and for DSA after applying RC (plain lines) or the within-subject biospecimens pooling approach (dashed lines) on error-contaminated measures with an unbalanced design with 2,5 or 10 repeated measures per subject being collected in a subgroup of the study population, a single biospecimen being collected otherwise. Average values computed over all simulation runs (100 simulation runs per scenario, pooled results) are displayed.

DSA: Deletion/substitution/addition algorithm; FDP: False discovery proportion; ICC: Intra-class correlation; MAB: Mean absolute bias (variables that were not predictors were ignored); Pooling: Within-subject biospecimens pooling; RC: Regression calibration; TP: True predictors.

Repeated biospecimens collection could be restricted to exposures with an ICC lower than 0.6 (corresponding to 42 % of exposures) without losing the benefits of measurement error correction methods (Figure 12.4); that is, performances were similar to those obtained when collecting repeated biospecimens for all exposures. Restricting repeated biospecimens to exposures with an ICC below 0.4 inflicted a substantial performance decrease.

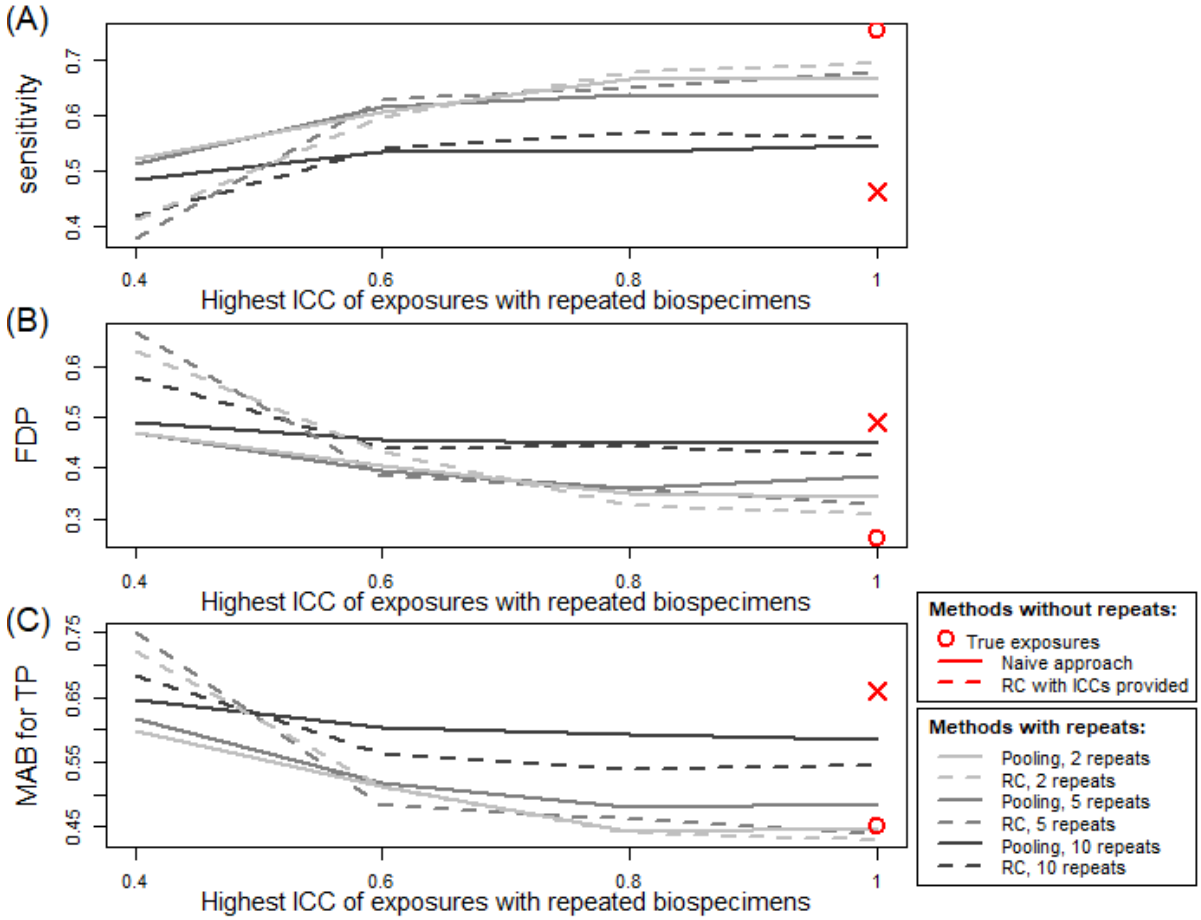


Figure 12.4: Sensitivity (A), FDP (B) and mean absolute bias for true predictors (C) when repeated biospecimens were collected in a subset of exposures, considering scenarios with the continuous health outcome being affected by 1, 3 or 10 true predictors randomly selected amongst a set of 237 exposures.

We considered values for the ICC threshold above which no repeated biospecimen is collected increased from 0.4 to 1 (1 meaning that repeated biospecimens were collected for all exposures). Results are given for DSA directly applied on error-free exposures (red circle) and on one error-contaminated measure (naive approach, x symbol); and for DSA after applying RC (plain lines) or the within-subject biospecimens pooling approach (dashed lines) on error-contaminated measures with an unbalanced design with 2,5 or 10 repeated measures per subject being collected in a subset of exposures, a single biospecimen being collected otherwise. Average values computed over all simulation runs (100 simulation runs per scenario, pooled results) are displayed.

DSA: Deletion/substitution/addition algorithm; FDP: False discovery proportion; ICC: Intra-class correlation; MAB: Mean absolute bias (variables that were not predictors were ignored); Pooling: Within-subject biospecimens pooling; RC: Regression calibration; TP: True predictors.

12.3.4 Correlation levels between exposures

