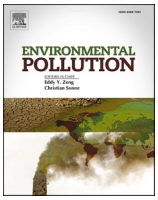
Environmental Pollution 313 (2022) 120091

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol



Suspect and non-targeted screening-based human biomonitoring identified 74 biomarkers of exposure in urine of Slovenian children☆

Žiga Tkalec [a,b], Garry Codling [c], Janja Snoj Tratnik [a], Darja Mazej [a], Jana Klánová [c], Milena Horvat [a,b], Tina Kosjek [a,b,*]

[a] Department of Environmental Sciences, Jožef Stefan Institute, Ljubljana, Slovenia
[b] Jožef Stefan International Postgraduate School, Ljubljana, Slovenia
[c] Research Centre for Toxic Compounds in the Environment, Masaryk University, Brno, Czech Republic

ARTICLE INFO

Keywords:
Exposure
Non-targeted
High resolution mass spectrometry
Exposome
HBM4EU
Biomarker

ABSTRACT

Human exposure to organic contaminants is widespread. Many of these contaminants show adverse health effects on human population. Human biomonitoring (HBM) follows the levels and the distribution of biomarkers of exposure (BoE), but it is usually done in a targeted manner. Suspect and non-targeted screening (SS/NTS) tend to find BoE in an agnostic way, without preselection of compounds, and include finding evidence of exposure to predicted, unpredicted known and unknown chemicals.

This study describes the application of high-resolution mass spectrometry (HRMS)-based SS/NTS workflow for revealing organic contaminants in urine of a cohort of 200 children from Slovenia, aged 6–9 years. The children originated from two regions, urban and rural, and the latter were sampled in two time periods, summer and winter. We tentatively identified 74 BoE at the confidence levels of 2 and 3. These BoE belong to several classes of pharmaceuticals, personal care products, plasticizers and plastic related products, volatile organic compounds, nicotine, caffeine and pesticides. The risk of three pesticides, atrazine, amitraz and diazinon is of particular concern since their use was limited in the EU. Among BoE we tentatively identified compounds that have not yet been monitored in HBM schemes and demonstrate limited exposure data, such as bisphenol G, polyethylene glycols and their ethers. Furthermore, 7 compounds with unknown use and sources of exposure were tentatively identified, either indicating the entry of new chemicals into the market, or their metabolites and transformation products. Interestingly, several BoE showed location and time dependency.

Globally, this study presents high-throughput approach to SS/NTS for HBM. The results shed a light on the exposure of Slovenian children and raise questions on potential adverse health effects of such mixtures on this vulnerable population.

1. Introduction

Humans are continuously exposed to a variety of external factors. Either physical, psychological or chemical, all of these factors make up the exposome. Chemical exposome involves exposure to hundreds of different chemicals, which usually occur at low levels. They originate from variety of sources, as for the example the environment, diet and personal care products (PCPs). With various biological effects and connected to genetic susceptibility, these chemicals are involved in etiology of many human diseases. As many as 80–85% of diseases have been reported to be linked to the exposure to environmental chemicals (Uppal et al., 2016). Children, with their biological systems and organs at various stages of development and less advanced elimination of contaminants are particularly sensitive to adverse effects caused by exposure (Ferguson et al., 2017), so it is of great importance to comprehensively describe their chemical exposome and subsequent biological responses.

To study the exposome, several approaches have been applied, such as stationary sensing, wearable devices and personal sensors (Niedzwiecki et al., 2019), however measuring external levels does not directly reflect the internal dose which actually affects biological endpoints. Hence, more answers have been offered by human biomonitoring (HBM)

☆ This paper has been recommended for acceptance by Dr Mingliang Fang.

* Corresponding author. Department of Environmental Sciences, Jožef Stefan Institute, Ljubljana, Slovenia.
E-mail address: tina.kosjek@ijs.si (T. Kosjek).

https://doi.org/10.1016/j.envpol.2022.120091
Received 24 April 2022; Received in revised form 6 August 2022; Accepted 29 August 2022
Available online 2 September 2022

that provides aggregated data of the exposures by targeting the chemicals of concern and their metabolites in biological matrices (Ganzleben et al., 2017). Conceptually, HBM can reflect temporal and spatial trends of exposure to environmental chemicals, and can serve as a measure of the efficiency of policy regulations (Pourchet et al., 2020). However, as a consequence of using traditional biomonitoring schemes and by following a limited number of only predetermined compounds, the chemicals that are not specifically targeted remain unidentified, with unknown levels, biological effects, and therefore uncharacterized risk. To define exposures holistically and to keep in time with rapidly emerging new chemicals in industrial and consumer products, non-targeted analysis (NTA) and suspect-screening (SS) approaches show promise to fill this gap. However, at present the analytical methodologies for SS/NTA human biological samples are still severely vague and represent a challenge for further development, optimization and validation (Pourchet et al., 2020).

Liquid chromatography (LC) hyphenated to high-resolution mass spectrometry (HRMS) is frequently used in non-targeted metabolomics (Schrimpe-Rutledge et al., 2016), however its employment in the exposomics involves several specificities. For example, a large number of endogenous metabolites are present in biological specimens in high quantities, while xenobiotics, being the focus of exposomics, are normally present at very low levels in the range of μg/L and lower. They often coelute and chromatographically overlap with endogenous metabolites, hindering their detectability. Furthermore, in many cases and particularly in urine biomarkers of exposure (BoE) are not parent compounds but their phase I and/or II metabolites, which may have not yet been identified and linked to their parent compounds. This increases the chemical diversity and complexity of compound identification, which is the bottleneck of NTA workflow. Given that BoE are present at very low levels, their tandem MS (MS2) spectra are difficult to obtain at sufficient quality. The MS2 spectra are used to identify the compounds by manual interpretation, by matching with spectral libraries, or by using *in-silico* fragmentation tools (Getzinger and Ferguson, 2020). The use of cheminformatics in structural annotation of exposome-related compounds is in detail elaborated in the recent review by Ljoncheva et al. (2020), whereas mass spectral libraries were addressed by Vinaixa et al. (2016).

The main objective of the study was to demonstrate a simple and high-throughput workflow for NTA and SS of HBM samples, and to provide a proof-of-concept for analysis of children's chemical exposome. In this context, we aimed to characterize the exposure of Slovenian children from urban and rural regions, the latter in two seasons, in order to distinguish between possible temporary and spatial trends.

2. Materials and methods

2.1. Study design and population

The study involved children aged from 6 to 9 years from urban and rural regions of Slovenia. We recruited 200 participants, of which 100 participants resided in the capital city of Slovenia and represented the urban part of the population. The remaining 100 participants resided in a region in the Eastern Slovenia, which is predominantly rural and characterized by intense agricultural activity. Due to the seasonal variations in agricultural practice, rural participants provided two samples, one in winter and the second at the beginning of summer, reflecting time periods with low and high intensity of nearby agricultural activity. Altogether, we obtained 300 samples of first morning urine.

Urban participants were recruited within CROME-LIFE + project ('Cross-Mediterranean Environment and Health Network', 2013–2017) (Stajnko et al., 2019), while rural participants within the national CRP project 'Exposure of children and adolescents to selected chemicals through their habitat environment (2016–2019)' (Stajnko et al., 2020). The eligibility criteria for the participants were to have resided in the selected region for at least 3 years and to not receive medication for any chronic liver or kidney disease. The population consisted of 50% female and 50% male participants, which was consistent also within location and sampling-time subgroups.

Parents or legal guardians provided the informed written consent and were able to withdraw from the study at any time. Ethical permissions were granted by Republic of Slovenia National Medical Ethics Committee (65/09/14 and 0120–118/2017/3).

2.2. Sample collection

Participants provided their first morning urine in the collection vessels, distributed by the organizer institution (Jožef Stefan Institute, Ljubljana, Slovenia). Samples were immediately aliquoted at the laboratory of the Clinical Chemistry and Biochemistry Institute of the University Medical Centre Ljubljana into 2 mL criovials, frozen on site with solid ice, and transported to Jožef Stefan Institute, where they were stored at −80 °C until the analysis.

2.3. Sample preparation

The sample preparation procedure was based on our in-house developed method (Tkalec et al., 2022). Briefly, 1 mL of urine was spiked with the internal standard ($^{13}C_3$-caffeine and ethyl paraben-$^{13}C_6$) at 3 ng/mL and deconjugated with 250 U/mL of β-glucuronidase (Abalone, purified) for 18 h at 37 °C. The samples were then extracted on Oasis HLB 60 mg 96-well plates, which were preconditioned with 1 mL of acetonitrile (ACN), methanol (MeOH) and water, respectively. After loading, the sorbent was washed with 1 mL of 5 vol/vol % MeOH in water and eluted with 1.5 mL 10 vol/vol % MeOH in ACN. The eluates were dried under a gentle stream of nitrogen and kept frozen, before being reconstituted in 50 μL of MeOH prior to the analysis.

2.4. UHPLC-HRMS2

Compounds were separated based on previously published method (Tkalec et al., 2022). In short, the separation was performed on the Shimadzu LC-30AD UHPLC using Waters Acquity HSS-T3 (2.1 × 100 mm, 1.8 μm) reversed phase column with water (A) and ACN (B) as mobile phases. Elution gradient was: 5–15% B (0.01–1 min), 15–25% B (1–5 min), 25–40% B (5–8 min), 40–60% B (8–18 min), 60–75% B (18–22 min), 75–85% B (22–24 min), 85–100% B (24–28 min), 100-5% B (28–30 min), 5% B, 30–35 min. Flow rate was 0.3 mL/min, column was heated to 35 °C. The injection volume was 1 μL.

The UHPLC was coupled to Orbitrap Fusion™ Tribrid™ Mass Spectrometer (Thermo Fisher Scientific Inc., Waltham, MA). Heated electrospray ionization (HESI) was used as the ionization source, at the spray voltage of +4500 V or −3500 V, sheath gas flow 40 L/min, nebulizer auxilliary gas 15 L/min, and sweep gas 2 L/min. The ion transfer capillary was heated to 350 °C. The full-scan mass acquisition covered the mass range of 100–900 m/z and was performed at the resolution of 120 000 FWHM (full width at half maximum at m/z 200) while MS2 data was acquired with the resolution of 60 000 FWHM at the collision energy of 20 eV. The automatic gain control (AGC) was 5×10^5 ions and maximum injection time 50 ms. Cycle time was 0.8 s.

2.4.1. Inclusion list for suspect screening

To enable efficient suspect screening, a suspect inclusion list was used during data acquisition. Suspect list included 801 entries of calculated m/z values for molecular ions of xenobiotic compounds compiled from Exposome Explorer (Neveu et al., 2017), T3DB (Wishart et al., 2015) and HBM4EU priority substances and pesticides and their 1st phase metabolites considered in the Harmonized SPECIMEN study (https://www.hbm4eu.eu/). The inclusion list contained entries of various classes of contaminants, including pesticides, plasticizers, plastic-related chemicals, PCPs and persistent organic pollutants. Each matching ion within the mass error of ±5 ppm was fragmented and its MS^2 was recorded in the ion trap (IT) segment of Orbitrap Fusion™

Tribrid™ Mass Spectrometer. Hence, the MS2 of the matching ions from our suspect list were recorded at the unit resolution. Such procedure enabled simultaneous SS/NTA screening.

2.5. Quality control (QC) procedures

QC was maintained at several levels, i.e. by QC samples, procedural blanks, solvent blanks and by the predetermined acquisition order.

QC samples were prepared by pooling 200 μL of samples, and spiking the pool with the isotopically labelled internal standards at 3 ng/mL and with a mixture of non-labelled standards to the final concentration of 20 ng/mL. The list of the standards used for QC is presented in Table SI-1. Caffeine-$^{13}C_3$ was used as an internal standard for following positive mode acquisition, while ethyl paraben-$^{13}C_6$ was used for the negative ionization mode. After spiking, the QC samples were processed following the same way as the investigated samples (see 2.3 Sample preparation).

The standards were monitored in real-time during data acquisition to ensure the control over the performance of the analytical system. After data acquisition, mass errors, RT drifts and trends in areas of native standards were checked to assess the quality of data acquisition. Batch acquisition was considered satisfactory when RT of the spiked standards were within ±0.4 min, and parent ion mass accuracy within ±5 ppm. Areas of the standards were plotted and screened for any increasing or decreasing trends. Figure SI-5 presents areas of native standards across QC samples, while SI-QC_parameters.xlsx tabulates areas, mass accuracies and retention time drifts for each monitored standard. Internal standards were checked after acquisition in the final aligned feature list and any sample without corresponding feature with RT drifts and mass error within specified limits of 0.4 min and 5 ppm, respectively, was rejected. Areas of internal standards for each sample are plotted in Figure SI-6. Accompanying plotting of standards in QCs, IS area plot served as an additional measure of quality for data acquisition.

Procedural blanks were prepared by storing frozen LC-MS water for three months in the same criovials as those used for storing the actual samples. They were subsequently spiked with isotopically labelled internal standards at the same concentration as the actual urine samples, and processed according to the standard sample preparation procedure (see 2.3 Sample preparation) to eliminate analytical artifacts. The procedural blanks were prepared in triplicate.

2.5.1. Acquisition order

At the beginning of the sequence, four solvent blanks and five QC sample replicates were injected to check the starting system's performance and the integrity of the QC samples. Afterwards, procedural blanks were injected. Randomly ordered samples were analysed in batches. At the beginning of each batch, a solvent blank and a QC sample were injected. QC samples and solvent blanks were injected per every 15 investigated samples. Each QC sample replicate was injected no more than ten times and the order of the injected QC replicates was alternating to avoid non-random trends.

Solvent blanks were used to control sample carryover and were injected after every 15th sample.

2.6. Data processing

Raw data was first transformed from raw format to mzXML using MSConvert and analysed using MzMine 2.53 (Pluskal et al., 2010). The processing parameters were optimized to allow for the detection of the spiked non-labelled standards and are presented in SI-1. Processing thresholds to detect BoE were rather low, resulting in computationally demanding amounts of data. Samples that lacked the signal of the internal standard caffeine-$^{13}C_3$ (9% of samples) were excluded from further processing. Any features present in the procedural blanks were removed from dataset. Only features with acquired MS2 data were filtered out and further subjected to compound annotation.

2.7. Compound identification

2.7.1. Non-targeted

Identification procedure is presented in Fig. 1. After processing we submitted all acquired MS2 spectra in one joined.mgf file to Sirius-CSI: Finger-ID (Dührkop et al., 2015, 2013). Once candidates were generated, we selected only the candidates with 100% elemental formula match within the mass error of 6 ppm, and highest library matches (KEGG, ChEBI or PubChem) with the minimum matching score of 60%.

The prioritized list of candidates, which passed the above defined criteria was exported as.csv file and subsequently manually filtered to retain only features matching potential BoE. This was done based on reviewing origin and use of candidates by compound databases T3DB (Wishart et al., 2015), Exposome Explorer (Neveu et al., 2020) and Chemistry Dashboard (Williams et al., 2017). MS2 of each filtered match was then examined by manually explaining MS2. For at least 3 most abundant fragment ions chemically logical structures were generated at the maximum mass error of 7 ppm. If MS2 was successfully explained, the candidate identity was tentatively assigned. The NT identification data along with molecular formulae of proposed fragments is presented in the document SI-NT_spectral_data.xlsx. The MS2 spectra of the tentatively annotated compounds are presented in SI-4.

The same process was repeated with MoNA (Mass Bank of North America), mass repository with experimental MS2 data (https://mona.fiehnlab.ucdavis.edu/). The criteria for successful identification were the mass error of $\leq$5 ppm for parent ion and 7 ppm for fragments and cosine similarity score $\geq$0.8, with at least 5 fragment ions matching. In case when less than 5 fragment ions were available, then all of them had to match. Matching of spectra with less than 3 fragments in MS2 was regarded insignificant.

All putatively identified BoE were then submitted to Feature Based Molecular Networking (FBMN) (Wang et al., 2017), clustering features with similar fragmentation patterns together. Clustering features were then identified by manual interpretation of their MS2 spectra, i.e. by comparing MS2 spectra of an unknown feature with structurally assigned MS2 spectrum of previously identified compound.

2.7.2. Suspect screening

The SS data included parent ion masses from the inclusion list (Section 2.4) determined at HR and the corresponding MS2 data at unit resolution. This was a consequence of instrumental setup which enabled simultaneous execution of NTA and SS; while NT MS2 data was acquired by data dependant acquisition (DDA) in Orbitrap analyser, and was therefore HR, the MS2 of ions corresponding to masses in suspect list were fragmented and acquired in linear ion trap and were LR. The reference MS2 spectra were collected from MoNa, but since they were scarce, we complemented them by generating an *in-silico* MS2 spectral library using CFM-ID (Allen et al., 2014). Compared to matching with MS2 library in NTA the identification parameters were adjusted for SS: mass error of molecular error was required to be $\leq$ 6 ppm, cosine similarity score was required to be $\geq$ 0.8, with at least 5 fragments matching. If any of MS2 spectra consisted of less than 5 fragments, then matching of all fragments was required. Matching of spectra which consisted of less than 3 fragments in MS2 was regarded as insignificant.

2.8. Regional and temporal differences

OPLS-DA model was constructed based on putatively identified BoE to check for differences between 1) exposure of children in rural and urban regions and 2) difference between winter and summer sampling of rural population. For statistical analyses urine dilution was corrected using specific gravity adjustment (Suwazono et al., 2005), while compound loss during sample preparation and data acquisition was corrected by normalizing the data to area of labelled internal standard caffeine-$^{13}C_3$. Models were constructed using Simca 15.0.2 (Sartorius Stedim Data Analytics) on natural logarithm transformed and Pareto

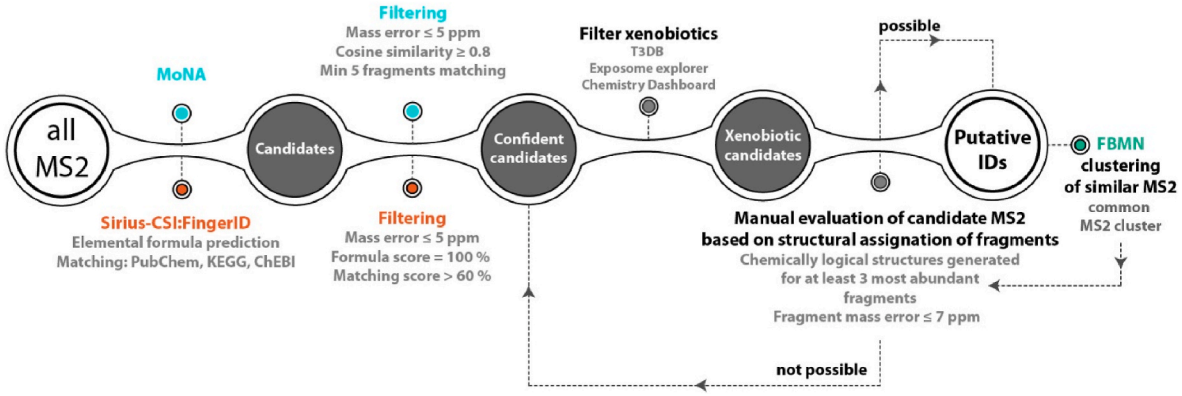


Fig. 1. Identification workflow for NTS.

scaled data (Figure SI-1 and Figure SI-2). Missing values were replaced by lowest non-zero value divided by two for each BoE. Number of components was determined using 7-fold cross-validation. Overfitting was checked with permutation test (N = 100, results are plotted in Figure SI-3 and Figure SI-4). Discriminatory features were determined using S-plot and Variable Importance for the Projection (VIP) score. Discriminatory features were checked using univariate non-parametric Mann-Whitney *U* test.