The degree of correlation amongst exposures had a large impact on the models' performances. When relying on an existing exposome correlation structure (median [1st quartile, 3rd quartile] absolute value of the coefficients of correlation, 0.06 [0.03, 0.15]), over all scenarios investigated (including scenarios with error-free exposures), average (range) sensitivity was 61 % (42 %;80 %) and average FDP was 38 % (24 %;57 %). When correlation levels amongst exposures were double the initial values ($\boldsymbol{\Sigma}^{+}$ correlation matric), performances diminished: sensitivity decreased to 37% (16 %; 66 %) and FDP increased to 62 % (41 %; 78 %). With correlation levels amongst exposures half the initial values ($\boldsymbol{\Sigma}^{-}$ correlation matric), performances were higher: sensitivity was 86 % (64 %; 98 %) and FDP was 14 % (6 %; 31 %). Variable selection methods were almost perfectly efficient when exposures were uncorrelated: sensitivity was 90 % (66 %; 100 %) and FDP was 6 % (2 %; 11 %). The impact of the number of repeated biospecimens and of TPs on the performances was similar across all correlation matrices tested.

12.4 Discussion

To our knowledge, this study is the first to describe the impact of exposure measurement error in realistic exposome-health studies. The decrease of the performance of variable selection models due to measurement error is known to be substantial in low-dimension settings (Perrier et al., 2016). We showed that this phenomenon is compounded in an exposome context, with increased false positive rates, false negative rates, and bias in dose-response functions; and that these problems aggravate when correlation within the exposome increases. Compared to single exposure studies of similar population size without measurement error, exposome studies ignoring issues related to measurement error suffer from a double source of performance loss, due to the increase in the number of exposures considered and to measurement error (Figure 12.5). The performance loss due to measurement error can in large part be recovered by collecting repeated biospecimens (the more biospecimens, the better the performances) and using methods such as regression calibration or within-subject biospecimens pooling. As expected, the impact of measurement error was larger for compounds with the largest variability, so that, for a given strength of association and assuming lack of other biases, exposome studies are more likely to correctly identify the effect of the most persistent compounds, compared to the least persistent ones – an issue that could be termed ICC-related differential sensitivity. On the other hand, most exposures, except the least persistent ones, had a similar risk of being mistakenly selected when they were not associated with the outcome. Limiting repeated samples to the least persistent exposures is an option to limit costs without affecting variable selection performances.