3. Results and discussion

3.1. Quality control

According to the described procedure, whose development and optimization are in detail discussed elsewhere (Tkalec et al., 2022) we aimed to characterize the exposures using UHPLC-HRMS in NT and SS mode. LC-MS data was acquired in both positive and negative mode. However, in the negative mode we did not observe labelled standard ethyl paraben-$^{13}C_6$, intended for monitoring quality of data acquisition in negative mode. This was attributed to inappropriately set data acquisition threshold. As the result, we rejected all the data acquired in the negative mode and hence continued only with the data acquired in the positive mode.

To ensure data quality, QC measures were held on several levels. Areas, mass accuracies and retention time drifts of nine native standards were checked after data acquisition (see Figure SI-5 and SI-QC_parameters.xlsx). No trends were observable in the areas of the native standards, except fluctuating trend which is the consequence of alternating order of injection of different QC samples. Due to the absence of a signal for phoxim, this compound was excluded from monitoring. Mass accuracies did not exceed 5 ppm, in fact for the majority of standards, the average mass error did not exceed 3 ppm, showing high mass accuracy of the acquired data. Retention times remained relatively constant throughout the run, with maximum retention time drift of 5% for the standard omethoate. Non-random trends in acquisition were monitored by observing areas of internal standard in each sample as presented in Figure SI-6, which shows no non-random trends.

To avoid erroneous conclusions about the exposure, we excluded any feature present in the solvent or procedural blanks from further consideration. Despite that certain analytical artifacts, for example phthalates, are potentially also the BoE, we avoided identifying parent compounds with possibly inconclusive sources.

No sample to sample carryover was observed from analysing solvent blanks which were injected following each QC within a batch.

3.2. Detection and identification of BoE in urine

Resulting data was processed using MzMine2.53 (Pluskal et al., 2010). Processing parameters were set to allow for detection of low abundance features, therefore very low thresholds for mass detection and subsequent steps were required. The criterion for the retention time (RT) drift was set according to the RT of standards spiked into QC samples (Tkalec et al., 2022). The maximum RT drift was for native standard omethoate, ± 0.31, min so accordingly maximum RT limit was set at ± 0.4 min. A too narrow RT limit could cause doubling of features, while a too high one can result in merging of non-identical features. Both events can negatively impact the identification ability, so this parameter should be set carefully. Still, we need to consider that the chromatographic separation, which is determined by the column stationary phase, mobile phases, elution gradient, etc., is rather a compromise than a universally ideal solution for any compound. This may in turn result in wider peaks and larger retention time drifts for some features and accordingly a compound might appear in final data matrix as two or more features and the resulting detection frequencies of BoE offer only a mere approximation.

Along with the stability of RT, the stability of instrument's mass accuracy was precisely followed (Tkalec et al., 2022), through monitoring of the drifts of standards' mass accuracies in the QC samples. The maximum mass error was determined for the protonated molecule of the standard ketoprofen at +3.53 ppm, demonstrating that errors were generally low. Nevertheless, to be on the safe side, we set the permitted mass error for the MS1 level at 5 ppm.

As we filtered out only the features with corresponding MS2 data, the final data matrix contained 34 015 features. This number is reasonably high, and arises from low thresholds of data acquisition and processing parameters thus maximizing the probability of low-level BoE detection. From this data matrix we tentatively identified 74 BoE in total, out of which 36 were found using NT analysis (Table 1) and 38 using SS (Table 2). As a result of identification workflow through mass spectral library matching or *in-silico* identification connected to the structural assignation of their MS2 spectra, BoE identified with NTA were at the Schymanski confidence level of 2 (Schymanski et al., 2014). Out of these, 16 were identified using Sirius-CSI:FingerID, 10 through mass spectral database MoNA, and another 10 by FBMN. The latter revealed the identification of two compound clusters, the penicillins and the polyethylene glycol cluster. Here, the primary BoE were phenoxymethyl penicillin (MH+ 351.1004) and decaethylene glycol (MH+ 459.2800), which were identified using Sirius-CSI:FingerID, and served as the basis for identification of other congeners within molecular networks. Assigned MS2 spectra with identification data and corresponding identification scores are presented in SI-NT_spectral_data.xslx.

The NT identification protocol considered only the highest scoring matches. Where the best match could not be confirmed based on the above stated criteria (see 2.7.1 Non-target screening), the candidate was rejected, with the next highest-scoring candidate not being considered anymore. This is a drawback of the presented workflow since it increases

Table 1
BoE identified using NT analysis. Presented are their elemental formulas, mass errors, RT, detection frequencies and classification.

Identity	Metabolite of	Biomarker group	Elemental formula	Theoretical mass $[M+H]^+$	Experimental mass $[M+H]^+$	Mass error (ppm)	RT (min)	Detection frequency (%)	ID level
3-Hydroxycotinine	Nicotine (Raja, 2016)	Biomarker of smoking	$C_{10}H_{12}N_2O_2$	193.0977	193.0972	2.589	2.36	24	2
Nonaethylene glycol	/	PCP	$C_{18}H_{38}O_{10}$	415.2538	415.2529	2.167	4.7	13	2
Undecaethylene glycol	/	PCP	$C_{22}H_{46}O_{12}$	503.3062	503.3091	−5.762	4.71	86	2
Decaethylene glycol	/	PCP	$C_{20}H_{42}O_{11}$	459.28	459.2821	−4.572	5.21	20	2
Dodecaethylene glycol	/	PCP	$C_{24}HO_1$	547.3324	547.3307	3.106	5.56	68	2
Tetradecaethylene glycol	/	PCP	$C_{28}H_{58}O_{15}$[c]	652.4152	652.4136	2.452	5.69	53	2
Pentaethylene glycol decyl ether	/	PCP	$C_{20}H_{42}O_6$	379.3054	379.3062	−2.109	20.15	24	2
Tetraethylene glycol decyl ether	/	PCP	$C_{18}H_{38}O_5$	335.2792	335.2794	−0.597	20.44	10	2
DEET	/	PCP, repellent	$C_{12}H_{17}N$	192.1383	192.1383	0.000	11.11	1	2
Icaridin	/	PCP, repellent	$C_{12}H_{23}NO_3$	230.1751	230.1756	−2.172	13.06	2	2
4-Hydroxybenzophenone	/	PCP, UV-filter	$C_{13}H_{10}O_2$	199.076	199.0758	1.005	10.61	33	2
Octabenzone	/	PCP, UV-filter	$C_{21}H_{26}O_3$	327.1955	327.1952	0.917	27.91	1	2
N-(2,6-dimethylphenyl)-2-hydroxyacetamide	Metalaxyl M (EPA, 1988)	Pesticide, Fungicide	$C_{10}H_{13}NO_2$	180.1019	180.1024	−2.776	2.94	31	2
N-(2,4-dimethylphenyl) formamide	Amitraz (Lazarus et al., 2021)	Pesticide, Fungicide	$C_9H_{11}NO$	150.0913	150.0918	−3.331	5.68	55	2
Naphthoxyacetic acid	/	PGR	$C_{12}H_{10}O_3$	203.0703	203.0707	−1.97	5.81	14	2
Trinexapac	/	PGR	$C_{11}H_{12}O_5$	225.0757	225.0763	−2.666	6.32	16	2
Phenoxymethyl penicillin/ Penicillin V	/	Pharmaceutical	$C_{16}H_{18}N_2O_5S$	351.1009	351.1004	1.424	6.18	1	2
Phenoxymethyl penicilloyl	Penicillin V (CHEBI:53 703)[a]	Pharmaceutical	$C_{16}H_{20}N_2O_6S$	369.1115	369.1116	−0.271	4.62	10	2
N-methylphenoxymethyl penicilloyl	Penicillin V[b]	Pharmaceutical	$C_{17}H_{22}N_2O_6S$	383.1271	383.1282	−2.871	5.14	2	2
Amoxycilloyl	Amoxycillin (CHEBI:53 705)[a]	Pharmaceutical	$C_{17}H_{23}N_3O_5S$	382.1431	382.1434	−0.785	5.83	11	2
Levetiracetam	/	Pharmaceutical	$C_8H_{14}N_2O$	171.1128	171.1133	−2.922	2.79	7	2
N-acetylphenoxymethyl penicilloyl	Penicillin V[b]	Pharmaceutical	$C_{18}H_{22}N_2O_7S$	411.122	411.1234	−3.405	4.49	11	2
Carbamazepine epoxide	Carbamazepine (Potter and Donnelly, 1998)	Pharmaceutical	$C_{15}H_{12}N_2O_2$	253.0972	253.0974	−0.790	5.13	10	2
Aminophenol	Paracetamol (Athersuch et al., 2018)	Pharmaceutical	C_6H_7NO	110.06	110.0604	−3.634	5.95	29	2
Lauramide DEA	/	Surfactant	$C_{16}H_{33}NO_3$	288.2533	288.2536	−1.041	18.82	11	2
Octylphenol	/	Surfactant byproduct	$C_{14}H_{22}O$	207.1743	207.1744	−0.483	19.7	2	2
Propenylaniline	/	Unknown use	$C_9H_{11}N$	134.0964	134.0968	−2.983	5.5	83	2
6-Phenylpicoline	/	Unknown use	$C_{12}H_{11}N$	170.0964	170.0968	−2.352	5.7	6	2
Cycloheptylamine	/	Unknown use	$C_7H_{15}N$	114.1277	114.1281	−3.505	8.28	5	2
Chloroisoquinoline	/	Unknown use	C_9H_6NCl	164.0262	164.0266	−2.439	8.56	2	2
Dicyclohexyl urea	/	Unknown use	$C_{13}H_{24}N_2O$	225.1961	225.1965	−1.776	12.26	1	2
Isoquinoline	/	Unknown use	C_9H_7N	130.0651	130.0654	−2.307	4.74	90	2
Methylacridine	/	Unknown use	$C_{14}H_{11}N$	194.0964	194.0968	−2.061	9.81	1	2
Naphthylamine	/	VOC	$C_{10}H_{90}N$	144.0808	144.0812	−2.776	5.93	15	2
Cresol	/	VOC	C_7H_8O	109.0648	109.0652	−3.668	4.31	41	2
Benzaldehyde	/	VOC	C_7H_6O	107.0498	107.0496	1.868	4.43	65	2

[a] CHEBI database entry (https://www.ebi.ac.uk/chebi/init.do, accessed June 2021), numerical value specifies CHEBI entry.
[b] Proposed identity based on mass spectrum, not available in mass spectral libraries.
[c] Only NH_4^+ adduct was found.

the loss of potential matches and the number of false negatives, however it enabled a more efficient and less time-consuming screening. We believe that many more BoE could be detected by considering, for example, top three or top five best-scoring candidate matches, which remains to be the future potential for identification of BoE from given data matrix.