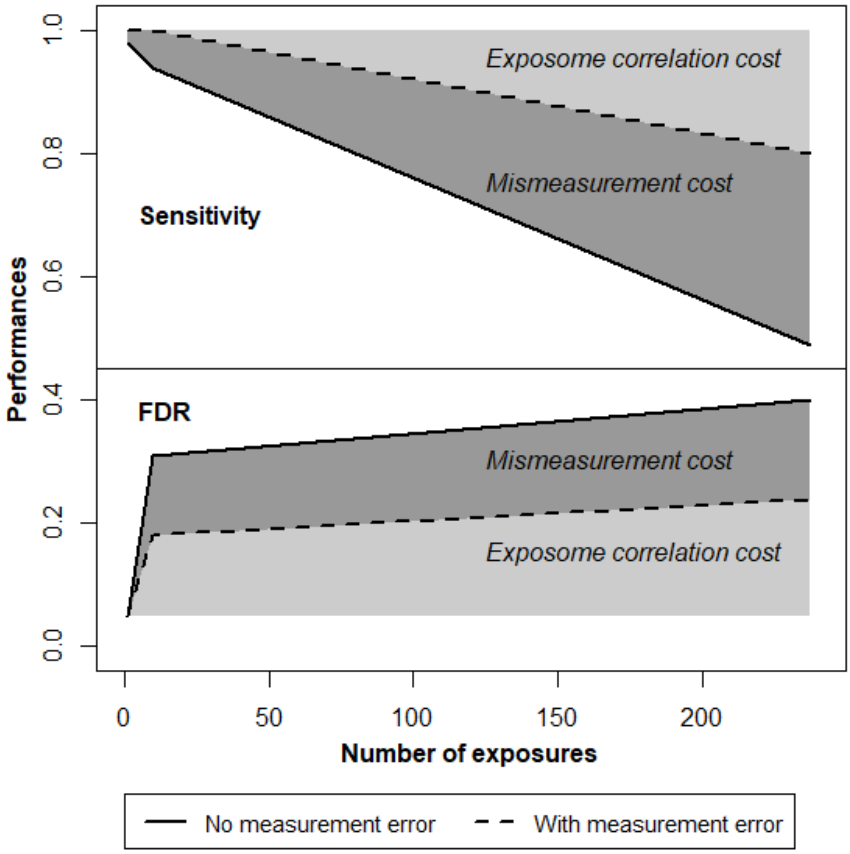


Figure 12.5: Sensitivity and FDP loss in exposome studies due to an increase in the number of exposures considered (the exposome correlation cost) and to exposures measurement error (the mismeasurement cost).

We considered scenarios with the continuous health outcome being affected by a single true predictor randomly selected amongst a set of 1, 10 or 237 exposures Results are given for DSA directly applied on the available data (in the presence of measurement error, on one error-contaminated measure, without applying a measurement error correction method (naive approach)). Average values computed over 200 simulation runs are displayed. Since FDP cannot be defined when there is a single exposure that is a true predictor (which means there is no variable that is not a true predictor), we assumed as a baseline value, a theoretical FDP of 5 % in the absence of measurement error.

DSA: Deletion/substitution/addition algorithm; FDP: False discovery proportion.

12.4.1 Mitigating the impact of measurement error in exposome studies

Regression calibration allowed some improvement over the naive approach ignoring measurement error when repeated biospecimens were available, and a good recovery of the impact of exposure measurement error when five repeated biospecimens were available per exposure and subject; there was little gain in further increasing the number of biospecimens per subject. When no repeated biospecimens are available, we cannot strongly recommend using RC, for the method did not always provide a clear improvement over the naive (uncorrected) approach; moreover “external” ICCs (i.e., stemming from another population) are needed in this case, and ICCs are not always transposable from one population to another (Vernet et al., 2019).

Our study can also be seen as a generalization in the exposome setting of the within-subject biospecimens pooling approach, which had previously been validated in a single-exposure context

(Perrier et al., 2016; Vernet et al., 2019). We assumed that no source of error was introduced when pooling biospecimens. The performances of pooling were marginally lower than those of RC in our simulations. However, two things should be kept in mind: first, estimation bias can be improved by applying *a posteriori disattenuation*, which is efficient when an unbiased estimate of the ICC is available (preferably from a subgroup of the study population) (Perrier et al., 2016; Vernet et al., 2019). Second, RC implies to assay exposures in all collected biospecimens, as opposed to one per subject in the pooling approach. Consequently, biospecimens pooling can be achieved at a much lower cost, and allows reducing the issue related to limits of detection (LOD), as pooling limits the proportion of samples below the LOD (Schisterman and Vexler, 2008; Vernet et al., 2019).