By using SS approach 11 BoE were identified through *in-silico* generated library employing CFM-ID (Allen et al., 2014) and one compound by comparison with the library spectrum from MoNA. The remaining 26 compounds were identified by manual interpretation of MS2 of potential suspect list candidates' MS2 spectra. IUPAC names with SMILES, InCHI, InCHI-Keys, matching scores, where applicable and characteristic ions are presented in SI-SS_data.xlsx. Based on the identification pathway the BoEs obtained *via* SS approach were identified at Schymanski confidence level of 3 (Schymanski et al., 2014). A comprehensive SS list with 801 entries covering potential BoE such as pesticides, industrial chemicals, personal care product ingredients, and their 1st phase metabolites generated a large number of hits. These hits were subsequently individually checked, which was timewise another bottleneck in data processing. Other SS studies employed even more extensive SS lists, for example with 1158 (Caballero-Casero et al., 2021) and 1450 (Chen et al., 2021) entries. To reduce the identification time and workload required for SS, a compromise between the number of entries and subsequently large number of hits and time-efficacy could be made via by prioritizing BoE according to probability of exposure. For example, one could include a larger number of pesticide-related entries in a study of rural population, while reducing the number of traffic-related ones, and the opposite in case of an urban population

Table 2
Compounds identified using SS approach. Presented are their elemental formulas, mass errors, RT, detection frequencies and mode of identification.

Identity	Metabolite of	Group	Elemental formula	Theoretical mass $[M+H]^+$	Experimental mass $[M+H]^+$	Mass error (ppm)	Compound_RT	Detection frequency (%)	ID level
Cotinine	Nicotine (Raja, 2016)	Biomarker of smoking	$C_{10}H_{12}N_2O$	177.1028	177.1028	−0.068	2.83	22	3
Caffeine	/	Biomarker of soda, tea	$C_8H_{10}N_4O_2$	195.0882	195.0882	0.005	2.36	98	3
Celestolide	/	PCP, fragrance	$C_{17}H_{24}O$	245.1905	245.1902	1.387	8.55	6	3
1-Methyl-alpha-ionone	/	PCP, fragrance	$C_{14}H_{22}O$	207.1749	207.1744	2.365	19.7	2	3
Ethylparaben	/	PCP, preservative	$C_9H_{10}O_3$	167.0708	167.0706	1.311	6.05	17	3
Butylparaben	/	PCP, preservative	$C_{11}H_{14}O_3$	195.1021	195.102	0.61	6.08	6	3
Dioxybenzone	/	PCP, UV-filter	$C_{14}H_{12}O_4$	245.0814	245.0811	1.159	7.53	4	3
Terbuthylazine	/	Pesticide, Algicide	$C_9H_{16}ClN_5$	230.1172	230.1173	−0.226	8.2	3	3
Pyrimethanil	/	Pesticide, Fungicide	$C_{12}H_{13}N_3$	200.1188	200.1186	0.859	6.37	4	3
Tebuconazole	/	Pesticide, Fungicide	$C_{16}H_{22}ClN_3O$	308.153	308.1535	−1.736	6.37	14	3
4-[(4,6-dimethylpyrimidin-2-yl) amino]phenol	Pyrimethanil (Faniband et al., 2019)	Pesticide, Fungicide	$C_{12}H_{13}N_3O$	216.1137	216.114	−1.448	8.27	11	3
4-Hydroxychlorpropham	Chlorpropham (Carrera et al., 1998)	Pesticide, Herbicide	$C_{10}H_{12}ClNO_3$	230.0584	230.0571	5.633	2.34	18	3
Metholachlor	/	Pesticide, Herbicide	$C_{15}H_{22}ClNO_2$	284.1417	284.1401	5.744	4.93	10	3
Desisopropyl atrazine	Atrazine (Joo et al., 2010)	Pesticide, Herbicide	$C_5H_8ClN_5$	174.0546	174.0554	−4.32	4.29	55	3
Desethyl atrazine	Atrazine (Joo et al., 2010)	Pesticide, Herbicide	$C_6H_{10}ClN_5$	188.0703	188.0709	−3.201	8.98	2	3
2-Isopropyl-6-methyl-pyrimidin-4-ol	Diazinon (Shemer and Linden, 2006)	Pesticide, Insecticide	$C_8H_{12}N_2O$	153.1028	153.1025	2.123	6.22	8	3
Prohexadione	/	PGR	$C_{10}H_{12}O_5$	213.0763	213.0765	−0.948	2.46	33	3
Bisphenol G	/	Plasticizer, bisphenol	$C_{11}H_{12}O_5$	313.2168	313.2169	−0.463	5.91	9	3
Bisphenol F	/	Plasticizer, bisphenol	$C_{21}H_{28}O_2$	201.0916	201.0911	2.263	7.77	6	3
Bisphenol A	/	Plasticizer, bisphenol	$C_{13}H_{12}O_2$	229.1229	229.1231	−1.069	8.23	2	3
Monobenzyl phthalate	BBP (Huang et al., 2021)	Plasticizer, phthalate	$C_{15}H_{16}O_2$	257.0814	257.0812	0.716	8.34	48	3
Monobutyl phthalate	DBP (Huang et al., 2021)	Plasticizer, phthalate	$C_{15}H_{12}O_4$	223.0965	223.0949	7.172	5.06	4	3
Monohydroxybutyl phthalate	DBP, BBP (Huang et al., 2021)	Plasticizer, phthalate	$C_{12}H_{14}O_4$	239.0919	239.092	−0.314	6.63	11	3
Monocyclohexyl phthalate	DCHP (Huang et al., 2021)	Plasticizer, phthalate	$C_{12}H_{14}O_5$	249.1127	249.1126	0.337	6.72	15	3
Mono-2-ethyl-5-hydroxyhexyl phthalate	DEHP (Huang et al., 2021)	Plasticizer, phthalate	$C_{14}H_{16}O_4$	295.1545	295.1551	−1.867	7.97	35	3
Mono-2-ethyl-5-oxohexyl phthalate	DEHP (Huang et al., 2021)	Plasticizer, phthalate	$C_{16}H_{22}O_5$	293.1389	293.1393	−1.368	8.44	24	3
Monoethylhexyl phthalate	DEHP (Huang et al., 2021)	Plasticizer, phthalate	$C_{16}H_{20}O_5$	279.1591	279.1581	3.582	9.67	14	3
Monoethyl phthalate	DEP (Huang et al., 2021)	Plasticizer, phthalate	$C_{16}H_{22}O_4$	195.0657	195.0658	−0.338	7.76	5	3
Monooxoisodecyl phthalate	DIDP (Huang et al., 2021)	Plasticizer, phthalate	$C_{10}H_{10}O_4$	321.1702	321.1701	0.308	6.34	6	3
Monocarboxyisodecyl phthalate	DIDP (Huang et al., 2021)	Plasticizer, phthalate	$C_{18}H_{24}O_5$	337.1651	337.1638	3.896	6.5	13	3
Monohydroxyisodecyl phthalate	DIDP (Huang et al., 2021)	Plasticizer, phthalate	$C_{18}H_{25}O_6$	323.1858	323.1858	0.151	8.12	6	3
Monoisononyl phthalate	DINP (Huang et al., 2021)	Plasticizer, phthalate	$C_{18}H_{26}O_5$	293.1753	293.1756	−1.078	7.17	6	3
Monohydroxyisononyl phthalate	DINP (Huang et al., 2021)	Plasticizer, phthalate	$C_{17}H_{24}O_4$	309.1702	309.1699	0.967	7.98	7	3
Monooxoisononyl phthalate	DINP (Huang et al., 2021)	Plasticizer, phthalate	$C_{17}H_{24}O_5$	307.1545	307.1541	1.461	8.06	8	3
Monocarboxyisooctyl phthalate	DIOP, DINP (Huang et al., 2021)	Plasticizer, phthalate	$C_{17}H_{22}O_6$	323.1494	323.1501	−2.089	7.21	11	3

(*continued on next page*)

Table 2 (*continued*)

Identity	Metabolite of	Group	Elemental formula	Theoretical mass $[M+H]^+$	Experimental mass $[M+H]^+$	Mass error (ppm)	Compound_RT	Detection frequency (%)	ID level
Monooctyl phthalate	DOP (Huang et al., 2021)	Plasticizer, phthalate	$C_{16}H_{22}O_4$	279.1591	279.1578	4.657	16.19	1	3
Monopentyl phthalate	DPP (Huang et al., 2021)	Plasticizer, phthalate	$C_{13}H_{16}O_4$	237.1127	237.1127	−0.067	6.39	29	3

monitoring. This would reduce the number of hits, however it might lead to a significant bias possibly arising from multiple sources of certain compounds, transit and relocation of people and dietary differences.