The number of biospecimens collected per subject can be reduced without affecting significantly the methods' efficiency, preferably by limiting the assessment of repeated biospecimens to exposures with the lowest ICCs (with 0.6 appearing as a relevant threshold); but also by assessing repeated samples in a subgroup of the study population (with 40 % of randomly chosen subjects appearing as a minimal threshold in our setting). In practice, one should collect repeated urinary biospecimens (urine being the matrix from which the compounds with high within-subject variability are generally assessed) but could afford to collect only one or a couple of blood samples per subject, compounds assessed in blood generally having higher ICCs. It is not possible to suggest an optimum number of biospecimens to be collected per subject, as this depends on several factors, including the ICC and likely the correlation with other exposures. Yet, in a single exposure context, about 18-35 biospecimens per subject collected in the relevant time window were required to decrease bias in the dose-response slope down to 10% (Perrier et al., 2016; Vernet et al., 2019).

12.4.2 Study limitations

We considered a simple simulation design; specifically, we did not assume the existence of confounders (although these could be incorporated using our simulation code), of LODs (the performance gain of repeated biospecimen sampling might actually be larger if the LODs are high for some exposures, because collecting more biospecimens is generally a good strategy to limit issues related to LODs (Mumford et al., 2006)), nor of measurement error affecting the health outcome. We considered classical-type error only, which is typically what can be expected for biomarker-based exposures and which most existing statistical methods for measurement error are designed for. Yet departure from this type of error may be observed for exposures assessed by other means, such as atmospheric pollutants and meteorological conditions; and correlation in the errors across exposures may also exist (which RC can account for if the corresponding information is provided (Carroll et al., 2006)). Finally, we only considered linear effects of exposures on the health outcome; non-linear effects are likely to make the identification of true predictors even more challenging.

Regarding statistical approaches, we only investigated measurement error correction approaches that transformed the exposure variables, allowing flexibility in the algorithm to assess the link between exposures and the health parameter. Although we did not quantify it, we have no reason to believe that the impact of measurement error or of measurement error correction techniques would be different if a different variable selection model was used at the exposome-health step. Several algorithms that jointly correct for measurement error and perform variable selection were recently developed (Liang and Li, 2009; Ma and Li, 2010; Sørensen et al., 2012; Wang et al., 2012), and can in principle be used in exposome studies. In practice, they are complex to

implement, usually not available in standard statistical software and sometimes not applicable to all types of regression models or to all settings (Guangning, 2014).

Regarding our first step, instead of using RC for correcting for measurement error, moment reconstruction and multiple imputation may be considered. In the context of regression models without variable selection, a simulation study showed that RC carried efficiency gains that were sometimes dramatic over these two substitution methods (Freedman et al., 2008). Regarding the way RC was applied, here we have conducted RC correction only once, and have included all available exposures. Yet, unlike the pooling approach, RC is a multivariate model-based correction whose results depend on the set of exposures that are included in the model. Hence, ideally, RC correction should be refitted at each step of the DSA procedure with the updated set of retained exposures.

This could improve the method performances but appears cumbersome in practice. One alternative would be to correct each exposure independently; in balanced design scenarios without adjustment factors, this procedure is equivalent to the pooling approach when combined with a variable selection procedure, as the resulting exposure proxies are a rescaled version of the pooled biospecimens estimate. Finally, in the subsequent exposure-health step, model's parameters standard errors are biased for both the RC and the pooling approaches (they do not account for the fact that exposures are measured with error (Armstrong and Basagaña, 2015; Spiegelman et al., 2001)). Proper standard error estimates may be obtained by bootstrap; yet estimation bias of standard errors does not affect our results, as the variable selection criteria in DSA does not depend on the coefficients standard error, but on the model fit.

12.4.3 Conclusions and possible strategy for future exposome studies

As demonstrated here and in previous studies, the factors that negatively affect the performances of statistical methods in studies linking exposome and health are notably (i) a large number of exposures; (ii) the existence of (even moderate) correlation between exposures; and (iii) differential measurement error (Figure 12.5). Schematically, an exposome study that considers both persistent and non-persistent biomarkers risks being short-sighted (underpowered) for the least persistent exposures if no specific efforts to correct the related unbalanced power is made. Collecting repeated biospecimens and using within-subject pooling or regression calibration on these repeated biospecimens allows improving performances up to a large extent when multiple biospecimens are collected per exposure and subject.

If assay cost is an issue, then the within-subject biospecimens pooling approach, which allows improving performances without increasing assay cost (but with an increase in biospecimens collection and handling costs compared to the approach with one biospecimen per subject) should be preferred; limiting repeated samples to the least persistent exposures is a further option to limit costs.

In conclusion, measurement error issues are compounded in exposome studies compared to smaller dimension exposure-health studies. Exposome research has the more or less overtly stated aim to hierarchize exposures in terms of strength of their association with health; as we illustrated here (see in particular Figure 12.2), this aim is unlikely to be achieved without implementing, from the step of study design, specific measurement error correction approaches such as those relying on the collection of repeated biospecimens in each subject.

12.5 References

Agier, L., Basagaña, X., Hernandez-Ferrer, C., Maitre, L., Tamayo Uria, I., Urquiza, J., Andrusaityte, S., Casas, M., de Castro, M., Cequier, E., Chatzi, L., Donaire-Gonzalez, D., Giorgis-Allemand, L., Gonzalez, J.R., Grazuleviciene, R., Gützkow, K.B., Haug, L.S., Sakhi, A.K., McEachan, R.R.C., Meltzer, H.M., Nieuwenhuijsen, M., Robinson, O., Roumeliotaki, T., Sunyer, J., Thomsen, C., Vafeiadi, M., Valentin, A., West, J., Wright, J., Siroux, V., Vrijheid, M., Slama, R., 2020. Association between the pregnancy exposome and fetal growth (in press). Int. J. Epidemiol.

Agier, L., Basagaña, X., Maitre, L., Granum, B., Bird, P., Casas, M., Oftedal, B., Wright, J., Andrusaityte, S., de Castro, M., Cequier, E., Chatzi, L., Donaire-Gonzales, D., Grazuleviciene, R., Haug, L.S., Sakhi, A.K., Leventakou, V., McEachan, R., Nieuwenhuijsen, M., Petraviciene, I., Robinson, O., Roumeliotaki, T., Sunyer, J., Tamayo, I., Thomsen, C., Urquiza, J., Valentin, A., Slama, R., Vrijheid, M., Siroux, V., 2019. Early life exposome and lung function in children from the HELIX cohort. Lancet Planet. Heal. 3, e81_e92. https://doi.org/10.1016/S2542-5196(19)30010-5

Agier, L., Portengen, L., Chadeau-Hyam, M., Basagaña, X., Giorgis-Allemand, L., Siroux, V., Robinson, O., Vlaanderen, J., González, J.R., Nieuwenhuijsen, M.J., Vineis, P., Vrijheid, M., Slama, R., Vermeulen, R., 2016. A Systematic Comparison of Linear Regression-Based Statistical Methods to Assess Exposome-Health Associations. Environ. Health Perspect. 124, 1848–1856. https://doi.org/10.1289/EHP172

Armstrong, B., Basagaña, X., 2015. Exposure Measurement Error: Consequences and Design Issues, in: Exposure Assessment in Environmental Epidemiology. Oxford University Press. https://doi.org/10.1093/med/9780199378784.003.0010

Armstrong, B.G., 1998. Effect of measurement error on epidemiological studies of environmental and occupational exposures. Occup. Environ. Med. 55, 651–656. https://doi.org/10.1136/oem.55.10.651

Brakenhoff, T.B., Mitroiu, M., Keogh, R.H., Moons, K.G.M., Groenwold, R.H.H., van Smeden, M., 2018a. Measurement error is often neglected in medical literature: a systematic review. J. Clin. Epidemiol. 98, 89–97. https://doi.org/10.1016/j.jclinepi.2018.02.023