The use of a tribrid instrument, which combines quadrupole, ion trap and Orbitrap mass analysers, enabled us to perform SS and NT screening simultaneously. Hence, for each scan, the detected *m/z* matching that of any exact mass of protonated BoE in the suspect list (within the predefined mass error of 5 ppm) was automatically fragmented in linear ion trap, while the Orbitrap analyser was simultaneously operated in DDA mode, fragmenting 10 most abundant ions. Therefore, MS2 spectra of compounds matching *m/z* values of those in the inclusion list are recorded at low resolution, whilst MS2 of other features are recorded at HR.

This process ensured the acquisition of MS2 for very low abundance BoEs from the suspect list, that would otherwise not be fragmented in a typical DDA mode. However, the main drawback of such acquisition is that the MS2 spectra of suspect hits are acquired at low-resolution. Likewise, since any hit from the SS is automatically excluded from the Orbitrap fragmentation and since the identification protocols use different thresholds and approaches there can be no overlap between the SS and NTS results.

Low-abundance compounds are hard to detect in complex matrices such as urine and even harder to identify due to sometimes insufficient quality of acquired MS2 spectra. Enzymatic hydrolysis enabled the cleavage of conjugate bonds, thus forming parent BoE or their 1st stage metabolites, which greatly benefitted their identification. First of all, deconjugation increases the concentration of BoE or its 1st stage metabolite, since different glucuronide or sulphate metabolites are transformed back into a single species, and with that the chance of acquiring high quality MS2 is increased. Furthermore, MS data on conjugated species are very scarcely included in databases and MS libraries resulting in the failure of identification. Third, conjugated species are inherently more prone to ESI negative ionization. On the contrary, parent BoE or their 1st stage metabolites, when ionised under electrospray positive conditions, yield more abundant and richer MS2 spectra, providing more structural information to enable successful identification. Furthermore, conjugated species and with that their MS2 spectra have increased molecular complexity, which in turn detrimentally affects identification through *in-silico* approaches or manual assignment of MS2 spectra. On the other hand, deconjugation step as a part of sample preparation leads to missing information on conjugated metabolites, which are crucially important in the toxicological research. Hence, ideally deconjugated and non-deconjugated samples should be analysed in parallel, yet this would significantly prolong the analysis time. Another concern is also that enzymatic deconjugation by glucuronidase/arylsulfatase may via non-specific activity produce new chemical species (Blount et al., 2000).

Sample preparation protocol involved solid-phase extraction and although broad polarity-range sorbent was used it is highly likely that very polar compounds have not been extracted. While this led to less chromatographic overlapping at low retention times, a chromatographic region which is inherently rich due to polar nature of urine, the drawback was the loss of information on very polar BoEs thus focusing medium to low polarity compounds.

In the applied identification approach, each feature was submitted for identification and only hits at the high degree of matching were retained. The high probability hits were screened for potential BoE. Due to that, this approach may have missed a large number of trace-level BoE, which were detected a low quality MS2. Thus the data matrix still holds a large amount of un-retrieved information and potentially much higher number of BoE. Furthermore, the identification procedure produced large amount of other tentatively identified compounds, as for example regarding dietary chemicals and their metabolites which were out of scope of this paper.

NTS/SS in this study were used as an initial screening operable within HBM to derive first overview of the populations' exposure. The compounds were therefore identified at confidence levels 2 and 3, and should be in future confirmed by reference standards.

3.3. Implications for the exposure of children

Compounds found by NT and SS are identified tentatively, acting as a basis for subsequent targeted studies to confirm the exposure and accurately quantify the exposure and explore exposure pattern in detail. However, within the scope of certainty of NT and SS, certain conclusions on the exposure can be implied.

Tentatively identified BoE indicate that children are exposed to a wide set of organic contaminants that belong to several compound classes (Fig. 2).

BoE to 11 pesticides were identified. We found the evidence of exposure to fungicides pyrimethanil, metalaxyl and tebuconazole, algicide terbuthylazine, herbicides metholachlor and chlorpropham. The latter was previously detected by similar SS approach in urine of pregnant women (Bonvallot et al., 2021). Further, we found BoE to several plant growth regulators (PGR), prohexadione, naphthyloxyacetic acid and trinexapac-ethyl. All of the above are EU approved products, however we identified BoE to three restricted pesticides. Atrazine, amitraz and diazinon were restricted from use in EU, atrazine and amitraz in 2004 (documents 2004/248/EC and 2004/141/EC), and diazinon in 2007 (document 2007/393), however the data here indicates continued exposure.

Personal care products (PCPs) are used on a daily basis and involve a number of different compound groups, many of which are continuously

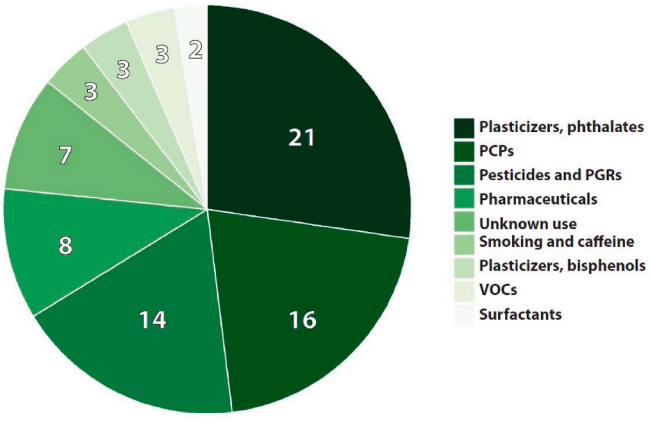


Fig. 2. Pie chart illustrating the main groups of identified BoEs with number of BoE in each group.

monitored in established HBM schemes. BoE to 16 PCPs were identified in the urine of children, including parabens, UV-filters, fragrances, repellents, surfactants and polyethylene glycols and polyethylene glycol ethers. Among parabens we identified two, ethyl and butyl paraben. This is in agreement with the results of previously published targeted analysis on the same population, where we quantified the levels of several parabens, among others also these two compounds (Tkalec et al., 2021). Parabens had also been previously identified by similar NT/SS approaches in urine (Caballero-Casero et al., 2021) and serum (Chen et al., 2021), as well as by a large number of targeted studies (Wei et al., 2021). Among the remaining compounds, the most commonly detected one was undecaethylene glycol (86%), a congener of polyethylene glycol (PEG) and polyethylene glycol ethers (PEGE). This class of compounds is used for various applications, particularly in cosmetics as emulsifiers, thickeners, humectants, cleansing agents and non-ionic surfactants, as excipients in pharmaceutical formulations and others (Fruijtier-Pölloth, 2005). Even though they are widely present in everyday life, there is scarcely any study considering human exposure to PEGs and PEGEs. The same is true for another tentatively identified compound, lauramide DEA, a surfactant widely used in cosmetics and other PCPs like shampoos and soaps (Mathews et al., 1996).

Similar to PCPs, a substantial group of compounds, which have been widely monitored using HBM are plasticizers and plastic-related compounds. Bisphenols are used in production of polycarbonate plastic and epoxy resins. In the samples of Slovenian children, we found BoE to BPA, and its alternatives BPF and BPG. Again, this is in agreement with the results of targeted analysis of same samples, where BPA and BPF were quantified in the majority of samples (Tkalec et al., 2021). Conversely, BPG was not analysed as a part of targeted analysis, and also the exposure to this compound has not yet been widely studied in literature. Thus, the current study demonstrates the capacity of SS/NT approach to reveal the exposure of children to BPG, and further suggests its inclusion in the existing HBM schemes to determine its distribution and levels in a wider population.

By far most detected compound class were BoEs to phthalates, out of which, monoisononyl phthalate was most commonly detected (93%). As for parabens, phthalates have likewise been largely monitored and quantified in many targeted studies (Eales et al., 2022). They were also identified by similar SS studies in urine, where, hydroxylated monoisononylphthalate was detected at even higher frequency (100%) (Caballero-Casero et al., 2021), and in serum, where the most commonly detected BoE was monodecyl phthalate (Chen et al., 2021). Both of them are members of high-molecular weight phthalates. Some of the phthalates, as for example isooctyl and nonyl phthalate are isomers. Isomers are particularly difficult to identify. In certain cases, structural and positional isomers will produce slightly different MS2 spectra due to producing different fragments, as the molecule cleaves differently. When structurally assigning fragments, one MS2 might be more chemically logical for a particular isomer. In case of other isomers, such as geometric isomers and especially stereoisomers, they cannot be differentiated by MS. Only a targeted method using standards and very well-designed separation method could unambiguously differentiate between such isomers and confidently identify the compounds.

Furthermore, we identified several volatile organic contaminants (VOCs), such as cresol, benzaldehyde and naphthylamine and BoEs to tobacco smoke, cotinine and 3-hydroxycotinine. All of these compounds can be linked to exposure to cigarette smoke, whereas cresol is released by automobile exhaust, and is found in air of areas of high traffic and vicinity of gas stations (Risne and Cash, 1990). Naphthylamine is used in various industrial applications, however it has been detected as a product of incomplete combustion in the cigarette smoke (Niu et al., 2018; Yu et al., 2014).