Brakenhoff, T.B., van Smeden, M., Visseren, F.L.J., Groenwold, R.H.H., 2018b. Random measurement error: Why worry? An example of cardiovascular risk factors. PLoS One 13, e0192298. https://doi.org/10.1371/journal.pone.0192298

Brown, B., Weaver, T., Wolfson, J., 2019. MEBoost: Variable selection in the presence of measurement error. Stat. Med. 38, 2705–2718. https://doi.org/10.1002/sim.8130

Buzas, J.S., Stefanski, L.A., Tosteson, T.D., 2005. Measurement Error, in: Handbook of Epidemiology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 729–765. https://doi.org/10.1007/978-3-540-26577-1_19

Calafat, A.M., Ye, X., Silva, M.J., Kuklenyik, Z., Needham, L.L., Kolossa, M., Tuomisto, J., Astrup Jensen, A., 2006. Human exposure assessment to environmental chemicals using biomonitoring, in: International Journal of Andrology. pp. 166–171. https://doi.org/10.1111/j.1365-2605.2005.00570.x

Carroll, R., Ruppert, D., Stefanski, L., Crainiceanu, C., 2006. Measurement Error in Nonlinear Models: A Modern Perspective, CRC Press. ed. https://doi.org/10.1111/j.1467-985X.2007.00528_4.x

Casas, M., Basagaña, X., Sakhi, A.K., Haug, L.S., Philippat, C., Granum, B., Manzano-Salgado, C.B., Brochot, C., Zeman, F., de Bont, J., Andrusaityte, S., Chatzi, L., Donaire-Gonzalez, D., Giorgis-Allemand, L., Gonzalez, J.R., Gracia-Lavedan, E., Grazuleviciene, R., Kampouri, M., Lyon-Caen, S., Pañella, P., Petraviciene, I., Robinson, O., Urquiza, J., Vafeiadi, M., Vernet, C., Waiblinger, D., Wright, J., Thomsen, C., Slama, R., Vrijheid, M., 2018. Variability of urinary concentrations of non-persistent

chemicals in pregnant women and school-aged children. Environ. Int. 121, 561–573. https://doi.org/10.1016/j.envint.2018.09.046

de Klerk, N.H., English, D.R., Armstrong, B.K., 1989. A review of the effects of random measurement error on relative risk estimates in epidemiological studies. Int. J. Epidemiol. 18, 705–712. https://doi.org/10.1093/ije/18.3.705

Donoho, D., Jin, J., 2008. Higher criticism thresholding: Optimal feature selection when useful features are rare and weak. Proc. Natl. Acad. Sci. U. S. A. 105, 14790–14795. https://doi.org/10.1073/pnas.0807471105

Freedman, L.S., Midthune, D., Carroll, R.J., Kipnis, V., 2008. A comparison of regression calibration, moment reconstruction and imputation for adjusting for covariate measurement error in regression. Stat. Med. 27, 5195–216. https://doi.org/10.1002/sim.3361

Guangning, X., 2014. Variable Selection in Measurement Error Models. https://doi.org/10.1017/CBO9781107415324.004

Guxens, M., Ballester, F., Espada, M., Fernández, M., Grimalt, J., Ibarluzea, J., Olea, N., Rebagliato, M., Tardón, A., Torrent, M., Vioque, J., Vrijheid, M., Sunyer, J., 2012. Cohort Profile: The INMA – INfancia y Medio Ambiente–(Environment and Childhood) Project. Int. J. Epidemiol. 930–940. https://doi.org/10.1093/ije/dyr054

Jurek, A.M., Greenland, S., Maldonado, G., Church, T.R., 2005. Proper interpretation of non-differential misclassification effects: expectations vs observations. Int. J. Epidemiol. 34, 680–687. https://doi.org/10.1093/ije/dyi060

LaKind, J.S., Idri, F., Naiman, D.Q., Verner, M.A., 2019. Biomonitoring and Nonpersistent Chemicals-Understanding and Addressing Variability and Exposure Misclassification. Curr. Environ. Heal. reports. https://doi.org/10.1007/s40572-019-0227-2

Liang, H., Li, R., 2009. Variable Selection for Partially Linear Models with Measurement Errors. J. Am. Stat. Assoc. 104, 234–248. https://doi.org/10.1198/jasa.2009.0127

Lyon-Caen, S., Siroux, V., Lepeule, J., Lorimier, P., Hainaut, P., Mossuz, P., Quentin, J., Supernant, K., Meary, D., Chaperot, L., Bayat, S., Cassee, F., Valentino, S., Couturier-Tarrade, A., Rousseau-Ralliard, D., Chavatte-Palmer, P., Philippat, C., Pin, I., Slama, R., 2019. Deciphering the impact of early-life exposures to highly variable environmental factors on foetal and child health: Design of SEPAGES couple-child cohort. Int. J. Environ. Res. Public Health 16, 3888. https://doi.org/10.3390/ijerph16203888

Ma, Y., Li, R., 2010. Variable Selection in Measurement Error Models. Bernoulli (Andover). 16, 274–300. https://doi.org/10.3150/09-BEJ205

Mumford, S.L., Schisterman, E.F., Vexler, A., Liu, A., 2006. Pooling biospecimens and limits of detection: effects on ROC curve analysis. Biostatistics 7, 585–598. https://doi.org/10.1093/biostatistics/kxj027

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, 378–388. https://doi.org/10.1097/EDE.0000000000000460

Preau, J.L., Wong, L.Y., Silva, M.J., Needham, L.L., Calafat, A.M., 2010. Variability over 1 week in the urinary concentrations of metabolites of diethyl phthalate and di(2-ethylhexyl) phthalate among eight adults: An observational study. Environ. Health Perspect. 118, 1748–1754. https://doi.org/10.1289/ehp.1002231

Schisterman, E.F., Albert, P.S., 2012. The biomarker revolution. Stat. Med. 31, 2513–5. https://doi.org/10.1002/sim.5499

Schisterman, E.F., Vexler, A., 2008. To pool or not to pool, from whether to when: Applications of pooling to biospecimens subject to a limit of detection. Paediatr. Perinat. Epidemiol. 22, 486–496. https://doi.org/10.1111/j.1365-3016.2008.00956.x

Sinisi, S.E., van der Laan, M.J., 2004. Deletion/Substitution/Addition Algorithm in Learning with Applications in Genomics. Stat. Appl. Genet. Mol. Biol. 3, 1–38. https://doi.org/10.2202/1544-6115.1069

Slama, R., Vrijheid, M., 2015. Some Challenges of Studies Aiming to Relate the Exposome to Human Health. Occup. Environ. Med. 0.

Sørensen, Ø., Frigessi, A., Thoresen, M., 2012. Measurement Error in Lasso: Impact and Correction 1–64. https://doi.org/10.5705/ss.2013.180

Spiegelman, D., Carroll, R.J., Kipnis, V., 2001. Efficient regression calibration for logistic regression in main study/internal validation study designs with an imperfect reference instrument. Stat. Med. 20, 139–160. https://doi.org/10.1002/1097-0258(20010115)20:1<139::AID-SIM644>3.0.CO;2-K

Tamayo-Uria, I., Maitre, L., Thomsen, C., Nieuwenhuijsen, M.J., Chatzi, L., Siroux, V., Aasvang, G.M., Agier, L., Andrusaityte, S., Casas, M., de Castro, M., Dedele, A., Haug, L.S., Heude, B., Grazuleviciene, R., Gutzkow, K.B., Krog, N.H., Mason, D., McEachan, R.R.C., Meltzer, H.M., Petraviciene, I., Robinson, O., Roumeliotaki, T., Sakhi, A.K., Urquiza, J., Vafeiadi, M., Waiblinger, D., Warembourg, C., Wright, J., Slama, R., Vrijheid, M., Basagaña, X., 2019. The early-life exposome: Description and patterns in six European countries. Environ. Int. 123, 189–200. https://doi.org/10.1016/J.ENVINT.2018.11.067

Vasquez, M.M., Hu, C., Roe, D.J., Halonen, M., Guerra, S., 2019. Measurement error correction in the least absolute shrinkage and selection operator model when validation data are available. Stat. Methods Med. Res. 28, 670–680. https://doi.org/10.1177/0962280217734241