Using NTS we identified BoEs to several pharmaceuticals. Pharmaceuticals are being used intentionally to treat or prevent chronic or acute diseases, or to weaken their symptoms. While an unintentional exposure to pharmaceuticals is possible through contaminated food or water, their doses are significantly lower and probably not detected by SS/NTS. Hence, pharmaceuticals play a rather special role in the exposure analysis, however we still considered them for several reasons. First of all, they served for the confirmation of the applied workflow. During sampling, participants answered an extensive questionnaire involving medication during time of sampling. This data was used *post-hoc* to connect self-reported data to the identified biomarkers and with that confirm the analytical workflow of the study. Second, due to high level of biological activity, pharmaceuticals influence a large array of metabolic pathways, which might be of great significance when considering health implications of exposures. We identified phenoxymethyl penicillin (Penicillin V) and its metabolites, phenoxymethyl penniciloyl, *N*-methyl and *N*-acetyl phenoxymethyl penniciloyl. They were identified in the participants' samples, who reported to take phenoxymethyl penicillin to treat acute tonsillitis. Similarly, amoxycilloyl being a metabolite of amoxycillin, was identified in urine of the participants, who reported using this drug for the treatment of acute ear infection. Further, carbamazepine epoxide is a pharmacologically active metabolite of an anticonvulsant carbamazepine. Both, carbamazepine epoxide and levetiracetam, another antiepileptic drug, were identified in samples of the participants reporting being treated for epilepsy. Finally, we found aminophenol, a metabolite of paracetamol in the corresponding urine samples of those participants, who reported the use of this pharmaceutical.

Along with compounds with well-defined source, we tentatively identified propenyl aniline, phenyl picoline, cycloheptyl amine, methyl acridine, isoquinoline, chloroisoquinoline and dicyclohexyl urea, for which we were unable to retrieve conclusive information about their use or sources of exposure. Dicyclohexyl urea was, however, also previously detected and identified in human serum using similar non-targeted approach (Hall et al., 2012). To the authors' knowledge these compounds are for the first time reported in human urine. This urges for monitoring of these chemicals in a larger population, to describe the extent of exposure, sources, toxicological parameters, and to assess health risks connected to the exposure.

3.4. Statistical differences between populations

Specific differences between the presence of BoE at two locations and at two sampling times were investigated using OPLS-DA, where the models demonstrated that the exposure of children differed according to the location and time (Table SI-2). Even though detection frequencies were low for certain BoEs, model parameters show sufficient explanation and predictability, while avoiding overfitting, confirmed by permutation test.

Discriminatory BoE with VIP-score higher than 0.5 were further validated for significance using non-parametric univariate Mann-Whitney test. The results are presented in Table SI-3 and Table SI-4 and visualized in Fig. 3 and Fig. 4.

3.4.1. Rural vs urban

Results of statistical analysis, presented in Table SI-3 show the difference between relative abundances of contaminants among individuals from urban and rural regions. Differential BoE are presented in S-plot of OPLS-DA (Fig. 3) demonstrating a handful of BoE that were differently represented in rural and urban populations. BoE to pesticides atrazine and diazinon were elevated in rural region, which could indicate higher intensity of local agricultural activity, typical for the region. PGR prohexadione was, however, higher in urban population, which might indicate dietary exposure of imported food items. The levels of cresol, normally found in the contaminated air of high traffic areas, were expectedly higher in samples of the individuals residing in urban region.

Several other BoE followed location-specific pattern, however the differences were difficult to interpret due reasons such as a large array of applications, for example of surfactant lauramide DEA and cosmetic ingredient decaethylene glycol, or inconclusive sources such as phenyl

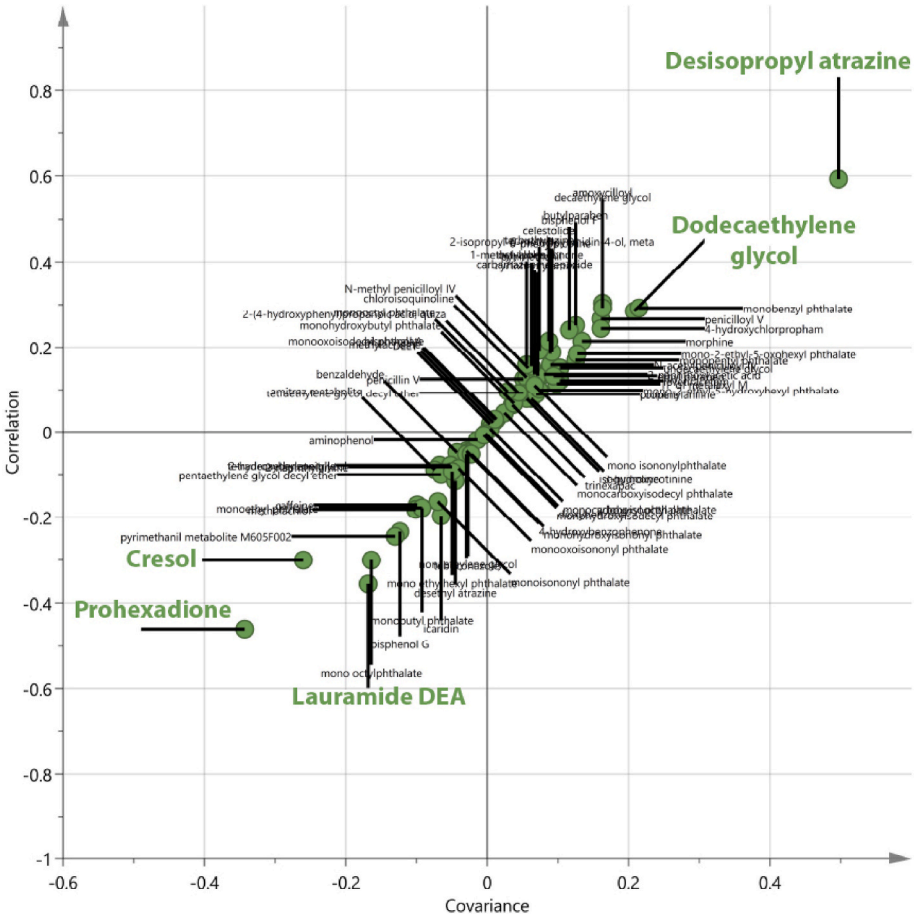


Fig. 3. S-plot (correlation vs covariance) visualizing discriminatory BoE between urban and rural samples. Significant features have higher non-zero covariance and correlation (top right and bottom left).

picoline, BPG and BPF.

3.4.2. *Winter vs summer*

Exposure is highly individual, some stressors can be present permanently, while other vary with season, changing of location and consumption of seasonal food. Based on the varying intensity of nearby agricultural activity, individuals from rural regions were sampled twice, once in winter/early spring 2019 and once during the summer of the same year. Discriminatory BoE for winter versus summer sampling are visualized in S-plot of OPLS-DA (Fig. 3), which shows that BoE to pesticides diazinon, terbuthylazine and pyrimethanil with PGR naphthoxyacetic acid were higher in the summertime samples (Tables SI–4). This coincides with higher agricultural activity in the region during summer season which is consistent with literature as proximity to fields, field area and spraying season have been identified as determinants of exposure to pesticides before (Teysseire et al., 2021). In contrast with what was expected the levels of BoE to UV-filter 4-hydroxybenzophenone were higher in winter. Seasonal differences of other compounds are, similarly to the regional differences, more difficult to discuss. Urinary BoE to organic contaminants are often short-lived, reflecting exposure that took place a few days before sampling at the most. Longitudinal samples would more accurately describe seasonal differences, however conclusions from spot samples might still reflect some differences in exposure of different groups.

4. Conclusions

Within the scope of this study the exposure of 300 urine samples from a cohort of Slovenian children aged 6–9 years (n = 200) were characterized using non-targeted analysis and suspect screening, which allowed tentative identification of a large number of xenobiotics and their BoE within one run. In contrast to metabolomics workflow, the thresholds for data processing of low-level biomarkers of exposure are required to be low, generating a large amount of noisy data, thus increasing the difficulty of confident compound identification. The latter remains the main bottleneck in applying NTA/SS in human biomonitoring.

In this study, 36 biomarkers of exposure were tentatively identified using non-targeted approach and another 38 using suspect screening. Biomarkers of exposure indicate environmental burden of children to several classes of chemicals such as personal care products, plasticizers and plastic production chemicals, volatile organic compounds, nicotine and caffeine, and pesticides, out of which three, atrazine, amitraz and diazinon were restricted in the EU due to their high toxicity. In addition, compounds not yet monitored in HBM schemes, such as bisphenol G, polyethylene glycols and polyethylene glycol ethers were tentatively identified. Alongside the compounds of known use, we tentatively identified 7 chemicals with unknown use, which might become in the future the chemicals of emerging concern and should be included in targeted HBM schemes in order to monitor their occurrence in a wider population. Due to sampling in two time periods and in two locations, we were able to demonstrate the transiency of the exposures and its location dependence. The results of the study show the complexity of the children's exposome, with them being exposed to many chemicals simultaneously. Overall, this work demonstrates the practical approach and emphasizes the potential of using non-targeted analysis and suspect screening in human biomonitoring.

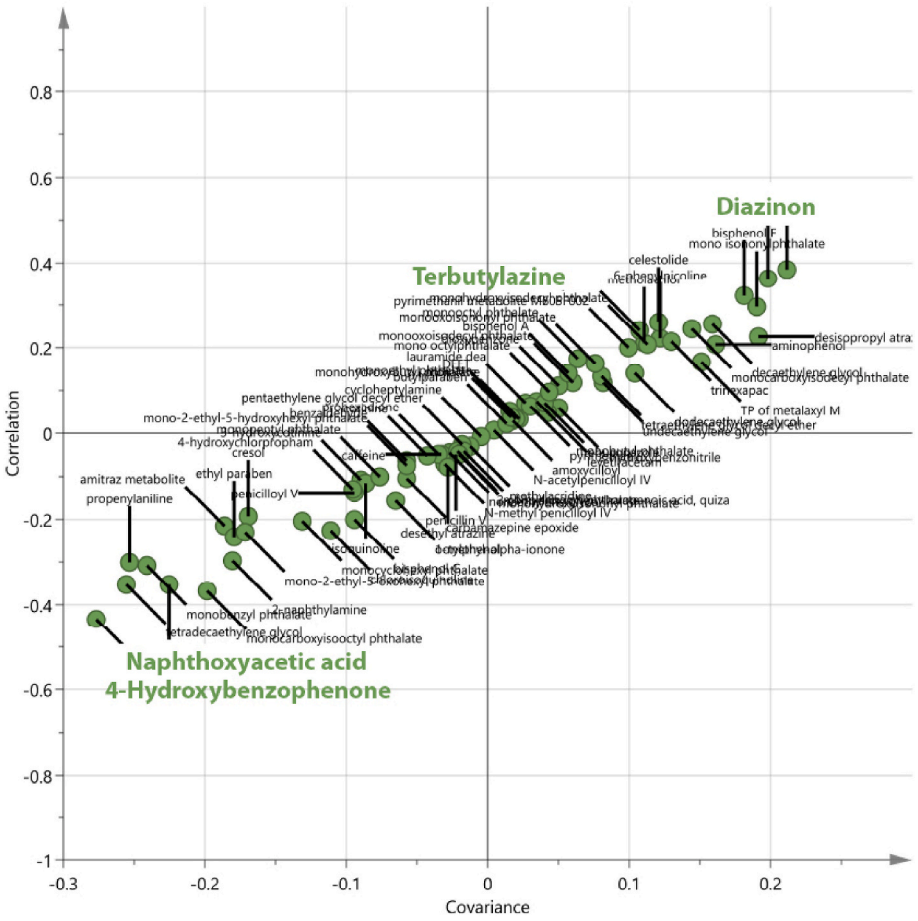


Fig. 4. S-plot (correlation vs covariance) visualizing discriminatory BoE between winter and summer sampling events. Significant features have higher non-zero covariance and correlation (top right and bottom left).

CRediT authorship contribution statement

Žiga Tkalec: Conceptualization, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Garry Codling:** Investigation, Methodology, Writing – review & editing. **Janja Snoj Tratnik:** Project administration, Writing – review & editing. **Darja Mazej:** Project administration, Writing – review & editing. **Jana Klánová:** Project administration, Funding acquisition. **Milena Horvat:** Project administration, Funding acquisition. **Tina Kosjek:** Conceptualization, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 733032 HBM4EU. The authors thank the Slovenian Research Agency - ARRS Program group P1-0143 and CRP V3-1722 funded by Chemicals Office of the Republic of Slovenia (CORS) and ARRS.

Part of the work was carried out with the support of RECETOX Research Infrastructure (ID LM2015051, MEYS CR, 2016–2019).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envpol.2022.120091.

References

Allen, F., Pon, A., Wilson, M., Greiner, R., Wishart, D., 2014. CFM-ID: a web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra. Nucleic Acids Res. 12, 94–99. https://doi.org/10.1093/nar/gku436.

Athersuch, T.J., Antoine, D.J., Boobis, A.R., Coen, M., Daly, A.K., Possamai, L., Nicholson, J.K., Wilson, I.D., 2018. Paracetamol metabolism, hepatotoxicity, biomarkers and therapeutic interventions: a perspective. Toxicol. Res. 7, 347–357. https://doi.org/10.1039/c7tx00340d.

Blount, B.C., Milgram, K.E., Silva, M.J., Malek, N.A., Reidy, J.A., Needham, L.L., Brock, J. W., 2000. Quantitative detection of eight phthalate metabolites in human urine using HPLC-APCI-MS/MS. Anal. Chem. 72, 4127–4134. https://doi.org/10.1021/ac000422r.

Bonvallot, N., Jamin, E.L., Regnaut, L., Chevrier, C., Martin, J.F., Mercier, F., Cordier, S., Cravedi, J.P., Debrauwer, L., Le Bot, B., 2021. Suspect screening and targeted analyses: two complementary approaches to characterize human exposure to pesticides. Sci. Total Environ. 786, 147499 https://doi.org/10.1016/j.scitotenv.2021.147499.

Caballero-Casero, N., Castro, G., Bastiaensen, M., Gys, C., van Larebeke, N., Schoeters, G., Covaci, A., 2021. Identification of chemicals of emerging concern in urine of Flemish adolescents using a new suspect screening workflow for LC-QTOF-MS. Chemosphere 280, 130683. https://doi.org/10.1016/j.chemosphere.2021.130683.

Carrera, G., Alary, J., Melgar, M.J., Lamboeuf, Y., Pipy, B., 1998. Metabolism and cytotoxicity of chlorpropham (CIPC) and its essential metabolites in isolated rat hepatocytes during a partial inhibition of sulphation and glucuronidation reactions: a comparative study. Arch. Environ. Contam. Toxicol. 35, 89–96. https://doi.org/10.1007/s002449900354.

Chen, Q., Zhang, Y., Ye, L., Gong, S., Sun, H., Su, G., 2021. Identifying active xenobiotics in humans by use of a suspect screening technique coupled with lipidomic analysis. Environ. Int. 157, 106844 https://doi.org/10.1016/j.envint.2021.106844.

Dührkop, K., Scheubert, K., Böcker, S., 2013. Molecular formula identification with SIRIUS. Metabolites 3, 506–516. https://doi.org/10.3390/metabo3020506.

Dührkop, K., Shen, H., Meusel, M., Rousu, J., Böcker, S., 2015. Searching molecular structure databases with tandem mass spectra using CSI:FingerID. Proc. Natl. Acad. Sci. USA 112, 12580–12585. https://doi.org/10.1073/pnas.1509788112.

Eales, J., Bethel, A., Galloway, T., Hopkinson, P., Morrissey, K., Short, R.E., Garside, R., 2022. Human health impacts of exposure to phthalate plasticizers: an overview of reviews. Environ. Int. 158, 106903 https://doi.org/10.1016/j.envint.2021.106903.

EPA, 1988. Metalaxyl Registration Standard.

Faniband, M., Ekman, E., Littorin, M., Maxe, M., Larsson, E., Lindh, C.H., 2019. Biomarkers of exposure to pyrimethanil after controlled human experiments. J. Anal. Toxicol. 43, 277–283. https://doi.org/10.1093/jat/bky091.

Ferguson, A., Penney, R., Solo-Gabriele, H., 2017. A review of the field on children's exposure to environmental contaminants: a risk assessment approach. Int. J. Environ. Res. Publ. Health 14, 1–25. https://doi.org/10.3390/ijerph14030265.

Fruijtier-Pölloth, C., 2005. Safety assessment on polyethylene glycols (PEGs) and their derivatives as used in cosmetic products. Toxicology 214, 1–38. https://doi.org/10.1016/j.tox.2005.06.001.

Ganzleben, C., Antignac, J.P., Barouki, R., Castaño, A., Fiddicke, U., Klánová, J., Lebret, E., Olea, N., Sarigiannis, D., Schoeters, G.R., Sepai, O., Tolonen, H., Kolossa-Gehring, M., 2017. Human biomonitoring as a tool to support chemicals regulation in the European Union. Int. J. Hyg Environ. Health 220, 94–97. https://doi.org/10.1016/j.ijheh.2017.01.007.

Getzinger, G.J., Ferguson, P.L., 2020. Illuminating the exposome with high-resolution accurate-mass mass spectrometry and nontargeted analysis. Curr. Opin. Environ. Sci. Heal. 15, 49–56. https://doi.org/10.1016/j.coesh.2020.05.005.

Hall, L.M., Hall, L.H., Kertesz, T.M., Hill, D.W., Sharp, T.R., Oblak, E.Z., Dong, Y.W., Wishart, D.S., Chen, M.H., Grant, D.F., 2012. Development of Ecom50 and retention index models for nontargeted metabolomics: identification of 1,3-dicyclohexylurea in human serum by HPLC/Mass spectrometry. J. Chem. Inf. Model. 52, 1222–1237. https://doi.org/10.1021/ci300092s.

Huang, S., Qi, Z., Ma, S., Li, G., Long, C., Yu, Y., 2021. A critical review on human internal exposure of phthalate metabolites and the associated health risks. Environ. Pollut. 279, 116941 https://doi.org/10.1016/j.envpol.2021.116941.

Joo, H., Choi, K., Hodgson, E., 2010. Human metabolism of atrazine. Pestic. Biochem. Physiol. 98, 73–79. https://doi.org/10.1016/j.pestbp.2010.05.002.

Lazarus, M., Tariba Lovaković, B., Orct, T., Sekovanić, A., Bilandžić, N., Đokić, M., Solomun Kolanović, B., Varenina, I., Jurič, A., Denžić Lugomer, M., Bubalo, D., 2021. Difference in pesticides, trace metal(loid)s and drug residues between certified organic and conventional honeys from Croatia. Chemosphere 266. https://doi.org/10.1016/j.chemosphere.2020.128954.

Ljoncheva, M., Stepišnik, T., Džeroski, S., Kosjek, T., 2020. Cheminformatics in MS-based environmental exposomics: current achievements and future directions. Trends Environ. Anal. Chem. 28, e00099 https://doi.org/10.1016/j.teac.2020.e00099.

Mathews, J.M., deCosta, K., Thomas, B.F., 1996. Lauramide diethanolamine absorption, metabolism, and disposition in rats and mice after oral, intravenous, and dermal administration. Drug Metab. Dispos. 24, 702–710.

Neveu, V., Moussy, A., Rouaix, H., Wedekind, R., Pon, A., Knox, C., Wishart, D.S., Scalbert, A., 2017. Exposome-Explorer: a manually-curated database on biomarkers of exposure to dietary and environmental factors. Nucleic Acids Res. 45, D979–D984. https://doi.org/10.1093/nar/gkw980.