Vernet, C., Philippat, C., Agier, L., Calafat, A.M., Ye, X., Lyon-Caen, S., Hainaut, P., Siroux, V., Schisterman, E.F., Slama, R., 2019. An empirical validation of the within-subject biospecimens pooling approach to minimize exposure misclassification in biomarker-based studies. Epidemiology. https://doi.org/10.1097/ede.0000000000001056

Vrijheid, M., Slama, R., Robinson, O., Chatzi, L., Coen, M., van den Hazel, P., Thomsen, C., Wright, J., Athersuch, T.J., Avellana, N., Basagaña, X., Brochot, C., Bucchini, L., Bustamante, M., Carracedo, A., Casas, M., Estivill, X., Fairley, L., van Gent, D., Gonzalez, J.R., Granum, B., Gražulevičienė, R., Gutzkow, K.B., Julvez, J., Keun, H.C., Kogevinas, M., McEachan, R.R.C., Meltzer, H.M., Sabidó, E., Schwarze, P.E., Siroux, V., Sunyer, J., Want, E.J., Zeman, F., Nieuwenhuijsen, M.J., 2014. The Human Early-Life Exposome (HELIX): Project Rationale and Design. Environ. Health Perspect. 122, 535–544. https://doi.org/10.1289/ehp.1307204

Wang, H.J., Stefanski, L.A., Zhu, Z., 2012. Corrected-loss estimation for quantile regression with covariate measurement errors. Biometrika 99, 405–421. https://doi.org/10.1093/biomet/ass005

Wielgomas, B., 2013. Variability of urinary excretion of pyrethroid metabolites in seven persons over seven consecutive days-Implications for observational studies. Toxicol. Lett. 221, 15–22. https://doi.org/10.1016/j.toxlet.2013.05.009

Wild, C.P., 2012. The exposome: from concept to utility. Int. J. Epidemiol. 41, 24–32. https://doi.org/10.1093/ije/dyr236

Ye, X., Wong, L.-Y., Bishop, A.M., Calafat, A.M., 2011. Variability of urinary concentrations of bisphenol A in spot samples, first morning voids, and 24-hour collections. Environ. Health Perspect. 119, 983–8. https://doi.org/10.1289/ehp.1002701

13 Overall conclusions

Collectively, the case studies conducted for this Deliverable highlight the need to take combined exposures into account in chemical risk assessment. It became evident that a disregard for combined exposures will lead to significant underestimations of the risks associated with chemical exposures. This is true both for the general population and for more specific exposure scenarios, e.g., those relevant to occupational settings.

In many cases, it was possible to identify drivers of mixture risks, chemicals that contribute more strongly than others to combined exposures. Often, the occurrence of drivers followed power laws, such as Pareto's rule where 20 % of the causes make up 80 % of the effects. The phenomenon arises because the frequency distributions that govern chemical exposures at the population level are independent of the potency of the chemicals that are evaluated together in the MRAs. Accordingly, the nature of drivers varies with the toxicological endpoint chosen for analysis.

The challenge was to assess risks associated with combined exposures on the basis of available toxicity and exposure data. Our case studies demonstrate that this is generally possible by using the HI method.

A bottleneck is in the provision of exposure data suitable for MRA. Ideally, such data should be from efforts where multiple chemicals were monitored together. However, data of suitable quality for this purpose are often hard to obtain. This has forced us to infer combined exposures from divergent data sources where single chemicals were monitored without regard for co-exposures to other substances, clearly a source of uncertainty in our assessments.

We suggest that the outcome of our investigations provides a valuable basis for future Human Biomonitoring strategies that can capture combined exposures. Such strategies should focus on the chemicals that we identified as drivers in several scenarios, and these chemicals should be monitored together. We realize that the implementation of this strategy will present challenges. For example, many chemicals identified as drivers are best monitored in urine samples, others in blood or hair. There is relatively little experience with sampling efforts across several matrices.

We conclude that chemicals currently regulated separately in different domains in the EU (pesticides, industrial chemicals, food contaminants, pharmaceuticals) can work together to produce significant combined effects that affect several important health endpoints. While there is the mandate to consider mixture effects in some sectorial regulations, the legal framework for assessing mixture risks for chemicals from multiple regulatory domains does not exist in the EU. It will be a challenge to develop, evaluate and implement suitable instruments that are fit for this purpose.