Neveu, V., Nicolas, G., Salek, R.M., Wishart, D.S., Scalbert, A., 2020. Exposome-Explorer 2.0: an update incorporating candidate dietary biomarkers and dietary associations with cancer risk. Nucleic Acids Res. 48, D908–D912. https://doi.org/10.1093/nar/gkz1009.

Niedzwiecki, M.M., Walker, D.I., Vermeulen, R., Chadeau-Hyam, M., Jones, D.P., Miller, G.W., 2019. The exposome: molecules to populations. Annu. Rev. Pharmacol. Toxicol. 59, 107–127. https://doi.org/10.1146/annurev-pharmtox-010818-021315.

Niu, J., Zhao, X., Jin, Y., Yang, G., Li, Zhonghao, Wang, J., Zhao, R., Li, Zhongyue, 2018. Determination of aromatic amines in the urine of smokers using a porous organic framework (JUC-Z2)-coated solid-phase microextraction fiber. J. Chromatogr. A 1555, 37–44. https://doi.org/10.1016/j.chroma.2018.04.059.

Pluskal, T., Castillo, S., Villar-Briones, A., Orešič, M., 2010. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinf. 11, 1–11. https://doi.org/10.1186/1471-2105-11-395.

Potter, J.M., Donnelly, A., 1998. Carbamazepine-10,11-Epoxide in therapeutic drug monitoring. Ther. Drug Monit. 20.

Pourchet, M., Debrauwer, L., Klanova, J., Price, E.J., Covaci, A., Caballero-Casero, N., Oberacher, H., Lamoree, M., Damont, A., Fenaille, F., Vlaanderen, J., Meijer, J., Krauss, M., Sarigiannis, D., Barouki, R., Le Bizec, B., Antignac, J.P., 2020. Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: from promises to challenges and harmonisation issues. Environ. Int. 139, 105545 https://doi.org/10.1016/j.envint.2020.105545.

Raja, M., 2016. Diagnostic methods for detection of cotinine level in tobacco users: a review. J. Clin. Diagn. Res. 10, 4–6. https://doi.org/10.7860/jcdr/2016/17360.7423.

Risne, C.H., Cash, S.L., 1990. A high-performance liquid chromatographic determination of major phenolic compounds in tobacco smoke. J. Chromatogr. Sci. 28, 239–244. https://doi.org/10.1093/chromsci/28.5.239.

Schrimpe-Rutledge, A.C., Codreanu, S.G., Sherrod, S.D., McLean, J.A., 2016. Untargeted metabolomics strategies—challenges and emerging directions. J. Am. Soc. Mass Spectrom. 27, 1897–1905. https://doi.org/10.1007/s13361-016-1469-y.

Schymanski, E.L., Jeon, J., Gulde, R., Fenner, K., Ruff, M., Singer, H.P., Hollender, J., 2014. Identifying small molecules via high resolution mass spectrometry: communicating confidence. Environ. Sci. Technol. 48, 2097–2098. https://doi.org/10.1021/es5002105.

Shemer, H., Linden, K.G., 2006. Degradation and by-product formation of diazinon in water during UV and UV/H2O2 treatment. J. Hazard Mater. 136, 553–559. https://doi.org/10.1016/j.jhazmat.2005.12.028.

Stajnko, A., Šlejkovec, Z., Mazej, D., France-Štiglic, A., Briški, A.S., Prpić, I., Špirić, Z., Horvat, M., Falnoga, I., 2019. Arsenic metabolites; selenium; and AS3MT, MTHFR, AQP4, AQP9, SELENOP, INMT, and MT2A polymorphisms in Croatian-Slovenian population from PHIME-CROME study. Environ. Res. 170, 301–319. https://doi.org/10.1016/j.envres.2018.11.045.

Stajnko, A., Snoj Tratnik, J., Kosjek, T., Mazej, D., Jagodic, M., Eržen, I., Horvat, M., 2020. Seasonal glyphosate and AMPA levels in urine of children and adolescents living in rural regions of Northeastern Slovenia. Environ. Int. 143, 105985 https://doi.org/10.1016/j.envint.2020.105985.

Suwazono, Y., Åkesson, A., Alfvén, T., Järup, L., Vahter, M., 2005. Creatinine versus specific gravity-adjusted urinary cadmium concentrations. Biomarkers 10, 117–126. https://doi.org/10.1080/13547500500159001.

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

Tkalec, Ž., Codling, G., Klánová, J., Horvat, M., Kosjek, T., 2022. LC-HRMS based method for suspect/non-targeted screening for biomarkers of chemical exposure in human urine. Chemosphere 134550. https://doi.org/10.1016/j.chemosphere.2022.134550.

Tkalec, Ž., Kosjek, T., Snoj Tratnik, J., Stajnko, A., Runkel, A.A., Sykiotou, M., Mazej, D., Horvat, M., 2021. Exposure of Slovenian children and adolescents to bisphenols, parabens and triclosan: urinary levels, exposure patterns, determinants of exposure and susceptibility. Environ. Int. 146 https://doi.org/10.1016/j.envint.2020.106172.

Uppal, K., Walker, D.I., Liu, K., Li, S., Go, Y.M., Jones, D.P., 2016. Computational metabolomics: a framework for the million metabolome. Chem. Res. Toxicol. 29, 1956–1975. https://doi.org/10.1021/acs.chemrestox.6b00179.

Vinaixa, M., Schymanski, E.L., Neumann, S., Navarro, M., Salek, R.M., Yanes, O., 2016. Mass spectral databases for LC/MS- and GC/MS-based metabolomics: state of the field and future prospects. TrAC, Trends Anal. Chem. 78, 23–35. https://doi.org/10.1016/j.trac.2015.09.005, 2016.

Wang, M., Carver, J.J., Phelan, V.V., Sanchez, L.M., Garg, N., Peng, Y., Nguyen, D.D., Watrous, J., Kapono, C.A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A. V., Meehan, M.J., Liu, W.-T., Crüsemann, M., Boudreau, P.D., Esquenazi, E., Sandoval-Calderon, M., Kersten, R.D., Pace, L.A., Quinn, R.A., Duncan, K.R., Hsu, C.-C., Floros, D.J., Gavilan, R.G., Kleigrewe, K., Northen, T., Dutton, R.J., Parrot, D., Carlson, E.E., Aigle, B., Michelsen, C.F., Jelsbak, L., Sohlenkamp, C., Pevzner, P., Edlund, A., McLean, J., Piel, J., Murphy, B.T., Gerwick, L., Liaw, C.-C., Yang, Y.-L., Humpf, H.-U., Maansson, M., Keyzers, R.A., Sims, A.C., Johnson, A.R., Sidebottom, A.M., Sedio, B.E., Klitgaard, A., Larson, C.B., Boya P, C.A., TorresMendoza, D., Gonzalez, D.J., Silva, D.B., Marques, L.M., Demarque, D.P., Pociute, E., O'Neill, E.C., Briand, E., Helfrich, E.J.N., Granatosky, E.A., Glukhov, E., Ryffel, F., Houson, H., Mohimani, H., Kharbush, J.J., Zeng, Y., Vorholt, J.A., Kurita, K.L., Charusanti, P., McPhail, K.L., Nielsen, K.F., Vuong, L., Elfeki, M., Traxler, M.F., Engene, N., Koyama, N., Vining, O.B., Baric, R., Silva, R.R., Mascuch, S.J., Tomasi, S., Jenkins, S., Macherla, V., Hoffman, T., Agarwal, V., Williams, P.G., Dai, J., Neupane, R., Gurr, J., Rodríguez, A.M.C., Lamsa, A., Zhang, C., Dorrestein, K., Duggan, B.M., Almaliti, J., Allard, P.-M., Phapale, P., Nothias, L.-F., Alexandrov, T., Litaudon, M., Wolfender, J.-L., Kyle, J.E., Metz, T.O., Peryea, T., Nguyen, D.-T., VanLeer, D., Shinn, P., Jadhav, A., Müller, R., Waters, K. M., Shi, W., Liu, X., Zhang, L., Knight, R., Jensen, P.R., Palsson, B.Ø., Pogliano, K., Linington, R.G., Gutierrez, M., Lopes, N.P.', Gerwick, W.H., Moore, B.S., Dorrestein, P.C., Bandeira, N., 2017. Sharing and community curation of mass spectrometry data with GNPS. Nat. Biotechnol. 34, 828–837. https://doi.org/10.1038/nbt.3597.Sharing.

Wei, F., Mortimer, M., Cheng, H., Sang, N., Guo, L.H., 2021. Parabens as chemicals of emerging concern in the environment and humans: a review. Sci. Total Environ. 778, 146150 https://doi.org/10.1016/j.scitotenv.2021.146150.

Williams, A.J., Grulke, C.M., Edwards, J., McEachran, A.D., Mansouri, K., Baker, N.C., Patlewicz, G., Shah, I., Wambaugh, J.F., Judson, R.S., Richard, A.M., 2017. The CompTox Chemistry Dashboard: a community data resource for environmental chemistry. J. Cheminf. 9, 1–27. https://doi.org/10.1186/s13321-017-0247-6.

Wishart, D., Arndt, D., Pon, A., Sajed, T., Guo, A.C., Djoumbou, Y., Knox, C., Wilson, M., Liang, Y., Grant, J., Liu, Y., Goldansaz, S.A., Rappaport, S.M., 2015. T3DB: the toxic exposome database. Nucleic Acids Res. 43, D928–D934. https://doi.org/10.1093/nar/gku1004.

Yu, J., Wang, S., Zhao, G., Wang, B., Ding, L., Zhang, X., Xie, J., Xie, F., 2014. Determination of urinary aromatic amines in smokers and nonsmokers using a MIPs-SPE coupled with LC-MS/MS method. J. Chromatogr., B: Anal. Technol. Biomed. Life Sci. 958, 130–135. https://doi.org/10.1016/j.jchromb.2014.03.023.